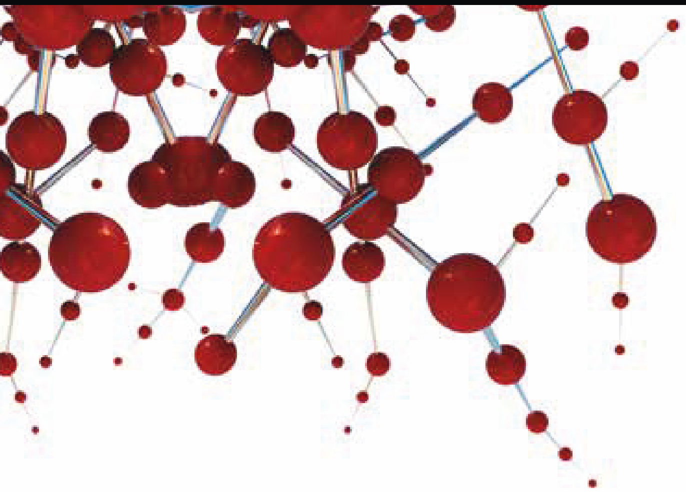


Zhenxing Li

Efficient and Accurate Hand-based Kinesthetic Interaction for Virtual Reality

Zhenxing Li

Efficient and Accurate Hand-based Kinesthetic Interaction for Virtual Reality

ACADEMIC DISSERTATION
To be presented with the permission of the Faculty of Information Technology and
Communication Sciences of Tampere University, for public discussion at Tampere
University on August 13th, 2021, at noon.

Faculty of Information Technology and Communication Sciences
Tampere University

Dissertations in Interactive Technology, Number 35
Tampere 2021

ACADEMIC DISSERTATION IN INTERACTIVE TECHNOLOGY

Supervisor:	Professor Roope Raisamo, Ph.D. Faculty of Information Technology and Communication Sciences Tampere University Finland
Opponent:	Senior Research Professor Nigel W. John, Ph.D. Department of Computer Science University of Chester The United Kingdom
Reviewers:	Professor William Harwin, Ph.D. School of Biological Sciences, Biomedical engineering University of Reading The United Kingdom
	Dr. Ken Pfeuffer, Ph.D. Research Institute CODE Bundeswehr University Munich Germany

ISBN 978-952-03-2052-2 (pdf)
http://urn.fi/URN:ISBN: 978-952-03-2052-2

Dissertations in Interactive Technology, Number 35

Faculty of Information Technology and Communication Sciences
FIN-33014 Tampere University
FINLAND

ISBN 978-952-03-2051-5 (print)
ISSN 1795-9489

PunaMusta Oy – Yliopistopaino
Tampere 2021

Abstract

Current virtual reality (VR) technology is mainly based on visual and auditive modalities. Kinesthetic interaction, as a form of human-computer interaction, allows users to apply touch-based hand behaviors to interact with virtual objects and simultaneously provides force to simulate the feeling of touch. Implementing such interaction may largely improve the immersion of VR and extend the range of its applications. There are many devices that can be used to implement kinesthetic interaction, such as wearable haptic gloves and grounded force-feedback devices. Among them, grounded force-feedback devices are more mature, which can provide a reliable kinesthetic interface with robust and realistic force feedback. However, there is a major challenge while using force-feedback devices for VR interaction, that is, their small device workspace cannot be directly used to explore a large virtual environment. The current solution to this challenge is to employ a large control-display (CD) gain. While holding the device arm, this technique scales a small hand motion to a large motion of the cursor in the virtual world and thus increases the device workspace.

The aim of this dissertation is to enable efficient and accurate hand-based kinesthetic interaction for VR. The research was divided into three steps: problem understanding, development and application. First, the research was focused on the technique of CD gain. Multiple studies have argued that the mismatch between the hand motion and the cursor motion caused by using the CD gain method may affect kinesthetic interaction. However, it is unclear how it affects kinesthetic tasks in terms of task performance and user experience. The present research filled this gap and examined the effects of CD gain based on a kinesthetic task. Second, to address the issue of limited workspace for force-feedback devices, three multimodal kinesthetic interfaces were developed by using the user's gaze as an input modality. These novel kinesthetic interfaces avoided the use of CD gain, and they also potentially relieved hand fatigue from prolonged operation. Third, the research explored medical applications of a kinesthetic VR interface and a vibrotactile VR interface associated with current popular VR equipment and force-feedback devices. To explore their practical usability, the two VR interfaces were compared with the state-of-the-art 2D interface (2D display + mouse) as the baseline.

The dissertation provides the following main contributions: First, the research experimentally demonstrated the effects of CD gain on kinesthetic interaction in terms of task performance and user experience.

Thus, it provided an empirical basis for designing new kinesthetic interfaces using the CD gain method. Second, the research explored a novel design space for kinesthetic interfaces by using eye gaze as a kinesthetic input modality. It also contributed to understanding human kinesthetic perception in the virtual environment and simultaneously identified the critical factors for designing high-quality kinesthetic interfaces in terms of kinesthetic perception accuracy. Third, the research revealed the strengths and weaknesses of the kinesthetic VR interface and the vibrotactile VR interface. Therefore, it provided an empirical understanding for developing efficient, accurate and user-friendly interactive VR systems. More importantly, the research demonstrated the potential of the kinesthetic VR interface to be the next-generation user interface in the field of medical diagnosis and planning, and it encouraged further research in this area.

Acknowledgements

First and foremost, I would like to express my deep gratitude to my supervisor, Professor Roope Raisamo. He guided me in the planning and execution of the research and provided an excellent research environment and resources. In addition, he always believed in me and gave me enough research independence so that I could try different study ideas and thus find my own research path for this dissertation.

I would like to take this opportunity to thank Professor Nigel W. John for agreeing to be the opponent for my dissertation defense and thank the pre-examiners Professor William Harwin and Dr. Ken Pfeuffer for their insightful comments.

The studies for this dissertation involved different prototype systems with multiple interactive techniques such as haptics and gaze tracking. I received much help from my colleagues in the Multimodal Interaction Research Group. I wish to thank all my co-authors. Deepak Akkil, the co-author of the first three publications, not only provided technical support for gaze tracking but also shared invaluable research experience with me. Maria Kiiveri and Jussi Rantala, the co-authors of the fourth publication, provided great assistance in conducting the study and completing the manuscript. Without your contributions, I could not have achieved my research goals and published our research in highly regarded scientific journals. I also want to thank my other colleagues, Jari Kangas, Ahmed Farooq and Tomi Nukarinen, for joining in the pilot studies and providing valuable feedback to improve the prototype system of each study. For all of you, again please accept my cordial thanks for your help in my research.

And finally, I would like to thank my family for your love and support. To my dear parents, thank you for encouraging me in my academic pursuit. Please forgive me for being away from you for such a long time. I was once confused and lost, and I could not see the way forward in my research. My loving wife, you are the one who is always there with me in those darkest nights. My adorable daughters, you are smart, diligent and independent so that I could have more time for my work. Words will never be enough to express just how thankful I am that I have you.

Tampere, July 1, 2021

Zhenxing Li

Contents

List of Publications

This dissertation is composed of a summary and the following original publications, reproduced here by permission.

I. Zhenxing Li, Deepak Akkil and Roope Raisamo (2020). The Impact of Control-Display Gain in Kinesthetic Search. In *Nisky I., Hartcher-O'Brien J., Wiertlewski M., Smeets J. (eds) Haptics: Science, Technology, Applications. EuroHaptics 2020. Lecture Notes in Computer Science*, vol 12272. Springer, Cham, Pages 158-166. https://doi.org/10.1007/978-3-030-58147-3_18 83

II. Zhenxing Li, Deepak Akkil and Roope Raisamo (2019). Gaze Augmented Hand-Based Kinesthetic Interaction: What You See is What You Feel. *IEEE Transactions on Haptics*, Volume 12, Issue 2, Pages 114-127. https://doi.org/10.1109/TOH.2019.2896027 95

III. Zhenxing Li, Deepak Akkil and Roope Raisamo (2020). Gaze-Based Kinaesthetic Interaction for Virtual Reality. *Interacting with Computers*, Volume 32, Issue 1, Pages 17–32. https://doi.org/10.1093/iwcomp/iwaa002 111

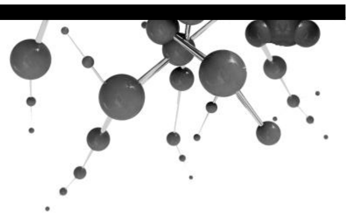
IV. Zhenxing Li, Maria Kiiveri, Jussi Rantala and Roope Raisamo (2021). Evaluation of Haptic Virtual Reality User Interfaces for Medical Marking on 3D Models. *International Journal of Human Computer Studies*, Volume 147, 102561. https://doi.org/10.1016/j.ijhcs.2020.102561 129

The Author's Contribution to the Publications

All publications reported in this dissertation were co-authored, and collaboration took place in the different stages of each work.

The present author was the main author in all publications. The contributions to these studies included formulating the goals of the research, developing methods and software, conducting experimental investigation, providing formal analysis of the data, verifying the achievement of research outputs, writing original drafts and revising manuscripts based on critical review and commentary.

For Studies I, II and III, Deepak Akkil helped develop research methods and provide technical support for eye tracking. For Study IV, Maria Kiiveri assisted the work in the development of software and the experimental investigation. Jussi Rantala was responsible for verifying the research outputs and revising the manuscript. For all studies, Professor Roope Raisamo supervised the research planning and execution. He provided study materials, financial support and valuable feedback, which made the publications reported in this dissertation possible.



1 Introduction

Virtual reality (VR) technology provides a computer-generated three-dimensional (3D) environment with which users can explore and interact. This technology has been used in many fields, such as entertainment, education (Radianti et al., 2019), e-commerce and manufacturing (Mujber et al., 2004). For example, in education, the implementation of interactive virtual environments to simulate various physical tasks brings great benefits for teaching and training. It can save time and cost, and more importantly avoid physical risks. Current interactive VR systems mainly focus on visual and auditive modalities. The sense of touch is one of the most important human interaction channels. Implementing realistic touch-based interaction in the virtual environment could largely improve the immersion of VR and potentially extend the range of VR applications. However, it is often missing or underused in current interactive VR systems.

In a traditional computer interface, virtual objects are modelled by a host computer and displayed on a 2D screen. A standard mouse and a keyboard are commonly used to interact with the objects. This traditional interface has been considered as a non-immersive VR system (Robertson et al., 1993) because of its unnatural interactive methods as well as the absence of some important senses like peripheral vision and touch. Technological advances have led to the development of VR equipment. Current popular VR devices are related mainly to the visual sense, such as HTC Vive[1], Oculus Rift[2] and Sony PlayStation VR[3]. These VR systems use a head-mounted display (HMD) as the visual output channel which can

---

[1] https://www.vive.com/eu/(accessed 1 July 2021)

[2] https://www.oculus.com/rift/(accessed 1 July 2021)

[3] https://www.playstation.com/en-gb/ps-vr/ (accessed 1 July 2021)

provide the user with flexible peripheral vision. In addition, they use VR motion controllers as the manipulation tools. These controllers allow the user's hand motions as the input modality and optionally provide vibration feedback to the hands, enabling tactile interaction.

In human-computer interaction (HCI), tactile interaction as a form of haptic interaction focuses on cutaneous sensations like vibration and pressure (El Saddik et al., 2011). However, when using tactile feedback such as the vibration from the motion controllers, it is difficult to implement realistic touch-based interaction in VR. For example, while people are pushing a wall, they feel a reaction force on their hands and arms. Tactile sensations only on the skins cannot recreate such interaction. Kinesthetic interaction is another branch of haptic interaction, focusing on movement-based sensations from the muscles, tendons and joints (El Saddik et al., 2011). Kinesthetic techniques can support realistic bidirectional touch behaviors. They allow the user's hand motions to be the kinesthetic input and provide force feedback as the kinesthetic output to simulate the feeling of touch.

Multiple devices can be used to implement kinesthetic interaction in a virtual environment. They include wearable haptic gloves (e.g., CyberGlove[4] and HaptX[5]), ungrounded kinesthetic pens (e.g., Kamuro et al., 2011) and grounded force-feedback devices (e.g., Geomagic Touch[6] and Novint Falcon[7]). Compared with other kinesthetic devices, grounded force-feedback devices are more mature. They typically have a mechanical arm with six degrees of freedom. Holding the arm allows the user's natural hand motions to be the input and thus provides a flexible desktop-based kinesthetic user interface. More importantly, the kinesthetic feedback generated by the devices have three or six degrees of freedom with resolution, up to 1 kHz (Massie and Salisbury, 1994). Therefore, these devices can enable realistic kinesthetic exploration as close as we have in the physical world.

Although using grounded force-feedback devices in VR is promising, it has a major challenge: the small device workspace. The length of the mechanical arm of the force-feedback devices is fixed, which limits the area that the user's hand can reach. This makes kinesthetic exploration in a large VR environment difficult. In addition, the prolonged operation of the devices can easily fatigue the user's hand (Ott et al., 2005; Hamam and El Saddik, 2015), which negatively influences user performance and experience during kinesthetic tasks (Allen and Proske, 2006; Cortes et al., 2014).

---

[4] http://www.cyberglovesystems.com/(accessed 1 July 2021)
[5] https://www.haptx.com/(accessed 1 July 2021)
[6] https://www.3dsystems.com/haptics (accessed 1 July 2021)
[7] https://www.hapticshouse.com/(accessed 1 July 2021)

1.1 RESEARCH CONTEXT AND OBJECTIVE

This dissertation focuses on kinesthetic interaction in the field of HCI. The objective **is to implement efficient and accurate kinesthetic interaction for VR and explore its application.** There are three steps in the research:

- Problem understanding: Identifying the issues with the traditional technique used to implement kinesthetic VR interaction.
- Development: Developing new interfaces which can enable fast kinesthetic interactions with good user control (i.e., efficient) and simultaneously maintain the accuracy of kinesthetic perception (i.e., accurate).
- Application: Exploring the significance of kinesthetic VR interaction.

Problem understanding

To overcome the issue of limited workspace for force-feedback devices, the current solution is to scale a small motion of the mechanical arm to a large motion of the haptic interaction point (HIP, the cursor in the virtual environment), that is, employing a large control-display (CD) gain (Argelaguet and Andújar, 2013). Multiple studies (Conti and Khatib, 2005; Dominjon et al., 2005a) have mentioned that applying a large CD gain may affect the use of force-feedback devices because of the mismatch between the hand motion and the HIP motion. To address this issue, they proposed new techniques that used large gains (>1) for only reaching targets and maintained the unit gain (=1) during the interaction.

The effects of CD gain on pointing devices like the mouse have been explored for many years (e.g., Casiez et al., 2008). Kinesthetic interaction involves more complex hand behaviors and interaction feedback. To the best of this researcher's knowledge, there have been no experimental studies to investigate how a large CD gain affects kinesthetic tasks in terms of task efficiency, accuracy and user experience. It is critical for the following research of the dissertation, because its answer can reveal the issues of the current technical solution for implementing kinesthetic VR interaction and motivate the following studies. This leads to the first research question of this research:

RQ1: How does a large CD gain (>1) affect task measures such as task completion time, task accuracy and user experience in kinesthetic tasks?

Development

Eye tracking is an emerging hands-free input mechanism in HCI. Using eye gaze as a mechanism for pointing and selection has been shown to be beneficial in numerous HCI scenarios in terms of task efficiency (Zhai et al., 1999; Majaranta et al., 2009) and user experience (Nukarinen et al., 2018). Previous studies (Stellmach and Dachselt, 2012; Pfeuffer et al., 2014) have used eye gaze to improve object acquisition and manipulation while using tactile input devices such as touchscreens.

To address the workspace issue for force-feedback devices, the research for this dissertation developed multimodal kinesthetic interfaces by using eye gaze as an input modality. The methods of using eye gaze in previous studies can be categorized into two types: directly replacing hand motions with eye gaze for pointing (e.g., Majaranta and Räihä, 2002; Kumar et al., 2007; Pfeuffer et al., 2014; Pfeuffer et al., 2017) and combining hand motions and eye gaze for pointing (e.g., Zhai et al., 1999; Stellmach and Dachselt, 2012). The present research investigated both the gaze-based methods in the context of kinesthetic interaction. The resulted multimodal kinesthetic interfaces have the potential to implement efficient kinesthetic interaction in VR. In addition, kinesthetic perception is important for our manual tasks. It is necessary to investigate that whether these gaze-based interfaces affect kinesthetic perception. These lead to the second research question:

RQ2: How do the multimodal kinesthetic interfaces that use eye gaze as an input modality affect user performance in terms of task efficiency, kinesthetic perception accuracy and user experience?

Application

Kinesthetic interaction using grounded force-feedback devices has been proposed for many professional applications. For example, it has a variety of applications in the field of medicine, including education (Kinnison et al., 2009), surgery simulation and training (Bielser and Gross, 2000; Webster et al., 2004; Steinberg et al., 2007; Alaraj et al., 2015), robot-assisted surgery (Okamura, 2004) and medical diagnosis and planning (Medellín-Castillo et al., 2016). In addition, as the visual output channel, VR HMDs have been used for many medical services, such as treating chronic pain (Jones et al., 2016), anxiety disorders and phobias (Maples-Keller et al., 2017). They have also been used to present 2D slices (King et al., 2016; Wirth et al., 2018) and volumetric imaging data (Randall et al., 2016; Sousa et al., 2017; Venson et al., 2017) for medical diagnosis and planning.

The present research combined a VR headset with a force-feedback device to create a kinesthetic VR interface that provided the user with a flexible viewing perspective, natural hand-based input and kinesthetic feedback together. The research also involved the vibrotactile VR interface that used a VR headset and a VR controller. The two VR interfaces are promising for tasks involving 3D manipulation (Bowman et al., 2004), such as the tasks of medical diagnosis and planning. However, it remains unknown how well the two haptic VR interfaces compare with the traditional 2D interface that uses a mouse and a 2D display in medical tasks. The main objective of this research was to explore the usability of the kinesthetic VR interface to determine if it has the potential to replace the traditional 2D interface in the field of medicine. Therefore, there is the research question three:

RQ3: How does the kinesthetic VR interface affect user performance in medical 3D manipulation tasks in terms of task efficiency, accuracy and user experience?

To answer these research questions, multiple experimental studies have been conducted. The basic methodology for these studies is described in the following section.

1.2 METHODOLOGY

The research for this dissertation required applying both constructive and empirical research methods. There are four studies included in this dissertation, all conducted in a controlled lab environment. Before each study, a prototype system was constructed based on the experiment. This included setting up the experimental hardware, developing the required software and designing the experimental task. Both quantitative and qualitative research methods were used in the research. The quantitative measures were collected by logging the objective data of users' task performance as well as the rating data of users' subjective feelings. Qualitative measures were collected by recording users' general responses to the interfaces and other free comments related to the experiment.

Study I was conducted to answer RQ1. It included a real-world kinesthetic task as the experimental task to test the effects of a large CD gain on kinesthetic interaction. The study began by implementing the experimental task and establishing the test platform. We used quantitative measures in the study. The effects of CD gain were explored by investigating objective data (task completion time, interaction accuracy and hand movement pattern) and subjective data (hand tiredness, naturalness, pleasantness and user confidence to successfully perform the task).

Studies II and III were conducted to answer RQ2. For each study, a prototype system was first implemented by developing new multimodal kinesthetic interfaces to address the challenge of kinesthetic VR interaction. The multimodal kinesthetic interfaces were evaluated by comparing them with the traditional hand-based kinesthetic interface. The experimental task was to reach and touch objects and then detect their physical properties. The system implementation was followed by a pilot study to define key system parameters and design choices. We used quantitative research method in the two studies. Task completion times, task errors and users' subjective data were collected to evaluate the new kinesthetic interfaces. In Study II, qualitative research method was also used. Participants completed a post-experiment questionnaire to vote for their favorite interfaces and give the reasons for their preferences.

Study IV was conducted to answer RQ3. It began by implementing a real-world task in computer-aided medical diagnosis and planning. The experimental system was then developed involving two haptic VR interfaces along with the traditional mouse-based interface. A pilot study was conducted to determine the system parameters. Both quantitative and qualitative research methods were used. As in the previous studies, task completion times, task errors, users' subjective data and comments were collected to evaluate the usability of the two haptic VR interfaces.

1.3 CONTRIBUTIONS

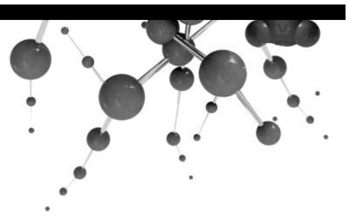
This dissertation makes the following main contributions: First, the research evaluated the technique of CD gain based on a kinesthetic task. To the best of this researcher's knowledge, it is the first known study to provide scientific evidence for the effects of CD gain on kinesthetic interaction. The results of studies in terms of user performance and experience provided an empirical basis for the following studies of this dissertation as well as future design of kinesthetic interfaces based on the CD gain method.

Second, the research explored a novel design space for kinesthetic interfaces that used eye gaze as a kinesthetic input modality, and three new multimodal kinesthetic interfaces were designed and developed (two in Study II and one in Study III). These interfaces successfully addressed the limited workspace of force-feedback devices without using the CD gain method, and they also reduced the hand workload from prolonged kinesthetic exploration. In addition, the research explored users' ability to detect the physical properties of objects while using the multimodal kinesthetic interfaces. The results increased our understanding of human kinesthetic perception in the virtual environment, and simultaneously they identified the critical factors for designing high-quality kinesthetic interfaces in terms of the accuracy of kinesthetic perception.

In parallel, the research developed and proposed two haptic VR interfaces for improving current medical diagnosis and planning work. The research revealed their strengths and weaknesses and thus provided an empirical understanding for developing efficient, accurate and user-friendly interactive VR systems. More importantly, by comparing them with the state-of-the-art 2D interface, the research demonstrated the potential of the haptic VR interfaces, particularly the kinesthetic VR interface, to replace the 2D user interface in the field of medical diagnosis and planning.

1.4 STRUCTURE

This dissertation consists of a summary and four research articles. Three of the articles were published in highly ranked peer-reviewed international journals, and the fourth was published in peer-reviewed international conference proceedings. Chapter 2 contains an introduction of VR, including the immersion levels of VR, current VR systems and their applications. Chapter 3 provides basic information about kinesthetic sensation as well as a review of the literature on kinesthetic interaction. The challenge of interaction in VR, the previous work to address the challenge and professional applications are described in detail in Chapter 3. Chapter 4 provides an overview of gaze-based haptic interaction, including the sense of vision, hand-eye coordination and existing gaze-based haptic interactions. The methodology, procedure and resources of the research are described in Chapter 5. In Chapter 6, the four publications are presented in detail in terms of research goal, methodology, results and discussion. The key findings of the dissertation, research limitations and potential future work are described in Chapter 7. Chapter 8 concludes the dissertation with a summary of the contributions of these efforts.



2 Virtual Reality

Three-dimensional computer graphics have entered many areas of our lives. To interact with the 3D models displayed on a 2D screen, users commonly use a mouse and a keyboard to manipulate these models and perform various tasks. However, instead of just seeing the images on the screen, it is possible to let the users step into this artificial world and interact with the scenes and objects as naturally as they can in the real world. Such interactive technology is called VR technology. Generally, the main goal of VR is to create a world with which users can interact and that their senses can perceive as realistically as the physical world.

This chapter provides an overview of current VR technology and applications. The factors that influence the immersion of interactive VR systems is described. Then, commercial VR systems and their existing applications are presented. The chapter concludes with the future development required for VR and the research direction of this dissertation.

2.1 THE IMMERSION LEVELS OF INTERACTIVE VR SYSTEMS

The goal of interactive VR systems is to allow users to interact with the virtual environment as naturally and realistically as they do in the physical world. Therefore, immersion is a key element for VR systems. The immersion level of an interactive VR system describes its capacity to deliver an illusion of reality to the senses of human users. It can be defined based on five aspects of interaction: *inclusive*, *extensive*, *surrounding*, *vivid*, and *matching* (Slater and Wilbur, 1997). In this section, the traditional 2D interface that uses a mouse and a 2D display is used to illustrate these factors.

First, *inclusive* refers to how strong the signals indicating the existence of the physical world are. These signals are related mainly to the interaction tools, such as the input method, the device weight and the noise generated. For example, in the traditional 2D interface, using a mouse to rotate 3D objects requires a mouse-based rotation technique, such as Virtual Sphere (Chen et al., 1988) and ArcBall (Shoemake, 1992). These techniques transfer the movement of the mouse on a 2D plane to the rotation axis and angle of the 3D object. However, this manner of interaction is different than our hand behaviors when we rotate an object in the physical world. Thus, it clearly indicates the presence of the mouse and negatively affects the aspect of inclusion.

Second, *extensive* refers to how many sensory modalities are involved in the interaction. Human sensory systems are mainly the visual system (sense of sight), auditory system (sense of hearing), somatosensory system (sense of touch), olfactory system (sense of smell) and gustatory system (sense of taste). Both visual and auditory modalities are commonly present in the traditional 2D interface. However, haptic modality as well as olfactory and gustatory modalities are often missing in this interface.

Third, *surrounding* refers to the size of the field of the user's peripheral vision in the virtual environment. A traditional 2D display has a clear boundary for its view of the virtual environment. This largely limits the user's peripheral vision, so it negatively affects the aspect of surrounding.

Fourth, *vivid* refers to how well the virtual environment performs in terms of vision, including fidelity and resolution. Visually presenting 3D models using a 2D display requires 3D projection techniques (Foley et al., 1995). The resulting visual effect of 3D objects on a high-definition 2D screen is as realistic as human visual effect in the real world. Thus, a 2D visualization system performs well in this aspect.

Fifth, *matching* refers to how close visual feedback is to proprioceptive feedback through motion capture. The traditional 2D interface that uses a mouse has a limited motion capture on the 2D plane, which gives it a low level in this aspect.

To sum up, two immersion factors of VR, *surrounding* and *vivid*, are related closely to visualization systems and the other three factors of VR, *inclusive*, *extensive* and *matching*, are intimately bound up with the manipulation tools used. The traditional 2D interface performs less well in terms of being *inclusive*, *extensive*, *surrounding* and *matching*. Thus, it is considered a low level of immersion or even a non-immersive VR system (Robertson et al., 1993). To improve the immersion of VR, multiple new VR systems have been developed and are available on the market. The next section introduces these commercial VR systems and their applications.

2.2 INTERACTIVE VR SYSTEMS AND APPLICATIONS

Commercial interactive VR systems

The earliest interactive VR system can be traced back to 1962. Sensorama, a multi-sensory simulator, was designed to show a film on a stereoscopic color display, enhanced by simple feedback with stereo sound, scent, wind and vibration (Burdea and Coiffet, 2003). In the following decades, numerous interactive systems have been developed for VR. The most popular commercial VR equipment is related mainly to the visual modality.

Of the five human senses, vision is considered the most important for interaction (Hutmacher, 2019). Thus, the development of visualization systems is a significant research area that is growing rapidly. Traditional human-computer interaction systems for VR mainly adopted a 2D display as the visual output channel. This display provides a fixed screen-based viewing perspective with limited peripheral vision for the surrounding environment. To provide a more flexible and natural viewing perspective to users, multiple new visualization systems have been developed. For example, the Cave Automatic Virtual Environment (CAVE) presents 3D virtual environments by employing projectors to create rear-projection screens on the walls of a room (Cruz-Neira et al., 1992). Users can move freely and see the virtual objects around them by wearing 3D glasses (see Figure 2.1).

An important visualization system for VR is the HMD. The first prototype VR HMD was designed as early as 1968 (Sutherland, 1968). HMDs continue to be developed rapidly, and they are widely available on the market, such as HTC Vive and Oculus Rift (see Figure 2.1). These headsets use motion tracking sensors to give users a head movement-based viewing perspective. In addition, unlike the traditional computer interfaces that use a mouse to manipulate 3D objects, the VR motion controller was developed to allow users to use hand gestures to interact with objects. This interactive method is close to our hand behavior in the

HTC Vive | Oculus Rift | CAVE[8]

Figure 2.1. Examples of interactive VR systems.

---

[8] https://www.steantycip.com/projects/virtual-reality-cave/(accessed 1 July 2021)

physical world and thus largely improves the naturalness of interaction.

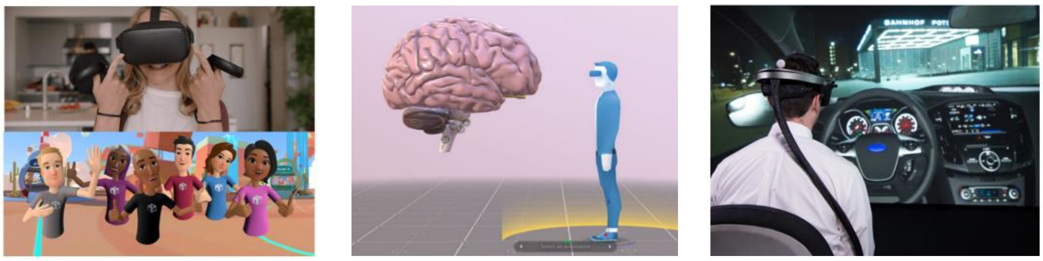
Applications

VR equipment has been widely used in both the entertainment and professional fields (see Figure 2.2). In entertainment, one important application is in the gaming industry. The interactive virtual environment that VR equipment brings to video games can largely improve gaming satisfaction and enjoyment (Shelstad et al., 2017). Furthermore, the use of VR headsets provides a novel experience to watch movies, sports games and concerts. For example, the VR display can present an active cinematic 360-degree virtual environment for movies instead of the traditional screen-based passive storytelling (Dooley, 2017). In addition, VR technology can allow users to virtually travel to different places (Guttentag, 2010), improve visiting experience in museums (Jung et al., 2016) or meet people in a virtual world for social activities (e.g., Facebook Horizon[9]).

In professional fields, VR equipment has been considered a low-cost, easy-to-use and powerful tool for professional education and training (Kamińska et al., 2019). It can help learners acquire knowledge and train skills without taking physical risks.

In medicine, VR equipment such as the CAVE (Al-Khalifah et al., 2006; Shen et al., 2008) and VR headsets (Sousa et al., 2017; Venson et al., 2017; Wirth et al., 2018) has been used to present medical imaging data of human skeletons and organs. This could aid learning and teaching in anatomy education (e.g., Alfalah et al., 2019; Erolin et al., 2019) and provide an effective tool in surgery training (e.g., Ros et al., 2017; Pulijala et al., 2018).

In addition, VR equipment has been used in the education of civil engineering, and it helped visualize different civil engineering disciplines with virtual models and motivate students during the learning process



Facebook Horizon

Anatomy Education (Erolin et al., 2019)

Ford FIVE

Figure 2.2. Examples of VR applications.

---

[9] https://www.oculus.com/facebook-horizon/community (accessed 1 July 2021)

(Dinis et al., 2017). Similarly, VR equipment has been used to promote mechanical and electrical engineering education by providing effective training in the applications (Kamińska et al., 2017). VR equipment has many other applications in fields that have physical risks, such as military and astronaut training. For instance, VR vehicle and flight simulators have been used to train military staff (Lele, 2011) and projector-based virtual environments have been proposed for astronaut assembly sequence training (Rönkkö et al., 2006).

Besides education and training, VR technology can directly assist professional activities and improve the quality of their services. For example, VR equipment has been proposed for medical purposes, such as treating chronic pain (Jones et al., 2016), anxiety disorders and phobias (Maples-Keller et al., 2017). Multiple companies have used VR technology in manufacturing. For example, The Ford Motor Company has developed Ford Immersive Vehicle Environment (FIVE[10]). It allows engineers to inspect and develop car models. The Visual Components [11] company has employed VR equipment in production design. In retail, companies like eBay[12] have used VR to enhance the shopping experience.

2.3 CURRENT STATUS AND FUTURE DEVELOPMENT

VR equipment has numerous applications in various fields because it enhances the user's immersion in vision and object manipulation. For instance, compared with the 2D display, the CAVE and VR HMD can present a wide peripheral vision for the environment surrounding the user, thus providing a higher level in *surrounding*. Furthermore, Both VR visualization systems could perform well in *vivid*. The projector-based CAVE environment has high resolution because it uses small pixel sizes to retain the illusion of reality (Cruz-Neira et al., 1992). Current VR headsets often provide a lower visual effect, because of issues like the screen door effect (Cho et al., 2017). However, they generally maintain fine fidelity and resolution for visual effects. In addition, VR controllers can track the user's hand and allow natural hand motions to be the input modality. This interaction method helps reduce the presence of the VR controllers, and it also makes visual feedback match proprioceptive feedback closely. Thus, both *inclusive* and *matching* aspects have been improved. Moreover, vibrotactile feedback from VR controllers is helpful for improving *extensiveness*. In sum, although some issues remain, such as the weight of

---

[10] https://www.at.ford.com/en/homepage/news-and-clipsheet/news/2014/1/how-ford-uses-virtual-reality-in-product-development.html (accessed 1 July 2021)

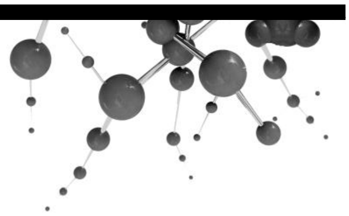
[11] https://www.visualcomponents.com/visual-components-experience/(accessed 1 July 2021)

[12] https://www.chaostheorygames.com/work/ebay-vr (accessed 1 July 2021)

VR devices and the discomfort from wearing a VR headset, the immersion of VR has been largely improved by current VR equipment.

However, VR technology still has a long way to establish a realistic interactive virtual environment. One important interaction channel in the physical world is the sense of touch. VR controllers permit hand-based gestures as the input modality, but they provide only simple vibration feedback. This kind of haptic feedback is unrealistic for touch behaviors. Other haptic technologies used for VR such as electrical muscle stimulation (e.g., Takahashi et al., 2019) are still under development.

As mentioned before, kinesthetic interaction implemented by using force-feedback devices permits the user′ hand motions as the input modality and provides kinesthetic feedback that can realistically simulate the feeling of touch. Implementing such bidirectional touch behaviors can further improve the immersion of VR and greatly extend its applications. For example, allowing the customers to freely select the commodities in a large virtual shopping mall and simultaneously allowing them to realistically touch can largely improve their shopping experience. Therefore, implementing such efficient and accurate kinesthetic VR interaction is the primary focus of the research for this dissertation. The next chapter provides background information about current kinesthetic interaction in HCI.



3 Hand-based Kinesthetic Interaction

While modalities such as taste and smell remain largely underexploited (Obrist et al., 2016), haptic (touch-based) technology has gained increasing interest and attention in HCI. Tactile interaction as a type of haptic interaction concentrates on sensations like perceiving texture, pressure and vibration, whereas kinesthetic interaction as another type focuses on sensations like perceiving hardness, weight, mechanical compliance and inertia (El Saddik et al., 2011). Humans naturally receive both tactile and kinesthetic feedback during touch-based interactions in the physical world. In HCI, tactile interaction by itself could not provide realistic sensations of touch for virtual objects. However, although kinesthetic feedback lacks the rich information of the contact surface associated with tactile feedback, it can realistically simulate the feeling of touch. So, kinesthetic interaction is promising for VR.

This chapter presents an overview of kinesthetic interaction in human-computer interaction. This includes introducing the movement-based kinesthetic sensation, current hand-based kinesthetic interfaces and their applications.

3.1 KINESTHETIC SENSATION

Our somatosensory system can process several modalities of somatic sensation: pain, temperature, proprioception, the cutaneous sense and kinesthesis (Goldstein and Brockmole, 2016). Because the communicative value of temperature and pain variations has been considered very low (Geldard, 1960), current haptic technology in HCI mainly focuses on

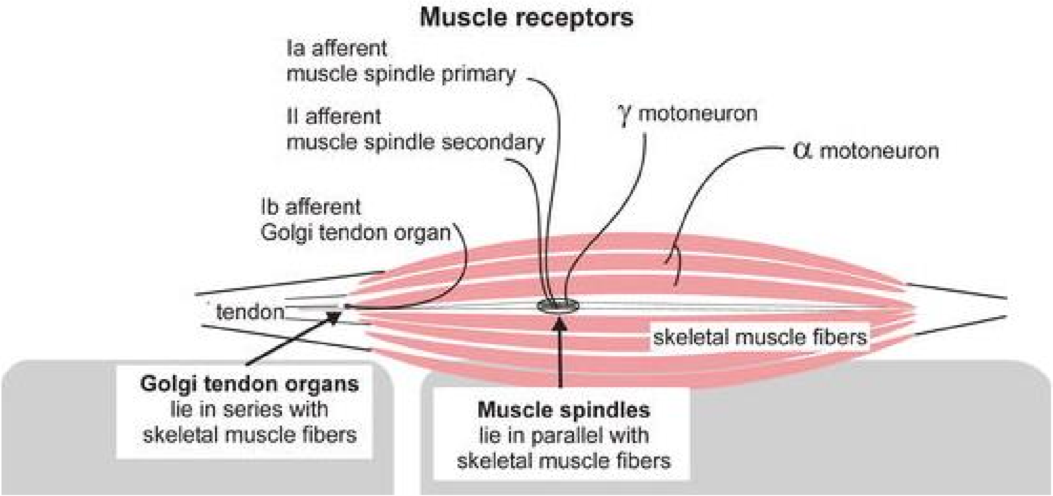


Figure 3.1. The structure of muscle receptors including Golgi tendon organs and muscle spindles (Taylor, 2016).

communicating proprioception, the cutaneous sense and kinesthesis between the user and the computer. The cutaneous sensation, also called tactile sensation, is based on the receptor stimulation in the skin. Both proprioception and kinesthesis are related to force sensation. Proprioception concentrates on the position of the limbs, and it is critical for maintaining posture and balance. Kinesthesis is related to the movement-based force sensation (Goldstein and Brockmole, 2016). To implement bidirectional touch behaviors in the virtual environment, the present research focused on the kinesthetic sensation.

According to physiology (Kaas, 2012; Taylor, 2016), sensory signals from the receptors in the muscles, tendons and joints significantly contribute to kinesthetic sensation. These receptors mainly include muscle spindles, Golgi tendons (see Figure 3.1) and encapsulated endings in the joints (see Figure 3.2).

Muscle spindles are fluid-filled capsules of connective tissue, which are 0.5 cm to 1 cm long and lie parallel to the skeletal muscle fibers. These fibers contain different motoneurons. Because of this structure, the muscle spindles are sensitive to the length of the skeletal muscle fibers. They have two sensory endings which can perceive the muscle stretch: primary (Ia) and secondary (II) spindle afferents. Both spindle afferents possess static sensitivity, but the primary spindle afferents are also dynamically sensitive to the stretching of the muscles. Therefore, the primary spindle afferents contribute to the sense of limb position and movement, whereas the secondary spindle afferents contribute only to the sense of position. In addition, Golgi tendon organs are attached to the tendon at one end and to many muscle fibers at the other end. Thus, these organs can be stretched or slackened by force when the muscles move, and they have both static and dynamic sensitivity.

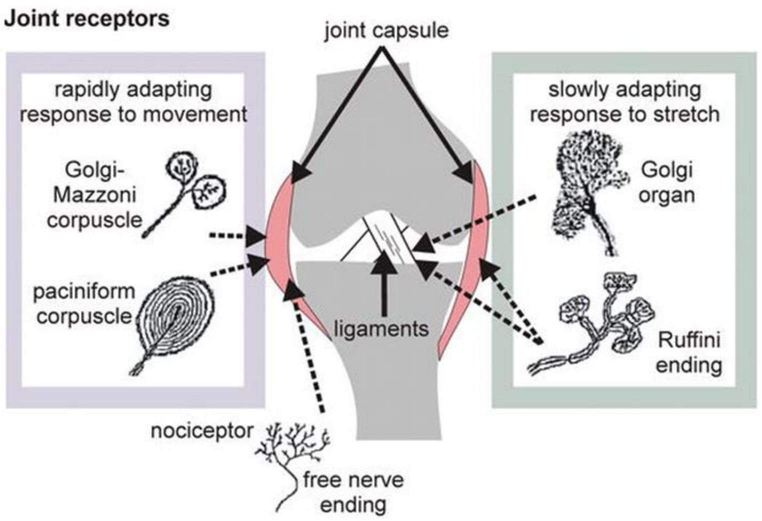


Figure 3.2. The structure of Joint with different encapsulated endings (Taylor, 2016).

In a joint, there are several types of encapsulated endings, which contribute to kinesthetic sensation. They include Ruffini endings, Golgi organ endings, Golgi-Mazzoni corpuscle endings, paciniform endings and some free nerve endings. The Ruffini endings are mainly in the joint capsule and the Golgi organ endings are in the ligaments of the joint. Both the Ruffini endings and Golgi organ endings belong to the slowly adapting mechanoreceptors, and they are responsive to the mechanical deformation within the joint. The Golgi-Mazzoni corpuscles and paciniform endings are primally located in the joint capsule. They are also in articular fat pads. Both endings are rapidly adapting receptors, sensitive to the compression of the joint. Our kinesthetic sensation mostly relies on the slowly adapting receptors. The skin also contains the Ruffini endings, which supports kinesthetic sensation.

Kinesthetic sensation is important for our daily manual tasks. We naturally perceive several physical properties of objects, such as hardness and weight, based on kinesthetic sensation. In addition, the perception of surface and object properties is closely bound to the nature of contact (Lederman and Klatzky, 2009). For example, we commonly employ lateral hand motions for texture detection and press the object for hardness identification. Therefore, the usability of kinesthetic interfaces may directly affect this perception process. While developing new kinesthetic interfaces, it is necessary to investigate their effects on the accuracy of kinesthetic perception.

3.2 EXISTING HAND-BASED KINESTHETIC INTERFACES

Kinesthetic mechanoreceptors are widely distributed in the muscles, tendons, joints and skin. Therefore, kinesthetic sensation occurs throughout the human body. We naturally use our hands to actively perform touch behaviors for perceiving and exploring the physical world. Therefore, the research for this dissertation concentrates on hand-based kinesthetic interaction. Many devices have been developed to implement hand-based kinesthetic interaction. These devices can be generally categorized as grounded and ungrounded kinesthetic devices.

Hand-based kinesthetic interfaces that use ungrounded devices

Current popular ungrounded kinesthetic devices are haptic gloves. They use an exoskeleton to provide kinesthetic feedback to the user’s hand. Existing haptic gloves are mostly developed based on hydraulic, pneumatic and electromechanical systems (Hinchet et al., 2018).

For example, CyberGlove and SenseGlove [13] are based on the electromechanical system. Motors or brakes are used in these gloves, and tendon-based cables are positioned over the top of the hand to control finger position and generate kinesthetic feedback. Other electromechanical wearable haptic devices such as HapThimble (Kim et al., 2016) use a servo motor to connect a cap contacted with the fingertip to provide kinesthetic feedback. Some other haptic gloves are developed based on hydraulic/pneumatic systems that use pumps and valves. Two examples are ExoHand[14] and HaptX. Winter and Bouzit (2007) and Zubrycki and Granosik (2017) proposed different types of hydraulic haptic gloves that use the jamming principle or magnetorheological fluid to provide kinesthetic feedback.

Besides haptic gloves, there are other ungrounded kinesthetic devices. For example, Kamuro et al. (2011) developed a pen-shaped device that can provide kinesthetic feedback to the user’s fingers to simulate the feeling of touch. Some examples of ungrounded devices are shown in Figure 3.3.

Theoretically, kinesthetic interfaces that use ungrounded devices are promising for VR. They have a large arm-based workspace and thus allow users to have more free kinesthetic interaction with virtual objects.

CyberGlove

ExoHand

Kinesthetic pen
(Kamuro et al., 2011)

Figure 3.3. Examples of ungrounded kinesthetic devices.

---

[13] https://www.senseglove.com/(accessed 1 July 2021)

[14] https://www.festo.com/group/en/cms/10233.htm (accessed 1 July 2021)

However, current ungrounded devices have several shortcomings. For instance, haptic gloves that use motors or pumps are difficult to generate a consistently sufficient force (Hinchet et al., 2018). Furthermore, the kinesthetic feedback from the gloves has limited degrees of freedom, and it often applies only to the user's fingers because of their mechanical structure. In addition, the hydraulic/pneumatic haptic gloves normally have bulky and heavy hardware systems (e.g., ExoHand and HaptX) because the hydraulic systems with pumps and valves are difficult to miniaturize. Electromechanical haptic gloves like CyberGlove are very costly. For the kinesthetic pen (Kamuro et al., 2011), kinesthetic feedback was provided by the built-in springs and motors. It also has low resolution and fewer degrees of freedom, similar to the haptic gloves. These technical shortcomings limit the application of current ungrounded kinesthetic interfaces.

Hand-based kinesthetic interfaces that use grounded devices

Grounded kinesthetic devices are also available on the market, such as Geomagic Touch, Novint Falcon and Force Dimension Omega[15]. Some examples of grounded devices are shown in Figure 3.4.

Grounded devices such as Touch X often have a mechanical arm that allows the user to perform natural hand touch behaviors. The inside encoders track the x, y and z coordinates of the user's hand, and the built-in actuators generate the force and torque with three or six degrees of freedom to simulate the feeling of touch (Massie and Salisbury, 1994). Although using grounded force-feedback devices lacks detailed cutaneous information, the hand motions imparted by our limbs and the reaction force generated by the devices can provide sufficient information about the physical structure and characteristics of the virtual objects touched, such as shape, stiffness, weight, and even friction and texture.

Kinesthetic interfaces that use grounded force-feedback devices have two advantages (Massie and Salisbury, 1994). First, the mechanical system has little back-drive friction, low inertia and no unbalanced weight. This

Geomagic Touch X

Novint Falcon

Force Dimension

Figure 3.4. Examples of grounded force-feedback devices.

---

[15] https://www.forcedimension.com/ (accessed 1 July 2021)

means that, while holding the mechanical arm, the user can freely move through the virtual space without any external force that can restrict hand movement. In contrast, ungrounded devices like hydraulic haptic gloves are bulky and heavy. Second, the feelings of touching virtual objects can be implemented as realistically as our feeling of touch in the physical world. For example, if there is a virtual cube made of stone, the user can feel stiffness and shape by using the force-feedback device to touch the cube from different angles. This is because the built-in actuators provide stable and sufficient force with high resolution (up to 1 kHz) and three degrees of freedom. Force-feedback devices with six degrees of freedom can generate kinesthetic feedback for additional hand motions: yaw, pitch, and roll.

Compared with current ungrounded devices, grounded force-feedback devices are more promising for VR because of two reasons. First and most important, the grounded force-feedback devices could support flexible hand input with six degrees of freedom and generate high-resolution force with three or six degrees of freedom. In contrast, current ungrounded devices provide kinesthetic feedback with low resolution and limited degrees of freedom. Second, ungrounded devices such as wearable haptic gloves (e.g., CyberGlove, ExoHand and HaptX) require the user to wear devices that are often bulky and heavy, whereas the force-feedback devices are placed on a desk and are thus more user-friendly. In addition, although grounded force-feedback devices only provide a desktop-based use interface, it is our primary working environment and has many potential applications.

However, there is a major challenge while using grounded force-feedback devices for kinesthetic VR interaction, that is, the limited workspace. The length of the mechanical arm of the force-feedback device limits the space for interaction, whereas the size of the virtual space provided by a VR headset or even a 2D display can be limitless. The first proposed solution to this challenge was a manual clutching technique (Johnsen and Corliss, 1971). It used a button on the force-feedback device to declutch HIP in the virtual environment, and then the user could move the device arm back to the center of the physical workspace. This method is close to how a mouse is used when it reaches the edge of the mousepad. The user lifts the mouse and places it back in an appropriate position. However, for a large virtual space, reaching and touching a distant virtual object may require a great number of re-clutching operations. So, the clutching technique is not well-suited for kinesthetic VR interaction.

The CD gain (Argelaguet and Andújar, 2013) is currently the main technique used to address this challenge. It defines how the translations and rotations of input devices (x) transfer into the translations and rotations of their virtual representants (X), that is:

$$\text{CD gain} = \Delta X / \Delta x \qquad (1).$$

The technique of CD gain has been used with many input devices to address the issue of limited workspace for interacting with large environments. For example, while using a mouse to interact with objects on a large screen, we commonly increase the CD gain to allow the user to quickly reach a distant target without moving the mouse a long distance on the table (Argelaguet and Andújar, 2013). Techniques such as the PRISM (Frees and Kessler, 2005) and Adaptive Pointing (König et al., 2009) are implemented based on dynamic CD gains. Furthermore, multiple studies have investigated the effects of CD gains on user performance for pointing tasks using the devices like the mouse. Using a large CD gain could reduce task completion time (e.g., Frees and Kessler, 2005; Casiez et al., 2008; König et al., 2009). However, a very large gain may increase task completion time (e.g., Kwon et al., 2011). In addition, previous research (e.g., Kwon et al., 2011) also noted that using a large gain may affect subjective feelings such as ease of use and preference.

The bare-hand gesture input and handheld input devices also have a limited workspace because of the length of the user's arm. Poupyrev et al. (1996) proposed the go-go technique to expand the hand workspace by applying a CD gain on the virtual hand. Similarly, Vogel and Balakrishnan (2005) proposed a freehand pointing and clicking method for interacting with a large display from a distance. They adopted a CD gain to increase the range of cursor movement on the screen.

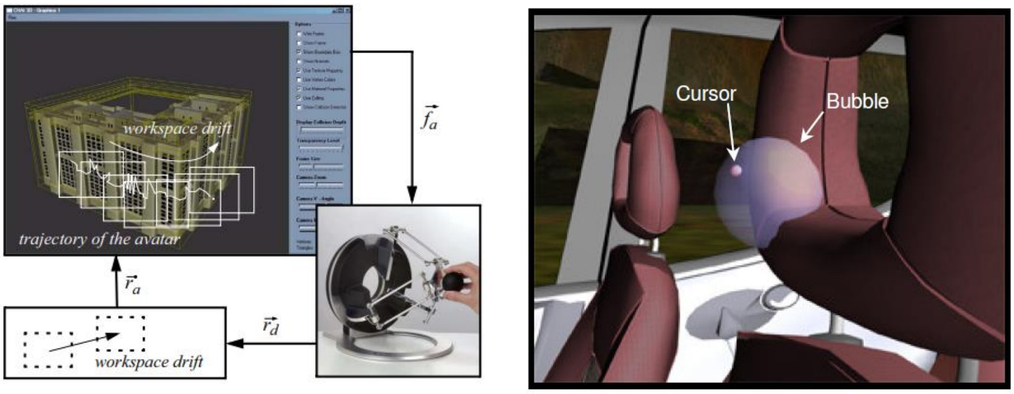
For force-feedback devices, when the explored virtual space is larger than the physical device workspace, a large CD gain (>1) also can be used so the workspace of the force-feedback device can be increased to match the required size of the virtual space. This feature is available in most haptic APIs and plugins, such as the opensource H3D API[16] and the Unity haptic plugin[17]. Although the technique of CD gain can help explore a large VR environment, it leads to an additional issue. Applying a CD gain (≠1) causes a mismatch between the hand motion and the HIP motion. Several studies (Conti and Khatib, 2005; Dominjon et al., 2005a) have argued that for kinesthetic tasks that require fine control of the HIP, this mismatch is likely to affect user performance. To avoid this issue, these studies developed new techniques to maintain the unit CD gain (=1) while manipulating objects and doing tasks (see Figure 3.5).

Conti and Khatib (2005) proposed a Workspace Drift Controller. It uses large CD gains for the HIP movement when the user wants to traverse a large virtual space. During object manipulation, this technique maintains

---

[16] http://www.h3dapi.org/(accessed 1 July 2021)

[17] https://www.assetstore.unity.com/packages/essentials/tutorial-projects/unity-5-haptic-plugin-for-geomagic-openhaptics-3-3-hlapi-hdapi-34393 (accessed 1 July 2021)



Workspace Drift Controller
(Conti and Khatib, 2005)

Bubble
(Dominjon et al., 2005a)

Figure 3.5. Example techniques to address the workspace issue.

the unit gain and progressively centers the device workspace during the interaction. Dominjon et al. (2005a) proposed another technique: Bubble. It puts a sphere around the HIP and then determines the CD gain based on the relative positions of the HIP and the sphere. When the user moves the HIP out of the sphere, a large CD gain is applied. So, the HIP can move fast to reach the target object. When the HIP stays inside the sphere, the unit CD gain is applied so that the motion of the HIP matches the user's hand motion. In this case, the user can precisely control the HIP to do kinesthetic tasks.

These two techniques have been examined with a 2D display. However, their feasibility and usability in the large virtual environment provided by VR equipment such as VR headsets remain unknown. Moreover, hand fatigue is another issue for the scaling-based kinesthetic interfaces. When the user holds the mechanical arm to do kinesthetic tasks, he or she must move the hand repetitively to reach and touch objects. This easily fatigues the user's hand, especially during a prolonged interaction (Ott et al., 2005; Hamam and El Saddik, 2015). Hand fatigue will negatively affect user performance in spatial position and haptic manipulation (Allen and Proske, 2006; Cortes et al., 2014) and thus kinesthetic task performance. Therefore, reducing hand fatigue is necessary in the design of new kinesthetic interfaces.

3.3 PROFESSIONAL APPLICATIONS OF KINESTHETIC INTERFACES THAT USE FORCE-FEEDBACK DEVICES

Force-feedback devices have been used in many professional applications, such as industry and medicine. In the field of industry, kinesthetic feedback from force-feedback devices can support product development.

For example, Seth et al. (2006) proposed a SHARP system that used a force-feedback device as the manipulation tool. This system could simulate various interactions during industrial manufacturing, leading to faster product development and faster identification of design issues. Read et al. (2016) developed a haptic sketch-based system to conduct concurrent design for manufacture and assembly. This could decrease the time taken for design process, reduce the cost and lead to a better design solution. Teklemariam and Das (2017) conducted a case study that used a force-feedback device in virtual product design. The study explored and demonstrated the advantages of kinesthetic feedback in product development and evaluation.

In the field of medicine, kinesthetic feedback benefits various medical practices, such as anatomy education (Kinnison et al., 2009), robot-assisted surgery (Okamura, 2004; Ortmaier et al., 2007), diagnosis and analysis using medical images (Medellín-Castillo et al., 2016) and surgery training (Bielser and Gross, 2000; Webster et al., 2004; Steinberg et al., 2007; Alaraj et al., 2015; Ribeiro et al., 2016). Some examples are shown in Figure 3.6.

In medical education, students often face psychological problems, such as anxiety and fear, when learning anatomy with real tissues and organs. This negatively affects the learning process. Kinnison et al. (2009) proposed using an interactive system integrated with a force-feedback device for anatomy learning. The results showed that the haptic interaction system provided an engaging and efficient learning method while addressing the psychological issues with using real samples.

Force-feedback devices have also been proposed for robot-assisted surgery. Okamura (2004) studies whether the lack of force feedback plays a critical role in mis-operation problems during robot-assisted surgery. The results demonstrated that adding force feedback could improve the surgeon's performance to control the robotic systems and reduce errors in the operation. Similarly, Ortmaier et al. (2007) compared manual and robot-

Anatomy education
(Kinnison et al., 2009)

Surgery simulation
(Alaraj et al., 2015)

Figure 3.6. Examples of the hand-based kinesthetic interfaces in medicine.

assisted minimally invasive surgery techniques, and they used a force-feedback device as the operational interface in the robot-assisted system. The results showed that force feedback could significantly reduce the injuries from minimally invasive surgery.

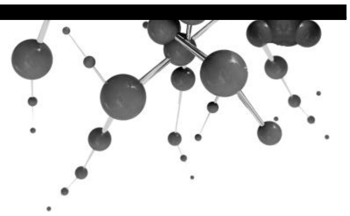
Furthermore, medical diagnosis and planning using images of human skeletons and organs is important work for radiologists (Kula and Ghoneima, 2018). Medellín-Castillo et al. (2016) have used a force-feedback device as the input tool for such work. They applied kinesthetic feedback for 2D, 2.5D and 3D cephalometry marking and examined their differences. The results showed that haptic-based 3D cephalometry marking provided better performance in terms of task accuracy than haptic-based 2D and 2.5D cephalometry marking.

Surgery training is another important application of force-feedback devices in the field of medicine. In the virtual surgeries, these devices can be used to interact with various virtual human tissues and organs and provide kinesthetic feedback for them. Such training can be repeated without wasting surgical samples or risking failure (Coles et al., 2011). Webster et al. (2004) developed a haptic simulator for cataract surgery with a force-feedback device as the input tool. The kinesthetic feedback enriched the surgeons' interaction channel and reduced operational errors. Similarly, Steinberg et al. (2007) developed a haptic dental training simulator using a force-feedback device. They demonstrated that such a simulator could help students develop dental tactile skills.

Researchers have also implemented other virtual surgical tasks with force-feedback devices. For example, Bielser and Gross (2000) developed a mathematical model for surgical incisions by using a force-feedback device as the scalpel. Similar studies include modelling the haptic device as a needle for surgical sutures (Webster et al., 2001) and modelling the device as different surgical instruments to simulate aneurysm clipping (Alaraj et al., 2015).

In addition, Ribeiro et al. (2016) provided an up-to-date survey to investigate haptic techniques and devices used in palpation simulation. Force-feedback devices have been used widely in such areas as tube thoracostomy (Everett et al., 2009), cancer detection (Widmer and Hu, 2012), osteopathic medicine and physiotherapy (Williams II et al., 2004; Howell et al., 2008) and dermatology (Lee et al., 2013).

Kinesthetic interaction using force-feedback devices has been found to be beneficial for various medical work, and it has even been proposed for the medical tasks commonly performed by using the traditional mouse-based interfaces, such as the 3D manipulation tasks of medical diagnose and planning (Medellín-Castillo et al., 2016). Previous studies have argued that, compared with the mouse-based interface, using force-feedback devices could lead to shorter task completion time for some 3D manipulation tasks such as 3D alignment (e.g., Scali et al., 2003; Gauldie et al., 2004). However, the mouse-based interfaces are still powerful user interfaces and dominant in the field of medicine. To explore the practical usability of kinesthetic interfaces in these medical fields, it is necessary to compare them with the mouse-based interfaces as the baseline in the real-world medical tasks.



4 Gaze-based Haptic Interaction

Eye gaze has been widely used as a hands-free input mechanism in HCI, including haptic interaction. This chapter presents an overview of gaze-based haptic interaction. It starts from introducing the structure of human eye and various eye movements. Then, the principles of hand-eye coordination are presented, followed by the description of current eye tracking techniques and existing gaze-based haptic interactions.

4.1 HUMAN VISION

Structure of human eye

The sense of vision is the most important sense for perceiving the physical world. Our eyes as the visual sense organ permit perceptions of, for example, light, color and depth. The components of the eyes are mainly the cornea, iris, lens and retina (Oyster, 1999), shown in Figure 4.1.

The cornea is a transparent tissue in the front of the eye that focuses the incoming light and transfers it to the iris. Aqueous humor is a thin transparent fluid located between the cornea and the iris. Its function is to maintain intraocular pressure and transport nutrients to the eye. When the light reaches the iris, it passes through an opening called the pupil. The pupil expands or contracts to control the amount of light entering the eye. After it passes through the pupil, light arrives at the crystalline lens, an important component that can enhance our ability to see the objects at different distances. For example, when an object is far from the eye, the muscles holding the lens expand and flatten the lens to make the eye see the remote object clearly. When an object is close to the eye, the muscles

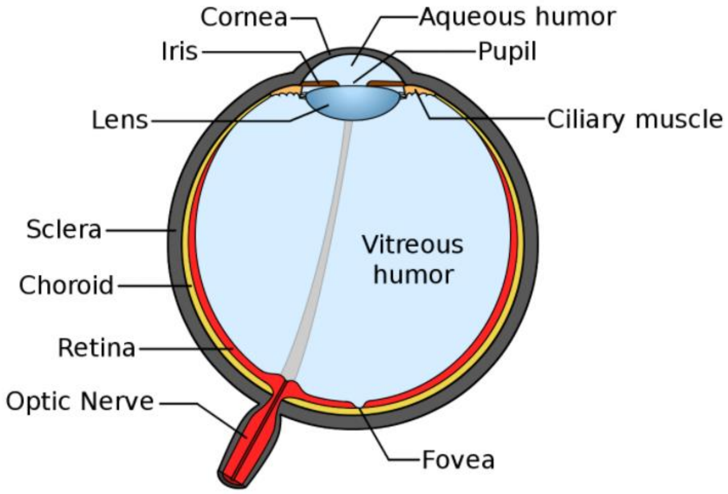


Figure 4.1. Anatomy of the human eye. Public Domain Image, source: Christopher S. Baird.

contract to thicken the lens. Vitreous humor is a fluid that protects the eye and maintains its shape. Light passes through it to the retina. The retina, with other two layers, sclera and choroid, constitutes the basic shape of the human eye. The sclera is an outer white layer that protects the eye, and the choroid in the middle layer has blood vessels to transfer nutrients and oxygen to the retina. The retina is the main light-sensitive tissue, and it contains millions of photoreceptors. There are two types of photoreceptors: rods and cones.

Rod photoreceptors have low visual acuity and color sensitivity. They are useful for monochrome vision in a low-light environment. In contrast, cone photoreceptors have high visual acuity and color sensitivity. A tiny pit directly behind the retina, called the fovea, contains a great number of cone photoreceptors, so it has the highest visual acuity. The fovea region is an area of about 1.5 to 2 degrees at the center of the visual field. Rod photoreceptors are distributed over the rest of the retina, which reduces visual acuity outside the fovea region. When light contacts the rods or the cones of the retina, it generates an electric signal, which is relayed to our brain for visual processing through the optic nerve.

The structure of the human eye is introduced here as a basis for understanding the movements of the eyes and gaze-tracking techniques. The following section describes the eye movements we often perform.

Eye movements

In our daily lives, we perform different types of eye movements, such as saccades and fixation, smooth pursuit movements, vergence movements and vestibula-ocular movements (Purves et al., 2001). This section introduces these eye movements and their characteristics.

Saccades are rapid and ballistic movements that allow us to move the fovea quickly from one point of view to another. We often use saccades to roughly scan a visual scene and bring an object of interest to the foveal region. Saccades are mostly involuntary. A saccade typically lasts 30–120 milliseconds (Jacob, 1995), and its velocity can be as high as 800 degrees/second (Bahill et al., 1975). Saccades also can be elicited voluntarily. For example, when we read, we fixate the words one by one and voluntarily perform saccades for each transition. Furthermore, saccades naturally suppress our visual perception (Matin, 1974). In contrast, the *fixation* of the eyes plays a key role in obtaining visual information. The duration of the fixation varies from milliseconds to a few seconds, based on the tasks (Rayner, 1998). During fixation, the eyes are relatively still, that is, there is little or no obvious movement for the gaze point. This keeps the object of interest in our foveal region and helps us detect visual details.

Smooth pursuit movements are tracking movements that keep a moving object on the fovea. They are much slower than saccades, and their velocity normally depends on the target's velocity (Meyer et al., 1985). If the speed of the target object is too fast for smooth pursuit movements, saccades may act instead to follow the moving object. If there are no moving targets in the environment, the eyes normally do not make smooth pursuit movements (Purves et al., 2001).

Vergence movements allow us to focus on the objects at different viewing distances by moving the two eyes in opposite directions. There are two types of vergence movements: convergence and divergence. Convergence movements rotate the eyes towards each other to focus on close objects, and divergence movements allow the eyes to focus on objects that are further away by rotating the eyes away from each other (Ciuffreda, 2014).

Vestibulo-ocular eye movements help stabilize our vision while the head or body is moving. For example, when we fixate on an object of interest and move our head, our eyes rotate in the opposite direction to compensate for the head movement (Purves et al., 2001). This allows our eyes to stay on the intended target.

From the perspective of HCI, eye gaze is promising as an input modality. For example, using eye gaze as a pointing mechanism has the potential to implement fast interactions because of the saccadic eye movements. Furthermore, the fixation of the eyes can be used as a selection mechanism. Eye gaze is also a potential input modality for kinesthetic interaction because of the coordination between the eyes and hand movements in our daily manual tasks. The following section presents the background information about hand-eye coordination.

4.2 HAND-EYE COORDINATION

Vision and touch are two important human senses, often linked with each other. For example, when we perform manual tasks, such as picking up a cup from a table, opening a sealed box or pushing an elevator button, we naturally use the sense of vision in parallel (Land et al., 1999). During these tasks, the eyes serve two distinct functions. First, they help locate the relevant task objects. After that, they guide the appropriate motor actions (Bowman et al., 2009). Additionally, previous studies have also found that hand movement can guide the saccades of the eyes (Ren and Crawford, 2009). They demonstrated that limb-based proprioception can be transformed into oculomotor coordinates to guide saccades.

In a typical manual task, we often look at the object we want to manipulate or the point of interest we want to touch before we move our hands to reach the object (Land and Hayhoe, 2001; Bowman et al., 2009). Land et al. (1999) accurately measured the time the hand moves after saccadic eye movements. They found that saccades precede the relevant touch-based interaction by 0.56 s. Once we reach the target object, the touch-based interaction we perform involves close hand and eye coordination. Hayhoe and Ballard (2005) and Foulsham (2015) found that our eyes fixate on the relevant objects at critical times during a manual task, based on the characteristics of the task. In addition, the fixation of the eyes often switches to the next object of interest before the touch-based interaction is completed (Foulsham, 2015).

A more detailed review of hand-eye coordination, such as look-ahead fixation (Pelz and Canosa, 2001) and diverse fixation routines (Land and Hayhoe, 2001), is beyond the scope of this dissertation.

Because of the close hand-eye coordination in manual tasks, it is possible to use eye gaze to augment current haptic interfaces. For example, our eye gaze guides motor actions. Thus, eye gaze can be used to replace hand motions as an input modality for haptic interaction. In the next section, gaze-tracking techniques and previous studies of gaze-based haptic interaction are introduced.

4.3 GAZE-BASED HAPTIC INTERACTION

Gaze tracking

To track the eyes, multiple gaze-tracking techniques have been developed. This section introduces two commonly used techniques: electro-oculography and video-oculography.

Electro-oculography is a gaze-tracking technique that measures the difference in electrical potential between the cornea and the retina (Blakley and Chan, 2015). Eye movement commonly leads to a variation of electrical potential transferring to the surrounding skin. This technique uses electrodes attached around the eyes to detect the small difference in electrical potential and thus eye movements. Compared with video-oculography, this approach is easy to use and requires lower computational power. However, this method for gaze tracking yields low accuracy because of interference signals from the surrounding mimic or chewing muscles (Aminoff, 2012).

Video-oculography, the other gaze-tracking technique reviewed here, uses a digital video camera to record the images of the eye and sends them to a computer for image processing (Aminoff, 2012). Based on specific computer algorithms, the positions of different points of the eyes, such as the center of pupil and the corner of the eye, can be identified and located. The orientation of the eyes is calculated based on the positions of these critical points. Further, the direction of the eye gaze can be determined by the horizontal and vertical orientations of the eyes. Video-oculography is more costly, but it is a more accurate gaze-tracking technique compared with the electro-oculography. The eye trackers used in HCI commonly adopt the technique of video-oculography. A more detailed review of both electro-oculography and video-oculography techniques has been provided by Haslwanter and Clarke (2010).

While using eye gaze as an input modality in HCI, the gaze data (i.e., spatial coordinates) generated from the eye tracker are directly related to the quality of interaction (Blignaut and Wium, 2014). There are many factors influencing the quality of the gaze data, such as different eye physiologies, body motions and recording environment, as well as the technique used for eye tracking, such as electro-oculography and video-oculography (Holmqvist et al., 2012). Eye trackers that use video-oculography often employ a gaze-tracking calibration procedure to reduce the effects of the recording environment and the physiologies of users' eyes. The calibration data can help accurately identify and locate the critical points of the eyes and thus improve the quality of gaze tracking. In addition, to measure the data quality, two performance factors have been commonly used in gaze tracking: accuracy and precision. Accuracy refers to the difference between the true gaze point and the gaze point generated by the gaze tracker. Precision refers to how consistent the generated gaze points are when the true gaze point is constant (Holmqvist et al., 2012).

Gaze-based tactile interaction

The studies of tactile interaction in HCI can be categorized into two research areas: tactile output and tactile input. The tactile output studies applied tactile feedback to enhance human-computer interactions (e.g., Bholat et al., 1999), whereas the tactile input studies used natural touch-based motion on tactile input devices as the input modality for interactions (e.g., Yu et al., 2010). Also, there have been studies that focused on both tactile input and output (e.g., Amberg et al., 2011). Gaze-based interaction has been involved in the two tactile research areas.

In the studies with tactile output, researchers explored how the two modalities of tactile and visual feedback can be used effectively in HCI (Rantala et al., 2020). They found that tactile feedback can be as good as visual feedback in gaze-based interactions in terms of task performance and user satisfaction (Kangas et al., 2016; Majaranta et al., 2016), even though vision has usually been considered as the dominant output modality (Ernst and Banks, 2002). Rantala et al. (2020) suggested that besides visual feedback, tactile feedback can be a compensable output modality while interacting with the devices with small visual displays. For example, studies have investigated the effects of tactile feedback in gaze-based events on a variety of small-screen devices, such as mobile phones (Kangas et al., 2014) and smartwatches (Akkil et al., 2015). Kangas et al. (2014) showed that gaze interaction with vibrotactile feedback on a mobile phone has a better task performance than the one without vibrotactile feedback in terms of efficiency and subjective factors. Akkil et al. (2015) showed that vibrotactile feedback is a more noticeable modality than visual feedback for interacting with a smartwatch.

The studies that involved tactile input used eye gaze to enhance tactile interaction. In HCI, eye gaze has been widely considered a fast pointing and selection mechanism (Sibert and Jacob, 2000; Majaranta and Räihä, 2002; Kumar et al., 2007). For example, Zhai et al. (1999) combined eye gaze with mouse movement for the pointing task. Majaranta and Räihä (2002) used the fixation of the eyes to type words on a virtual keyboard rather than using a physical keyboard. Kumar et al. (2007) used eye gaze to point to the target on the screen and the keys on the keyboard for selection. All these techniques could achieve efficient interactions because of using the eye gaze as an input modality.

In the tactile input studies, researchers used eye gaze to improve interaction with tactile input devices like touchscreens (e.g., Stellmach and Dachselt, 2012; Pfeuffer et al., 2014; Pfeuffer et al., 2015). For example, Gaze-Touch developed by Pfeuffer et al. (2014) and Look and Touch developed by Stellmach and Dachselt (2012) are two typical techniques that employed the user's gaze as a pointing and selection mechanism to improve the interaction with the tactile input devices. The technique of Gaze-Touch allowed the user to remotely select objects by gaze without using the hand to reach for the target. Then he or she could locally manipulate the objects by hand gestures on the touchscreen. The technique of Look and Touch used eye gaze for coarsely pointing the target. Then, accurate pointing and selection was done by hand motions on a handheld touchscreen. Two gaze-based techniques are shown in Figure 4.2.

Gaze Touch
(Pfeuffer et al., 2014)

Look and Touch
(Stellmach and Dachselt, 2012)

Figure 4.2. Examples of gaze-based tactile interactions.

The potential of eye gaze to improve tactile interaction has been explored. Eye gaze has also been used for kinesthetic interaction. Existing gaze-based kinesthetic techniques are introduced in the next section.

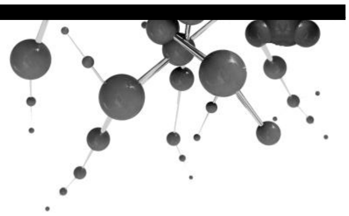
Gaze-based kinesthetic interaction

One of the most important fields where kinesthetic interaction is applied is medicine. Eye gaze has been proposed as a safety mechanism (e.g., Mylonas et al., 2008) or an additional communication channel (Leff et al., 2015) to assist users with medical tasks.

For example, robot-assisted minimally invasive surgery demands the surgeons' manual control because of the complexity of instrument manipulations (Mylonas et al., 2008). To enhance safety in these critical surgical tasks, a technique called Gaze-Contingent Motor Channeling has been proposed to assist surgeons while using a force-feedback device (Mylonas et al., 2008; James et al., 2010; James et al., 2013). Because our vision guides motor actions, they used eye gaze as a safety mechanism to prevent the instrument from inadvertently penetrating the tissue. The safety mechanism was established by applying a force on the instrument towards the eye fixation point. The magnitude of the force depended on the distance between the eye fixation point and the instrument. This technique made the surgeons aware of the incorrect location of the instrument by making them feel the force.

Another kinesthetic study involving eye gaze was done by Leff et al. (2015). Because of the difficulty of remote collaboration during kinesthetic interaction of surgery training, a gaze-based collaborative system was developed for remote partners. The eye gaze of the partners was used as a communication channel to foster collaboration.

These studies, however, used eye gaze only as an independent channel. The potential of eye gaze in kinesthetic interaction remains largely unexplored. Eye gaze has been proposed as a tactile input modality to augment tactile interaction (Stellmach and Dachselt, 2012; Pfeuffer et al., 2014). Using gaze as a kinesthetic input modality is also promising to enhance kinesthetic interaction. However, the potential of eye gaze as the kinesthetic input as well as its effects on task performance such as task completion time, kinesthetic perception and user experience have not been explored. The research of this dissertation is to develop multimodal kinesthetic interfaces by using eye gaze as an input modality to address the issue of limited interaction space and achieve efficient and accurate kinesthetic VR interaction.



5 Methodology, Procedure and Resources

The research in the dissertation was both constructive and empirical. The studies adopted qualitative and quantitative research methods, relying on laboratory experimentation and qualitative questionnaires.

Prototype systems were developed according to the research purposes of the studies. The experiment settings for each prototype system were adjusted in a pilot study with participants who were familiar with the technologies used. Each prototype system was then evaluated in a controlled laboratory environment. Volunteer participants were recruited from the local university community. Research ethics have been considered in all studies. The duration of each experiment and the type of experimental tasks were carefully designed to avoid any mental and physical risk to the participants. Before taking part in the experiment, all participants read and signed an informed consent form. The data from the experiments and the participants' personal information were handled confidentially and used only for research purposes. The names of the participants and other personally identifiable information were removed while reporting the experimental data.

5.1 RESEARCH METHODS

Constructive and empirical research methods were used in the research. Study I designed an experimental system that could be used to examine the effects of different CD gains on kinesthetic interaction. Studies II and III focused on designing and developing new techniques for kinesthetic VR interaction. Study IV involved the development of a prototype system to experimentally compare a kinesthetic VR interface and a vibrotactile VR interface with the traditional mouse-based interface in a medical task.

All studies examined the usability of the proposed interfaces, covering effectiveness, efficiency and user satisfaction (Frøkjær et al., 2000; Issa and Isaias, 2015). Effectiveness refers to the accuracy and completeness with which users achieve specified goals. Studies I and IV were conducted on real-world tasks, and effectiveness was evaluated based on task accuracy (i.e., searching accuracy for Study I and marking accuracy for Study IV). Studies II and III were conducted on a designed task to investigate the effects of the developed interfaces on kinesthetic perception. So, effectiveness was evaluated based on the accuracy of kinesthetic perception (i.e., perception of softness and smoothness).

Efficiency refers to the resources expended in relation to the accuracy and completeness with which users achieved goals. The efficiency of the examined interfaces for all studies was evaluated based on task completion time.

Satisfaction refers to the users' comfort and positive attitudes towards the use of the system. We evaluated satisfaction with the proposed interfaces based on the perceived mental effort, physical fatigue, naturalness and pleasantness. Study I also involved investigating the effects of different CD gains on user confidence about successfully completing the task. Study IV with the VR headset involved the subjective data of immersion.

The main research method was quantitative research based on laboratory experimentation as the data collection method. The prototype systems were designed to automatically record task completion times and accuracy data. Questionnaire were used to record rating data of participants' subjective feelings. Qualitative research method also was used to gather participants' opinions to help clarify their subjective responses. Their comments were collected through post-experiment questionnaires.

5.2 RESEARCH PROCEDURE AND RESOURCES

The research steps for conducting the studies included defining the purpose of the research, reviewing relevant literature, creating functioning prototypes, designing experiments and methods, defining hypotheses, collecting experimental data, performing formal data analysis and drawing conclusions.

Developing the prototype systems required both hardware and software support. A host computer, a Geomagic Touch X force-feedback device, an HTC Vive VR headset and a VR controller, several video-oculography-based Tobii eye trackers[18], a normal 2D display, a standard mouse and a keyboard were used as the experimental hardware for the research. They were available in the TAUCHI Research Center (Tampere Unit for Human-Computer Interaction) at Tampere University. To develop the software systems, Unity[19] with SteamVR[20] was used to connect the VR headset. Tobii SDK[18] was used for eye tracking. H3D and Unity haptic plugin with the OpenHaptics rendering system (Itkowitz et al., 2005) were employed to control the force-feedback device and generate kinesthetic feedback. The programming languages included X3D, C++, C# and Python.

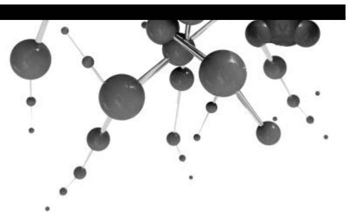
For the three studies that involved gaze tracking (Studies I, II and III), we calibrated the eye tackers for each participant using a five-point or nine-point calibration system integrated with the Tobii SDK. The eye trackers used could track both eyes. We controlled the quality of gaze data by measuring the accuracy and precision of gaze tracking, using an open-source gaze data quality measurement system, TraQuMe (Akkil et al., 2014).

In formal data analysis, we first conducted a normality test, such as the Shapiro–Wilk normality test, to check whether the data were normally distributed. If the data were normally distributed, we analyzed them based on the repeated-measures parametric ANOVA and used a T-test for post-hoc analysis. On the other hand, if the data were not normally distributed, we analyzed them based on the aligned rank transform (ART) repeated-measures non-parametric ANOVA (Wobbrock et al., 2011), and we used the Wilcoxon signed-rank test for post-hoc analysis. If multiple comparisons existed, the Holm-modified Bonferroni correction (Holm, 1979) was used to control the family-wise type-1 error.

In the next chapter, the studies of this dissertation, including their research objectives, methods, results and discussions, are described in detail.

---

[18] https://www.tobii.com/(accessed 1 July 2021)
[19] https://www.unity.com/(accessed 1 July 2021)
[20] https://www.store.steampowered.com/app/250820/SteamVR/(accessed 1 July 2021)



6 Introduction to the Studies

The studies of this dissertation concentrated on hand-based kinesthetic interaction for VR. The research included identifying the issues of the traditional kinesthetic interfaces, developing new interfaces and exploring their applications. Figure 6.1 shows the studies and their research directions.

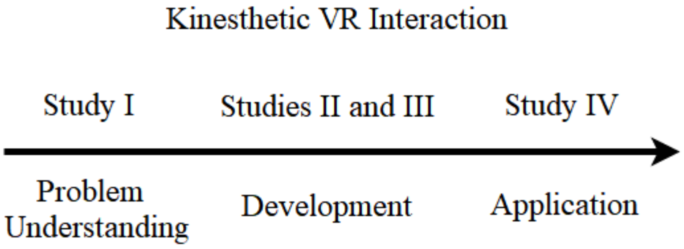


Figure 6.1. The studies organized along a timeline with their research directions.

The following sections summarize the four studies. For each study, the research objective and methodology are introduced, followed by the results and discussion. The main contributions of the research will be discussed in relation to the research questions of the dissertation in Chapter 7.

6.1 STUDY I: EFFECTS OF CONTROL-DISPLAY GAIN IN KINESTHETIC SEARCH

Reference

Zhenxing Li, Deepak Akkil and Roope Raisamo (2020). The Impact of Control-Display Gain in Kinesthetic Search. *In Nisky I., Hartcher-O'Brien J.,*

Wiertlewski M., Smeets J. (eds) Haptics: Science, Technology, Applications. EuroHaptics 2020. Lecture Notes in Computer Science, vol 12272. Springer, Cham. Pages 158-166.
https://doi.org/10.1007/978-3-030-58147-3_18

Objective and method

In HCI, the technique of CD gain has been widely used with many input devices, such as the mouse (Casiez et al., 2008) and the bare-hand gesture input (Poupyrev et al., 1996), for interacting with a large virtual space. In kinesthetic interaction, because the length of the mechanical arm limits the area of hand motions (Massie and Salisbury, 1994), grounded force-feedback devices are difficult to use for exploring a large virtual space. CD gain has been commonly used as the solution to this challenge. However, using a large gain brings a new issue, that is, the mismatch between the hand motion and the HIP motion. Because of this, several studies (e.g., Conti and Khatib, 2005; Dominjon et al., 2005a) have claimed that, for the tasks that require accurate positioning, interacting with the targets is physically challenging using the current scaling-based interfaces. However, in contrast to simple point-and-click tasks using a mouse, kinesthetic interaction involves more complex human behaviors and interaction feedback. To the best of this researcher's knowledge, there have been no experimental studies to investigate the effects of CD gains on kinesthetic task measures such as task completion time, accuracy and user experience.

To fill this gap, this study examined two different CD gains, a unit gain (=1) and a large gain (=3.25). The selection of the large gain was based on the size of the required virtual space and the selection of the unit gain could allow the user's hand motion to march the HIP motion. In the case of the unit gain, the workspace of the force-feedback device was the default 16 x 12 x 12 cm. In the case of the large gain, a 1 cm physical movement of the mechanical arm resulted in a 3.25 cm movement of the HIP along the x-, y- and z- axes. Thus, the workspace was increased to 52 x 39 x 39 cm.

The experimental task was kinesthetic search. Kinesthetic search is a real-world kinesthetic task that requires the user to touch an object and move his or her fingers back and forth along its surface to detect texture, abnormalities or non-uniformities on or under the surface. The participants were asked to employ a force-feedback device with the two CD gains to search lumps underneath a soft tissue. The lumps were of the same size, and they placed randomly inside the tissue and at the same depth. Before the task, the participants had time to familiarize themselves with the size and depth of the lumps. This was to ensure that the participants could identify the force needed to touch the lumps so they could easily complete the task.

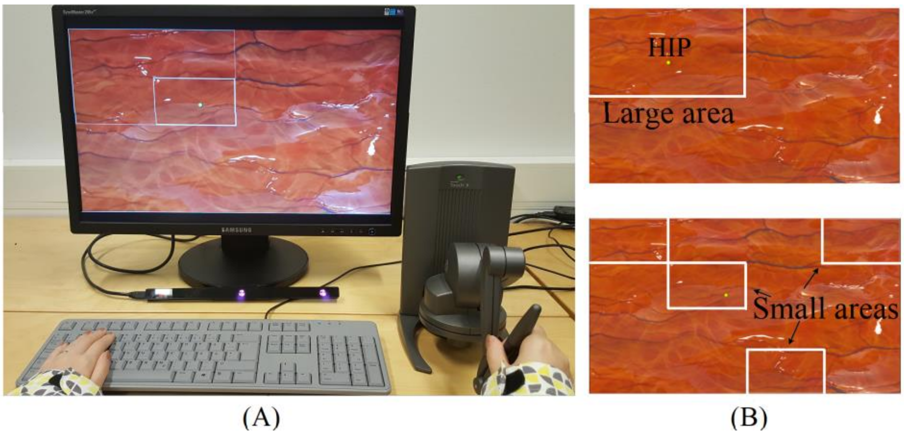


Figure 6.2. (A) shows the experiment environment. (B) shows the area types.

We varied the area of the kinesthetic search (i.e., a large area and multiple small areas) as an independent variable in the experiment, since the effects of CD gain may be influenced by the types of interaction area. The experiment environment is shown in Figure 6.2. Twenty-four participants were recruited from the local university community. We collected both objective and subjective data to examine the effects of the two CD gains. The objective data included the search time, the number of lumps that the participants missed, as well as the kinesthetic search pattern gathered from the data of the movement (along x-, y-, and z-axes) of the HIP during the search. The subjective data included the perceived hand fatigue, naturalness, pleasantness and user confidence in finding all the lumps.

Results and discussion

The results showed that using a large CD gain led to a better interaction speed than using the unit CD gain, regardless of the types of search area. Applying the large gain significantly increased the movement speed of the HIP, allowing users to kinesthetically search for lumps faster. Although the kinesthetic interaction involved complex hand motions and interaction feedback, this result suggested that applying a large gain could lead to a faster kinesthetic interaction. Further, we noted that the search time while using the large gain was influenced by the types of search area, whereas the search time while using the unit gain was not influenced by them. Participants typically adopted a search strategy that involved making long horizontal or vertical sweeping motions. While using the large gain, searching a large area easily made the participants perform fewer sweeping motions. In contrast, searching multiple smaller areas made the participants perform numerous shorter sweeping motions, potentially leading to longer search time. In the case of the unit gain, the participants always performed numerous shorter sweeping motions for searching a large area or small areas, so there was no statistically significant difference in search time between two types of search area.

For the search accuracy, using the large CD gain led the participants to miss more lumps than using the unit CD gain, according to our results. There were two possible reasons for that: First, while using the large gain, the participants might cover less search area than using the unit gain and thus had a higher probability of missing the lumps. Second, the lumps were fixed at the same depth inside the tissue. To find the lumps effectively, participants had to maintain a constant depth of the HIP that could touch the lumps while performing the sweeping motions. A more stable HIP depth presented a higher probability of finding the lumps. Using the large gain might lead to uneven hand stability in the search depth. The results of the kinesthetic search pattern gathered from the data of the HIP movement (along x-, y-, and z- axes) supported these two hypotheses. Thus, we could conclude that using a large CD gain could negatively influence user control and thus task accuracy in kinesthetic interaction. On the other hand, this result indicated the significance of using the unit CD gain for accurate kinesthetic interaction.

In addition, although we did not find differences between the two CD gains in terms of naturalness, pleasantness and the perceived hand fatigue, using the large gain negatively affected the user confidence in finding all lumps, especially in searching a large area.

6.2 STUDY II: GAZE AUGMENTED HAND-BASED KINESTHETIC INTERACTION

Reference

Zhenxing Li, Deepa Akkil and R. Raisamo (2019). Gaze Augmented Hand-Based Kinesthetic Interaction: What You See is What You Feel. *IEEE Transactions on Haptics*, Volume. 12, Issue 2, Pages 114-127. https://doi.org/10.1109/TOH.2019.2896027

Objective and method

Because of the mismatch between the hand motion and the HIP motion, the technique of CD gain used to increase the size of the workspace for force-feedback devices could negatively affect kinesthetic interaction. Furthermore, hand fatigue is another issue for current kinesthetic interfaces (Ott et al., 2005; Hamam and El Saddik, 2015). It could affect user performance, such as spatial positioning and haptic manipulation (Allen and Proske, 2006; Cortes et al., 2014). Study II contributed to this line of research by developing gaze-based kinesthetic interfaces to address the issue of limited workspace for force-feedback devices as well as the issue of hand fatigue from prolonged operation of devices.

We naturally look at the point of interest on the object before we move our hand to touch it (Land et al., 1999; Bowman et al., 2009). Unlike touch behaviors in the real world, which require physically moving the hand to reach and touch the objects, it is possible to use the user's gaze as an input modality for touch-based interaction in HCI. A similar principle has been demonstrated as feasible for tactile interaction using a touchscreen (e.g., Pfeuffer et al., 2014).

This study extended this concept to kinesthetic interaction (see Figure 6.3). The method was to use eye gaze to directly substitute for the components of hand motions as the kinesthetic input. Two new multimodal kinesthetic interfaces, *HandGazeTouch* and *GazeTouch*, have been developed.

HandGazeTouch employed the user's eye gaze to replace the reaching behavior of hand motions during kinesthetic interaction. This means that the HIP would follow the user's gaze point on the target and the user could move his/her hand along the viewing direction (i.e., the z-axis) for touching. *GazeTouch* employed the user's eye gaze to replace both the reaching and touching behaviors of hand motions in kinesthetic interaction. The user could move the HIP by moving the eyes and then stare at the point of interest to trigger touching. For both multimodal interfaces, the user received kinesthetic feedback about the object by holding the mechanical arm of the force-feedback device. To evaluate the two interfaces, we compared them with the traditional hand-based kinesthetic interface (i.e., *HandTouch*) as the baseline in an experiment involving softness discrimination.

The study focused on both interaction efficiency and kinesthetic perception, Fitts's Law is commonly adopted to measure pointing efficiency (e.g., Langolf et al., 1976). However, accurate kinesthetic perception may require participants to touch the same object or compare the two models repeatedly. Another experimental setup has been used, inspired by the Fitts' Law. The experimental task asked participants to touch two soft objects and identify which one was softer. The difficulty of the experimental task depended on two factors: difficulty in reaching the

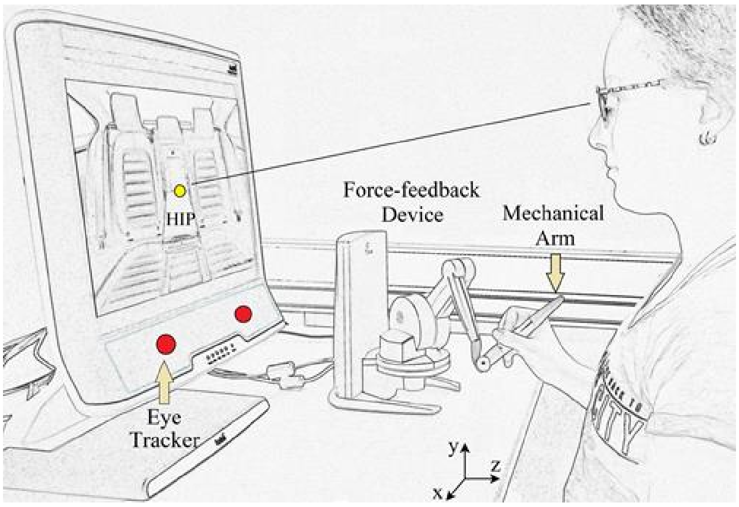


Figure 6.3. Gaze-based kinesthetic interaction.

targets and difficulty in perceiving the difference in softness. We varied the difficulty levels for reaching and perception as independent variables in the experiment (easy and difficult). Twenty-four participants were recruited from the local university community. We collected both objective and subjective results for evaluation. Objective data included the task completion time and the accuracy of softness detection. Subjective data included naturalness and pleasantness as well as the perceived physical and mental difficulties. We gathered the participants' comments about the two gaze-based interfaces after the experiment.

Results and discussion

The results of the study demonstrated the feasibility and usability of the multimodal kinesthetic interfaces. First, based on task completion time, *HandGazeTouch* was faster than *HandTouch* and *GazeTouch*. There was no statistically significant difference between *HandTouch* and *GazeTouch*. The analysis of task completion time with the difficulty levels of reaching showed that the task completion times while using both gaze-based interfaces were not influenced by the difficulty levels of reaching, whereas using *HandTouch* led to longer task completion times when the reaching difficulty was high. This result revealed that *HandGazeTouch* performed better because of the greater speed of saccadic eye moment (Jacob, 1995) to move the HIP. In the case of *GazeTouch*, reaching the target using eye gaze was as fast as *HandGazeTouch*. However, the additional dwell time to trigger kinesthetic cues likely slowed the interaction. Furthermore, the analysis of the difficulty levels of perception revealed that the task completion times of all three user interfaces were influenced by the difficulty levels of perception. The participants generally took more time to complete the tasks when the task perception difficulty was high. However, the increase in task completion times was less for *HandTouch* but substantially higher for the gaze-based interfaces. This situation was likely related to the different accuracy of kinesthetic perception while using three user interfaces.

To evaluate the accuracy of kinesthetic perception, we investigated the number of errors the participants made in discriminating softness. A higher number of errors indicated that the kinesthetic cues from using the interface were harder to interpret by our somatosensory system, and vice versa. The results showed that using *GazeTouch* led to more errors than using *HandTouch*. There were two possible explanations for this. First, the physical motion of the hand to provide haptic cues plays a decisive role in our cognitive process (Lederman and Klatzky, 1987). *GazeTouch* employed eye gaze to fully replace hand behaviors and thus broke the link between hand motions and our somatosensory system. This negatively affected the softness perception. Furthermore, we are better at touch perception using active touch rather than passive touch (Lamotte, 2000). *GazeTouch* triggered the haptic cues by eye gaze, so it belonged to passive touch. This situation further decreased the users' ability to perceive softness while using *GazeTouch*. In contrast, *HandGazeTouch* adopted a compromise design that used eye gaze only for reaching the targets. It continued to use hand motions for the touching behavior. This design method led to the competitive performance of *HandGazeTouch* compared with *HandTouch* in terms of the accuracy of kinesthetic perception.

For user experience, the subjective results showed that using both gaze-based kinesthetic interfaces led to less hand fatigue, but they potentially caused more fatigue on the eyes and increased mental effort. Generally, *HandGazeTouch* was reported to be as natural and pleasant as the traditional kinesthetic interface *HandTouch*.

6.3 STUDY III: GAZE-BASED KINESTHETIC INTERACTION FOR VIRTUAL REALITY

Reference

Zhenxing Li, Deepak Akkil and Roope Raisamo (2020). Gaze-based Kinaesthetic Interaction for Virtual Reality. *Interacting with Computers*, Volume 32, Issue 1, Pages 17–32.
https://doi.org/10.1093/iwcomp/iwaa002

Objective and method

Our previous study developed two gaze-based kinesthetic interfaces by using eye gaze as a kinesthetic input modality to replace hand motions. The approaches successfully addressed the issue of limited workspace for force-feedback devices as well as the issue of hand fatigue. The results of the softness perception experiment showed that using *HandGazeTouch* could maintain a high level of kinesthetic perception accuracy and achieve more efficient interactions compared with the traditional hand-based kinesthetic interface.

However, the previous gaze-based kinesthetic interfaces had several limitations. First, for other kinesthetic tasks, such as identifying the texture and shape of an object, the required touch behaviors were difficult to perform using eye gaze. Second, the two gaze-based interfaces required a consistent use of gaze as the input during kinesthetic exploration. This method limited the user's freedom to allocate visual attention to other objects, and it also led to more eye strain.

This study developed another gaze-based kinesthetic interface, Gaze-Switching Workspace (GSW). Its design method was inspired by the work of Stellmach and Dachselt (2012). The GSW interface adopted the user's gaze as a mechanism to switch the workspace of the force-feedback device (see Figure 6.4), so that the HIP will coarsely reach the target. The accurate reaching and touching behaviors were done by the user's hand motions.

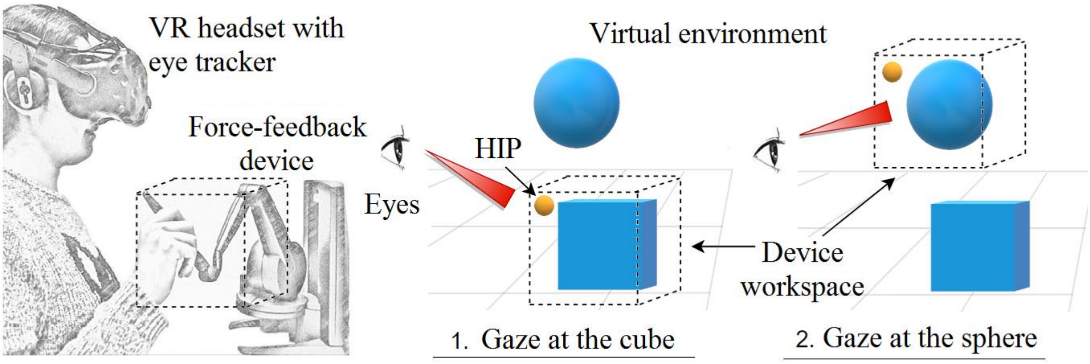


Figure 6.4. Gaze-switching workspace (GSW) kinesthetic interaction.

The process is that, when the user moved the mechanical arm backwards to make the HIP away from all objects and then gazed at a target for a predefined amount of time (500 ms), the device workspace would be moved to the target. The gaze dwell time helped avoid the *Midas Touch* issue (Jacob, 1991). The user would notice the relocation of workspace by observing the changed position of the HIP. Once the workspace locked on the target, the user could employ hand motions to control the HIP for reaching and touching objects. The user had to repeat this switch process to relocate the device workspace. This new gaze-based kinesthetic interface allowed robust and free relocation of the device workspace to any object in the virtual space. Thus, it acquired a large workspace for exploring the virtual environment. Moreover, it maintained hand motions for kinesthetic explorations and used eye gaze only for relocating the device workspace. This likely reduced the intentional use of the eyes.

The GSW interface is developed to address the kinesthetic interaction that requires an interaction space larger than the physical workspace of the force-feedback device. In other words, if the default size of the device workspace is enough for the interaction, the GSW interface works exactly as the traditional hand-based kinesthetic interface. Therefore, to examine the usability of the GSW interface, we compared the GSW interface with the traditional kinesthetic interface in a kinesthetic task that required a large interaction space. In this case, the technique of CD gain was used to increase the device workspace for the traditional kinesthetic interface. The kinesthetic task involved both softness and smoothness discrimination. There were thirty-two participants recruited from the local university community. Like the previous study, the difficulty levels for reaching and perception were used as independent variables (easy and difficult), and objective and subjective data were collected to evaluate the GSW interface.

Results and discussion

This study showed the strengths of the GSW interface when it is used to interact with a large VR environment. First, in both softness and smoothness tasks, using the GSW interface led to statistically significantly shorter task completion times than using the traditional interface. The analysis of the softness task with the reaching difficulty revealed that the better performance of the GSW interface was because of the faster saccadic eye movement (Jacob, 1995) to relocate the device workspace. When the target objects became more difficult to reach, using the traditional interface made the participants spend more time to do the softness task, but using the GSW interface did not.

The faster saccadic eye movement was also the reason of the greater interaction speed of the GSW interface in the task of discriminating smoothness. However, because of the complex hand behaviors needed to perceive the smoothness of an object (i.e., moving back and forth along the object's surface), the time required to reach the objects was likely only a small part of the overall completion time for the smoothness task. Therefore, the analysis with the reaching difficulty did not show a significant result for the smoothness task. In addition, the analysis of the smoothness task with the perception difficulty showed that, while using the GSW interface, the participants spent less time to complete the tasks when the perception difficulty was low. However, while using the traditional interface, there was no statistically significant difference in the task completion times between the levels of task perception difficulty. This result implied that the reduced user control caused by the CD gain, in the case of the traditional interface, might negatively affect kinesthetic perception. Thus, the participants spent much time to do the smoothness tasks even with the low perception difficulty.

For the accuracy of kinesthetic perception, the analysis with the reaching difficulty provided experimental evidences that the reduced user control negatively influenced kinesthetic perception. For both the softness and smoothness tasks, when the objects became more difficult to reach, the participants made statistically significantly more errors in identifying the softness and smoothness while using the traditional interface. While using the GSW interface, there was no statistically significant difference in the number of errors regarding the reaching difficulty levels. Because of the reduced user control caused by the CD gain in the case of the traditional interface, the participants had to move their hand to carefully control the HIP while detecting the softness and smoothness. The strenuous hand motions could have interfered with the hand's ability to perceive material properties, a phenomenon known as *tactile suppression* (Chapman and Tremblay, 2015; Juravle et al., 2017). This situation became worse when the objects became difficult to reach, so the participants committed more errors. While using the GSW interface, the device workspace was relocated to the target object by the user's eye gaze. Regardless of the reaching difficulty, the participants could easily control the HIP and focus on the discrimination tasks. Therefore, the GSW interface could provide more accurate interactions in terms of kinesthetic perception.

In addition, for the smoothness task, the analysis of task errors with the perception difficulty showed that, compared with using the GSW interface, the participants using the traditional interface made a substantially higher number of errors when the perception difficulty was high. Because, besides the reduced user control of the HIP, the hand motions required to detect smoothness (i.e., sliding motions along the surface) was also limited by the used CD gain, this might further lead the participants to make more errors especially in the tasks that required fine-level kinesthetic explorations.

In user experience, the results showed that the GSW interface was perceived to cause less mental effort and hand fatigue, and it was reported to be more natural and pleasant to use than the traditional interface. More importantly, the GSW interface used eye gaze as an input modality, but the users reported that it did not increase eye strain.

6.4 STUDY IV: THE USABILITY OF HAPTIC VIRTUAL REALITY USER INTERFACES IN THE FIELD OF MEDICAL DIAGNOSIS AND PLANNING

Reference

Zhenxing Li, Maria Kiiveri, Jussi Rantala and Roope Raisamo (2021). Evaluation of Haptic Virtual Reality User Interfaces for Medical Marking on 3D Models. *International Journal of Human Computer Studies*, Volume 147, 102561.
https://doi.org/10.1016/j.ijhcs.2020.102561

Objective and method

3D visualization has become a valuable technique in computer-aided medical diagnosis and planning. Volumetric images of human anatomic structures can be created based on computed tomography and magnetic resonance imaging scans (Sutton, 1993). These medical images have been widely used to assist diagnoses (Satava and Robb, 1997) and help surgeons simulate and plan surgical operations (Gross, 1998). For example, medical marking is a fundamental 3D manipulation task performed on volumetric medical images for diagnosis and planning. Medical practitioners manipulate the 3D model (i.e., rotate, pan and zoom) and mark critical points on the model for such things as later inspection, measurement and analysis of skeletal relationships (Kula and Ghoneima, 2018) and planning treatment (Harrell, 2007). The accuracy of the markers directly influences the following medical analyses and thus the overall quality of diagnosis and planning (Lindner et al., 2016).

Figure 6.5. Vibrotactile (left) and kinesthetic (right) VR interfaces.

To interact with these 3D models, current user interfaces in medical systems still rely on 2D interaction technique based on a mouse and a 2D display. However, several studies have argued that using a mouse-based interface for 3D manipulation can be difficult (e.g., Bowman et al., 2004). Researchers have begun to explore other interaction methods for 3D manipulation tasks (e.g., Hinckley et al., 1997; Yu et al., 2010; Besançon et al., 2017). However, these methods either could not outperform the mouse-based interface (e.g., the touchscreen-based interface (Yu et al., 2010)), or they had a limited application area compared with the mouse (e.g., the tangible interface (Hinckley et al., 1997; Besançon et al., 2017)).

In Study IV, we combined a head-mounted VR display with a VR controller and a force-feedback device to create two haptic VR interfaces (i.e., the vibrotactile VR interface using the VR controller and the kinesthetic VR interface using the force-feedback device, shown in Figure 6.5). These interfaces provided the user with a flexible viewing perspective, natural hand-based input and haptic feedback together. Because two interfaces could enable intuitive and realistic 3D interactions, they are promising for the tasks involving 3D manipulation (Bowman et al., 2004), such as the medical diagnosis and planning tasks. To explore their practical usability in this medical field, we compared them with the traditional mouse-based interface as the baseline in an experiment involving medical marking. Such an experiment could help understand the suitability of the interaction methods of two VR interfaces for 3D manipulation and reveal the effects of different types of haptic feedback for this medical task. More importantly, the comparison with the dominant 2D interaction technique could explore the potential of the two haptic VR interfaces to improve current medical diagnosis and planning work.

The experimental task asked the participants to manipulate and mark three medical models using three user interfaces. The marking positions for each model were predefined before the experiment. The two VR interfaces allowed the participants to use hand motions to manipulate the medical models and required them to reach and touch the models for marking. The mouse-based interface required the participants to employ a mouse-based rotate-pan-zoom technique to manipulate the models (i.e., rotation was implemented by using ArcBall technique (Shoemake, 1992) and initiated by pressing the mouse left button, and panning and zooming were done by using the mouse right button and the mouse wheel). It required them to use the mouse pointing for marking, based on the ray-casting technique (Gallo et al., 2010).

We varied the marking positions as an independent variable for the experiment (i.e., the task marking difficulty: easy-to-mark and difficult-to-mark). For the difficult-to-mark tasks, the user needed to manipulate the models carefully to search the marking positions and find an appropriate angle to mark them. For the easy-to-mark tasks, the user could easily find and mark the marking positions. Twenty-four participants were recruited from the local university community. We collected both objective data (task completion time and marking accuracy) and subjective data (the perceived mental effort, hand fatigue, naturalness, immersiveness and user preference) to evaluate the two VR interfaces.

Results and discussion

The study showed the strengths and weaknesses of the two haptic VR interfaces in terms of interaction efficiency, marking accuracy and user experience. In the analysis of task completion time, the comparison between the two VR interfaces showed that using the kinesthetic VR interface led to statistically significantly longer task completion times than using the vibrotactile VR interface, regardless of the task marking difficulty. Because two interfaces used the same HMD as the visual output channel, this result demonstrated the strength of the VR controller for 3D manipulation tasks.

In addition, according to data analysis, the two haptic VR interfaces yielded competitive performances in terms of task completion time compared with the mouse-based interface. A closer analysis of task completion time with the marking difficulty levels showed that using the vibrotactile VR interface led to statistically significantly shorter task completion times than using the mouse-based interface in the difficult-to-mark tasks. However, using the kinesthetic VR interface led to statistically significantly longer task completion times than using the mouse-based interface in the easy-to-mark tasks.

The haptic VR interfaces and the mouse-based interface have their own strengths and weaknesses for medical marking. For the easy-to-mark tasks that did not require complex 3D manipulation, the mouse-based interface performed well in terms of task completion time likely because of the fast pointing capability of the mouse (Kim and Choi, 2019). Two VR interfaces required the participants to reach and touch the target object for marking.

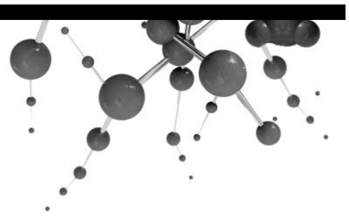
It has been known that reaching an object in the 3D environment is difficult, influenced by the size and position of the target (Berthier et al., 1996). This likely made the haptic VR interfaces, particularly the kinesthetic VR interface, perform less well in terms of task completion time.

For the difficult-to-mark tasks that required complex 3D manipulation, the participants could rotate, pan and zoom the model simultaneously when using the two haptic VR interfaces. In contrast, the mouse-based interface needed the participants to separately use three mouse buttons for each manipulation function. It might have slowed the interaction. Furthermore, using the VR headset likely facilitated the visual search of marking positions. These might bring advantages to the VR interfaces in the difficult-to-mark tasks.

In marking accuracy, the comparison between the kinesthetic VR interface and the vibrotactile VR interface showed the strengths of kinesthetic feedback for accurate medical marking. Regardless of the marking difficulty levels of the tasks, using the kinesthetic VR interface was statistically significantly better than using the vibrotactile VR interface in terms of marking accuracy. There were three possible reasons for this. First, the participants could find and mark the marking positions accurately relying on the kinesthetic feedback. Such an advantage was not available from using the vibrotactile VR interface. Second, the participants could not commit a depth error in the marking in the case of the kinesthetic VR interface, because they were difficult to penetrate the surface of the 3D model due to the kinesthetic feedback. However, the participants easily made depth errors in the marking when using the vibrotactile VR interface. Third, kinesthetic feedback could help maintain hand stability during the touch process and thus improved marking accuracy. In contrast, the issue of hand stability in mid-air negatively affected the marking accuracy of the VR controller.

Moreover, the comparison between the two VR interfaces and the mouse-based interface showed that, for the easy-to-mark tasks, the mouse-based interface and the kinesthetic VR interface performed well in marking accuracy, and using the vibrotactile VR interface led to a statistically significantly lower marking accuracy. For the difficult-to-mark tasks, the mouse-based interface provided the worst performance in marking accuracy among the three interfaces. These results indicated that the marking accuracy using the traditional mouse-based interface was more affected by the marking positions, whereas the two haptic VR interfaces that used natural hand motions with haptic feedback to touch the object for marking could provide more stable marking accuracy regarding the marking positions.

For user experience, the mouse-based interface caused the least hand tiredness. However, the two VR interfaces permitted changes in the head-based viewing perspective, and they allowed hand motions to manipulate objects with haptic feedback. These behaviors were close to our daily behaviors in the physical world, so the participants found them easy to use. It also contributed to that the perceptions of naturalness and immersiveness were better while using the two VR interfaces. Furthermore, compared with the vibrotactile VR interface, the kinesthetic VR interface provided more realistic feelings of touch and benefited the marking process. This made it the most preferred interface for the medical marking task.



7 Discussion

The studies reported in this dissertation constitute a complete set of research in the context of hand-based kinesthetic interaction in VR, including problem understanding, development and application. In this chapter, the main contributions of the studies are discussed in connection with the research questions, and future research directions for kinesthetic VR interaction are proposed.

7.1 RQ1: HOW DOES A LARGE CD GAIN (>1) AFFECT TASK MEASURES SUCH AS TASK COMPLETION TIME, TASK ACCURACY AND USER EXPERIENCE IN KINESTHETIC TASKS?

Study I examined the effects of a large gain (=3.25) and the unit gain (=1) based on a kinesthetic search task. For task completion time, we noted that the large CD gain not only increased the HIP movement but also modulated the users' searching behaviors (i.e., the large gain likely made the user perform fewer sweeping motions). Both effects helped reduce task completion times. As for search accuracy, the results for hand stability in search depth directly demonstrated that the large CD could negatively influence user control of the HIP and thus affect the search accuracy. For user experience, the study found out that applying a large gain made the users lose confidence about completing the task accurately.

This study examined two commonly used CD gains. Technically, the CD gain values between them may not be used because of the unsuitable size of workspace. Our results demonstrated that the high gain negatively affected the HIP depth stability while searching the lumps, even though kinesthetic feedback could help maintain the hand stability while touching. This result indicated that, following the increase of the value of CD gain, this technique could further decrease the stability of the HIP and thus negatively affect user control. From the perspective of this dissertation, this finding is important because it experimentally confirmed the claims made by previous studies (e.g., Conti and Khatib, 2005; Dominjon et al., 2005a) and motivated the following studies of this dissertation.

Moreover, the research provided a preliminary exploration about the effects of CD gain on kinesthetic interaction. Our study employed a kinesthetic search task as the experimental task. The effect of CD gain on task completion time (i.e., a large gain caused shorter task completion time) was consistent with the common effect of CD gain on pointing tasks (Frees and Kessler, 2005; Casiez et al., 2008; König et al., 2009). However, different kinesthetic tasks or a very large gain (i.e., the resulted device workspace is much larger than the required virtual space) may lead to different user performances (e.g., increase the task completion time, like the work by Kwon et al. (2011)). Furthermore, the technique of CD gain could influence subjective measures such as ease of use and preference in some HCI applications (e.g., Kwon et al., 2011). Our results did not show differences in most user factors between the large gain and the unit gain. For the decreased user confidence, the participants likely found that using the large gain gave them less control over the movement of the HIP, so it negatively affected their confidence about successfully completing the task. These indicated that the effects of CD gain in terms of task performance and user experience may depend on the specific context and the selected gain values. The effects of CD gain on kinesthetic interaction should be studied further.

In addition, future studies could explore the effects of other types of CD gain on kinesthetic interaction, such as dynamic gains. Dynamic gains have been used with pointing devices to improve interaction (e.g., Wobbrock et al., 2009; Argelaguet and Andújar, 2013). However, to the best of this researcher's knowledge, the effects of dynamic CD gains on kinesthetic interaction have not been examined. Moreover, besides efficiency, accuracy and user experience, using CD gain could have other effects on kinesthetic interaction. For example, the visual feedback provided by different CD gains can influence the users' weight perception (Dominjon et al., 2005b; Samad et al., 2019). Future studies could investigate other effects of CD gain on kinesthetic interaction.

7.2 RQ2: HOW DO THE MULTIMODAL KINESTHETIC INTERFACES THAT USE EYE GAZE AS AN INPUT MODALITY AFFECT USER PERFORMANCE IN TERMS OF TASK EFFICIENCY, KINESTHETIC PERCEPTION ACCURACY AND USER EXPERIENCE?

To avoid the reduction of user control from using the CD gain technique, Studies II and III used eye gaze as an input modality to implement kinesthetic VR interaction. The methods of using gaze in the two studies were inspired by previous gaze-based methods. Study II was inspired by the method of using eye gaze to point targets without using the hand motion (e.g., Pfeuffer et al., 2014). Two gaze-based kinesthetic interfaces were developed: *HandGazeTouch* and *GazeTouch*. *HandGazeTouch* used the user's gaze to replace the reaching behavior of hand motions in kinesthetic interaction, and *GazeTouch* replaced both the reaching and touching behaviors of hand motions with eye gaze. Both gaze-based kinesthetic interfaces could acquire a large workspace because of the free movement of eye gaze in the virtual world. Study III was inspired by the method of using the eye gaze to coarsely point the target and then using hand motions to accurately point (e.g., Stellmach and Dachselt, 2012). The resulting gaze-based interface, GSW, employed eye gaze to relocate the workspace of the force-feedback device. In this way, it moved the HIP close to the target object. The fine control of the HIP for kinesthetic exploration was implemented by using hand motions.

In terms of task completion time, compared with the traditional hand-based kinesthetic interface, the *HandGazeTouch* had a better performance because it used eye gaze to determine the position of the HIP. This method could quickly move the HIP to any position in the virtual space, so it provided greater interaction speed. Similarly, using the GSW interface could lead to shorter task completion times because it used fast eye movements to relocate the device workspace. This method eliminated the need for explicit hand motions to reach the targets. However, using the *GazeTouch* did not yield shorter task completion times. This was mainly because this gaze-based interface required additional gaze dwell time to trigger the touch behavior. This likely slowed the interaction. Overall, this research demonstrated that using eye gaze as a kinesthetic input modality is feasible and promising, which could largely increase the interaction speed.

As for the accuracy of kinesthetic perception, the research demonstrated that kinesthetic perception in the virtual environment is significantly affected by the kinesthetic interface used. Two design points for kinesthetic interface have been highlighted. First, we noted that our kinesthetic perception relies on the kinesthetic cues and more importantly the hand motions used (Lederman and Klatzky, 1987). This indicated that, to maintain the accuracy of kinesthetic perception, kinesthetic interfaces should allow the user to employ various active hand motions during the touching process. *GazeTouch* did not allow any hand motion and thus performed less well in kinesthetic perception, whereas *HandGazeTouch* still permitted active hand motions for touching and thus maintained a high accuracy of kinesthetic perception. Second, we noted that our somatosensory system has low kinesthetic perception while we are focusing on hand motions (Chapman and Tremblay, 2015). It demonstrated that a good user control of the HIP is critical to achieve better kinesthetic perception. This was the main reason that the GSW interface performed better than the traditional interface in kinesthetic perception.

For user experience, one on hand, both *HandGazeTouch* and *GazeTouch* reduced hand fatigue, but they increased eye fatigue and demanded more mental effort. Since our eyes are primarily perceptual organs, the fine control of the eyes for pointing was difficult while using *HandGazeTouch*, so it increased the eye fatigue and mental effort. *GazeTouch* provided the worst user experience because it required eye gaze for both pointing at the target and triggering kinesthetic cues. However, these negative user experience from using eye gaze may relieve with long-term usage and more practice. For example, Majaranta et al. (2009) reported that novice users suffered from eye fatigue while using gaze-based interfaces because eye movements were felt unnatural, such as long-term fixation of the eyes. But Chitty (2013) showed that experienced users did not report any eye fatigue during the gaze-based interaction.

On the other hand, the GSW interface was overall better than the traditional kinesthetic interface in terms of user experience. The GSW interface enabled robust kinesthetic exploration with fine control of the HIP. In contrast, using the traditional interface led to less user control of the HIP because of the CD gain used. Because of this, the users found it difficult to interact with the objects in a large VR environment, which negatively affected the user experience.

Compared with *GazeTouch* and *HandGazeTouch*, the main advantages of the GSW interface are that it did not fatigue the user’s eyes and it did not increase mental effort. Although all three interfaces used the principle of hand-eye coordination, the different ways they used eye gaze played a critical role in their different user experience. We look at the object we want to manipulate or the point of interest we want to touch before we move our hand to reach for it (Bowman et al., 2009). Based on this natural hand-eye coordination, *GazeTouch* and *HandGazeTouch* used the point of eye gaze as the interaction point. However, two interfaces required an intentional and consistent use of the user’s eyes, and this easily fatigued the eyes. Moreover, the users had to combine eye pointing and hand motions (*HandGazeTouch*) or staring at the target (*GazeTouch*) to trigger the kinesthetic cues, which increased the mental effort required. In contrast, the design of the GSW interface more closely followed natural hand-eye behaviors by using eye gaze only to relocate the workspace and then using hand motions for reaching and touching. In this case, the intentional use of eye gaze was reduced, so there was little eye fatigue. In addition, kinesthetic interaction with objects was done only by hand motions, which reduced the user’s mental workload.

Besides eye fatigue and mental effort, using the gaze point as the touch point, in the cases of *GazeTouch* and *HandGazeTouch*, would lead to the difficulty of performing certain touch behaviors, such as detecting textures and shapes (i.e., movements along the x-and y-axes that require fine control of eye gaze). Also, it caused the difficulty of touching the places that the user could not see (e.g., the sides and back of an object). In contrast, the GSW interface switched the device workspace to cover the target object. Thus, it allowed the user to easily perform various touch behaviors with hand motions, and it also could touch different faces of the object. In addition, the GSW interface permitted users to allocate visual attention to other objects while kinesthetically interacting with an object. Both *GazeTouch* and *HandGazeTouch* required eye gaze to consistently control the HIP, so eye free movements were limited. To address this issue, it may require the HIP to stop following the user's eye gaze during the interaction process. Further development for *GazeTouch* and *HandGazeTouch* interfaces is needed.

On the other hand, *GazeTouch* and *HandGazeTouch* had an advantage that both interfaces could lead to a simplified mechanical structure for the force-feedback device. For these two interfaces, the physical motion of the mechanical arm and the generated kinesthetic feedback were required only along the z-axis. In addition, *GazeTouch* did not require any active hand motion for reaching and touching behaviors. This might make it suitable for specific disabled users, such as the people who lose the ability of active hand movement or cannot do prolonged hand motions.

There is a major limitation to these three gaze-based kinesthetic interfaces, that is, the quality of eye tracking. Both *GazeTouch* and *HandGazeTouch* determined the interaction point by the user's gaze point, which relied heavily on the accuracy and precision of eye tracking. In the experiment, we used the research-quality eye trackers to track the user's gaze, and we calibrated the eye tracker for each participant. TraQuMe (Akkil et al., 2014) was used to ensure a high-quality gaze data (the measured mean gaze offset value for the participants was only 0.62 cm). However, we found that some participants still had difficulties in interacting with objects while using *GazeTouch* and *HandGazeTouch*.

The GSW interface did not show the issue with eye tracking in the experiment. The objects were large and unblocked in terms of visibility, so the quality of gaze data from the used eye tracker was sufficient to accurately switch the device workspace. However, when objects are small or densely distributed in a virtual space, switching the workspace to these objects would also be highly dependent on the accuracy and precision of eye tracking. Fortunately, researchers have started to develop technical solutions for this issue, such as the work by Mardanbegi et al. (2019). They have proposed using depth estimation to solve the issue of partial occlusion of the targets while using eye gaze for pointing and selection. The development of gaze-based tracking technology will make gaze selection more robust in complex VR environments.

To sum up, the development of the three gaze-based kinesthetic interfaces largely explored the design space of using eye gaze as an input modality for kinesthetic interaction. These multimodal interfaces can be compelling choices for future kinesthetic VR interaction. However, this is by no means a complete exploration of the design space of combining gaze and hand motions for kinesthetic interaction. For example, another way to enhance kinesthetic interaction with gaze is to use the gaze to select and move the target object to the workspace of the force-feedback device. Furthermore, interacting with moving objects using the traditional kinesthetic interface could be particularly difficult. Using the eye gaze as the kinesthetic input might be suitable for such interaction tasks because the HIP can remain at the object that has current visual attention. Future studies can explore new gaze-based kinesthetic interfaces and investigate their usability in different scenarios and use cases.

7.3 RQ3: HOW DOES THE KINESTHETIC VR INTERFACE AFFECT USER PERFORMANCE IN MEDICAL 3D MANIPULATION TASKS IN TERMS OF TASK EFFICIENCY, ACCURACY AND USER EXPERIENCE?

In professional fields such as medical diagnosis and planning, interacting with 3D models requires user interfaces with high levels of accuracy and efficiency. However, researchers have noticed that using traditional mouse-based interfaces to interact with 3D models can be difficult and using a multidimensional interaction method could acquire better task performance (Hinckley et al., 1997; Bowman et al., 2004). To contribute to this line of research, Study IV developed two haptic VR interfaces and compared them with the state-of-the-art 2D interface in a medical diagnosis and planning task. The results demonstrated that, the kinesthetic VR interface performed the best in terms of marking accuracy, and its interaction speed was competitive compared with the mouse-based interface. In addition, although the vibrotactile VR interface performed less well in terms of marking accuracy, it performed the best in task completion time among the three user interfaces. Overall, the participants preferred to use the kinesthetic VR interface for this medical task.

Study IV made three major contributions. First, the study experimentally compared two haptic VR interfaces associated with current popular VR equipment and haptic devices. The two VR interfaces used the same HMD but separately employed a VR controller and a force-feedback device as the input tool. Both devices allowed natural hand motions to manipulate 3D objects. However, compared with the free hand motions using the handheld VR controller, the mechanical arm of the force-feedback device, held by the user, not only limited the interaction space but also limited the flexibility of the hand motions that could be used. This likely influenced the manipulation of 3D objects. Thus, using the force-feedback device had longer task completion times than using the VR controller, as seen in the study results.

Second, the study demonstrated the strength of kinesthetic feedback for high-standard medical diagnosis and planning tasks. In previous studies, kinesthetic feedback has been used to enrich interaction channels in surgical simulation and training (Webster et al., 2004; Steinberg et al., 2007; Coles et al., 2011) and to improve surgeons' performance in robot-assisted surgery (Okamura, 2004; Ortmaier et al., 2007). For medical diagnosis and planning, kinesthetic feedback has also been proposed to use in cephalometric marking (Medellín-Castillo et al., 2016). However, it was unclear whether kinesthetic feedback was necessary for this medical 3D manipulation task. Our study directly demonstrated the significance of kinesthetic feedback for accurate medical marking by comparing kinesthetic feedback with vibrotactile feedback. The results showed that using kinesthetic feedback, which allowed users to realistically feel the model, could lead to more accurate marking than using vibrotactile feedback, which provided a simple tactile signal as confirmation of surface contact.

Third, by comparing two haptic VR interfaces with the traditional mouse-based interface, the study demonstrated the practical usability of two haptic VR interfaces for 3D manipulation tasks. Previous studies developed the tangible interfaces (e.g., Hinckley et al., 1997; Besançon et al., 2017) and the touchscreen-based interfaces (e.g., Yu et al., 2010) for 3D manipulation tasks. Tangible interfaces could lead to more efficient interaction than the mouse-based interface (Hinckley et al., 1997; Besançon et al., 2017). However, to manipulate 3D models with different shapes, multiple tangible interfaces must be created. Furthermore, other operations on the 3D model, such as measuring and highlighting an area, might require additional input devices. Thus, the application area of tangible interfaces has been limited. The touchscreen-based interface could be used to perform various operations on 3D models. However, its performance such as interaction speed often could not exceed the performance of the mouse (e.g., Yu et al., 2010).

In contrast, the two haptic VR interfaces proposed in this study could have a wide application area. Further, the vibrotactile VR interface could achieve faster interaction speed than the mouse-based interface in the difficult-to-mark tasks. Surprisedly, the kinesthetic VR interface did not exceed the performance of the mouse-based interfaces in terms of task completion time. Previous studies have demonstrated that kinesthetic interface that used the force-feedback device as the input tool could achieve faster interaction speed than the mouse-based interface in 3D alignment tasks (Scali et al., 2003; Gauldie et al., 2004). However, the medical marking task in our study required the users to manipulate the virtual medical models (i.e., rotate, pan and zoom) and then required them to reach and touch the models for marking. Compared with the mouse-based interface that used the mouse to point for marking, the reaching and touching the model for marking might slow the interaction. However, the competitive performance of the kinesthetic VR interface in terms of task completion time still demonstrated its strength for medical 3D manipulation tasks.

In addition, the analysis of marking accuracy with the two levels of marking difficulty showed the weakness of the traditional mouse-based interface for accurate medical marking. This mouse-based interface commonly uses a ray-casting technique to estimate the point of marking (Gallo et al., 2010). If the surface of the marking point was rotated parallel to the plane of the 2D screen, the mouse-based interface performed well in marking accuracy. However, if the surface of the marking point was not rotated parallel to the screen plane, the angle between the marking surface and the screen plane would make a small displacement error of the mouse pointer on the 2D screen become a large displacement error of the marker on the 3D model. This was the reason that the marking accuracy of the mouse-based interface was more affected by the marking positions, as shown in our results.

Study IV had a few limitations. For example, the medical marking task involved a series of sub-tasks such as rotation, panning and zooming. It is possible that the different experimental conditions we studied provide different levels of cost and benefits for each of the sub-tasks involved. These data were difficult to collect through the medical marking task, because the participants often rotated, panned and zoomed the model simultaneously when using the VR controller and the force-feedback device. More focused future studies are required to answer how the different subtasks affect user performance with different interfaces. In addition, the VR headset with a head movement-based viewing perspective had contributed to the experimental tasks, especially the difficult-to-mark tasks, but it was not fully explored in this study. To further explore the VR headset, a different comparative study would be needed. It would have to make the type of display (2D display and VR headset) an independent variable and employ the same input device as the manipulation tool. Future studies could explore this area.

To sum up, this study has demonstrated that the haptic VR interfaces, particularly the kinesthetic VR interface, have the potential to replace the traditional 2D interface for the high-standard medical marking task. Future studies can investigate their usability in other medical diagnosis and planning scenarios. Furthermore, the study also provided empirical data that can be useful for developing efficient, accurate and user-friendly haptic VR systems. Haptics will become more important in VR technology, and we hope our promising results will encourage future work in this research area.

7.4 OPEN QUESTIONS IN KINESTHETIC VR INTERACTION

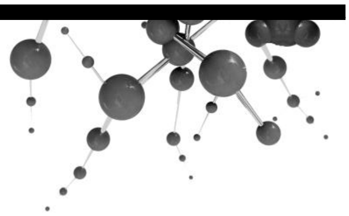
Kinesthetic interaction is an important part of VR. However, current kinesthetic VR interaction has been limited by the interaction devices. In the studies of this dissertation, a grounded force-feedback device (i.e., Touch X) was selected as the kinesthetic interface because of its reliable kinesthetic input and output. Multiple gaze-based kinesthetic interfaces were developed mainly to address the issue of limited workspace for force-feedback devices when they are used to interact with a large virtual space. The resulting interfaces could have a large workspace and allow the user to reach any object in a large virtual environment.

However, kinesthetic force-feedback devices have other limitations. First, the mechanical structure of the device arm also limits the flexibility of the user's hand motions. This negatively affects the interaction speed (demonstrated in Study IV). Second, the device supports only single-point haptic interactions but our everyday interactions in the real world are commonly multiple-point interactions. Furthermore, most current haptic rendering systems used with force-feedback devices to generate kinesthetic feedback still rely on the linear spring law. Realistic touch sensation for virtual objects with complex physical properties is thus difficult to implement. These issues negatively influence the application of kinesthetic devices in VR. This requires further development of kinesthetic techniques in both software and hardware.

In addition, kinesthetic interfaces using grounded devices are suitable for desktop-based VR applications. Using these devices for VR applications that need the user to move physically is a challenge. Current ungrounded kinesthetic devices such as wearable haptic gloves require users to wear bulky and heavy equipment. The low quality of kinesthetic output, the heavy weight and the discomfort while wearing them restrict their applications. Ungrounded kinesthetic devices may still need long-term development. Moreover, ungrounded kinesthetic devices also suffer the issue of a small workspace while using them in a large VR environment because of the length of our arms and the limited room space. Fortunately, gaze-based multimodal interfaces presented in this dissertation could also be suitable for addressing this issue with ungrounded kinesthetic devices. Future studies could investigate this area.

Besides the technical issues of kinesthetic devices, there are other factors that limit the spread and application of kinesthetic VR interaction. First, the required hardware, particularly the kinesthetic device, is expensive for typical computer users. It makes kinesthetic VR interaction still a niche technology used in the laboratory. Second, hand fatigue is an issue for current kinesthetic interfaces. Although the gaze-based interfaces presented in this dissertation can help reduce hand fatigue, we noted that, compared with 2D interaction while using a mouse, 3D manipulation while holding the arm of the kinesthetic device could easily fatigue the user's hand (Study IV). This may require extra arm-rest equipment to relieve the fatigue induced by prolonged operation. Third, kinesthetic interfaces are suitable for some specific use cases, such as medical training and 3D manipulation. For the common computer tasks such as simple dragging and selection, 2D interfaces using the mouse are still the most powerful user interfaces. Further, various software products have been developed to assist the mouse-based interfaces for different computer tasks. Advanced kinesthetic interfaces and the relevant software are required to address this issue.

Kinesthetic VR interfaces provide users with a head movement-based viewing perspective and allow them to employ hand motions to manipulate objects with force feedback. This interaction method is as natural and realistic as our everyday interactions in the physical world. Following the advances of kinesthetic techniques in both hardware and software, kinesthetic VR interfaces have great potential to become popular and powerful interfaces between humans and computers.



8 Conclusion

This dissertation had the goal of exploring hand-based kinesthetic interaction for virtual reality using grounded force-feedback devices. The research was divided into four individual studies. The first study experimentally investigated the issues of the current core technique for implementing kinesthetic interaction in VR. The second and third studies explored using eye gaze as an input modality for designing kinesthetic interfaces. The fourth study explored and highlighted the significance of kinesthetic VR interaction in the field of medicine. The following conclusions can be made for each of these studies:

- CD gain, the technical method currently used with force-feedback devices to implement kinesthetic VR interaction, could negatively influence user control over the HIP and thus reduce task accuracy (Study I).

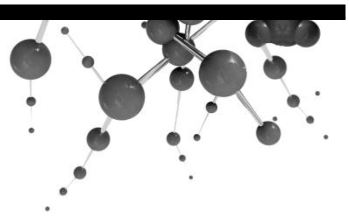
- Substituting the user's eye gaze for hand motions as the input modality for kinesthetic VR interaction is feasible and promising. The resulting multimodal kinesthetic interfaces successfully address the issue of limited workspace for force-feedback devices. The multimodal interface that uses eye gaze for reaching and hand motions for touching could provide more efficient kinesthetic interaction compared with the traditional hand-based kinesthetic interface. Another multimodal interface that uses eye gaze for both reaching and touching objects performed less well in terms of task performance and user experience, but it could be a promising kinesthetic interface for specific disabled users (Study II).

- Combining the user's eye gaze and hand motions as the input modalities for kinesthetic VR interaction is also feasible and promising.

The resulting multimodal kinesthetic interface that employs eye gaze to relocate the workspace of the force-feedback device successfully addresses the issue of limited workspace for force-feedback devices. Overall, this gaze-based interface was better than the traditional hand-based kinesthetic interface in terms of interaction efficiency, kinesthetic perception accuracy and user experience (Study III).

- Haptic VR interfaces implemented by combining a head-mounted VR display with a force-feedback device and a VR controller are promising for computer-aided medical diagnosis and planning. By comparing with the traditional 2D interface that uses a mouse and a 2D display in a medical marking task, the study showed that using the vibrotactile VR interface could lead to the best performance in terms of interaction speed, and using the kinesthetic VR interface led to the best performance in terms of task accuracy. In addition, the kinesthetic VR interface was perceived as being the most suitable user interface for the medical task. This study demonstrated the potential of two haptic VR interfaces to replace the traditional 2D interface for medical diagnosis and planning (Study IV).

Kinesthetic interaction is an essential part of VR technology, and it can support a realistic sense of touch in virtual environments. The studies of this dissertation contributed to this research area by implementing efficient and accurate kinesthetic VR interaction and exploring its application. Future research could focus on the development of kinesthetic techniques in both hardware and software to create robust and user-friendly kinesthetic VR interfaces for HCI.



9 References

Akkil, D., Isokoski, P., Kangas, J., Rantala, J., & Raisamo, R. (2014). TraQuMe: A tool for measuring the gaze tracking quality. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '14),* 327-330. Safety Harbor, FL, USA: ACM. DOI:10.1145/2578153.2578192

Akkil, D., Kangas, J., Rantala, J., Isokoski, P., Špakov, O., & Raisamo, R. (2015). Glance awareness and gaze interaction in smartwatches. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA'15),* 1271-1276. Seoul, Republic of Korea: ACM. DOI:10.1145/2702613.2732816

Alaraj, A., Luciano, C.J., Bailey, D.P., Elsenousi, A., Roitberg, B.Z., Bernardo, A., Banerjee, P.P., & Charbel, F.T. (2015). Virtual reality cerebral aneurysm clipping simulation with real-time haptic feedback. *Neurosurgery*, 11 Suppl 2(0 2), 52-58. DOI:10.1227/NEU.0000000000000583

Alfalah, S.F.M., Falah, J.F.M., Alfalah, T., Elfalah, M., Muhaidat, N., & Falah, O. (2019). A comparative study between a virtual reality heart anatomy system and traditional medical teaching modalities. *Virtual Reality*, 23, 229-234. DOI:10.1007/s10055-018-0359-y

Allen, T.J., & Proske, U. (2006). Effect of muscle fatigue on the sense of limb position and movement. *Experimental Brain Research*, 170(1), 30-38. DOI:10.1007/s00221-005-0174-z

Al-Khalifah, A., McCrindle, R., Sharkey, P.M., & Alexandrov, V.N. (2006). Using virtual reality for medical diagnosis, training and education.

International Journal on Disability and Human Development, 5(2), 187-193. DOI:10.1515/IJDHD.2006.5.2.187

Amberg, M., Giraud, F., Semail, B., Olivo, P., Casiez, G., & Roussel, N. (2011). STIMTAC: A tactile input device with programmable friction. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology (UIST '11 Adjunct*), 7-8. Santa Barbara, CA, USA: ACM. DOI:10.1145/2046396.2046401

Aminoff, M.J. (2012). *Aminoff's Electrodiagnosis in Clinical Neurology,* 6th Edition. USA: Elsevier Saunders.

Argelaguet, F., & Andújar, C. (2013). A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, 37(3), 121-136. DOI:10.1016/j.cag.2012.12.003

Bahill, A.T., Clark, M.R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences,* 24(3-4), 191-204. DOI:10.1016/0025-5564(75)90075-9

Berthier, N.E., Clifton, R.K., Gullapalli, V., McCall, D.D., & Robin, D.J. (1996). Visual information and object size in the control of reaching. *Journal of Motor Behavior*, 28(3), 187-197. DOI:10.1080/00222895.1996.9941744

Besançon, L., Issartel, P., Ammi, M., & Isenberg, T. (2017). Mouse, tactile, and tangible input for 3D manipulation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17),* 4727-4740. Denver, CO, USA: ACM. DOI:10.1145/3025453.3025863

Bholat, O.S., Haluck, R.S., Murray, W.B., Gorman, P.J., & Krummel, T.M. (1999). Tactile feedback is present during minimally invasive surgery. *Journal of the American College of Surgeons,* 189(4), 349-355. DOI:10.1016/s1072-7515(99)00184-2

Bielser, D., & Gross, M.H. (2000). Interactive simulation of surgical cuts. In *Proceedings the Eighth Pacific Conference on Computer Graphics and Applications (PCCGA '00),* 116-125. Hong Kong, China: IEEE. DOI:10.1109/PCCGA.2000.883933

Blakley, B.W., & Chan, L. (2015). Methods considerations for nystagmography. *Journal of Otolaryngology - Head & Neck Surgery,* 44(1), Article 25. DOI:10.1186/s40463-015-0078-2

Blignaut, P., & Wium, D. (2014). Eye-tracking data quality as affected by ethnicity and experimental design. *Behavior Research Methods,* 46(1), 67-80. DOI:10.3758/s13428-013-0343-0

Bowman, D.A., Kruijff, E., LaViola Jr., J.J., & Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice,* 1st Edition. Boston, MA, USA: Addison-Wesley Professional.

Bowman, M.C., Johannson, R.S., & Flanagan, J.R. (2009). Eye-hand coordination in a sequential target contact task. *Experimental Brain Research,* 195(2), 273-283. DOI:10.1007/s00221-009-1781-x

Burdea, G.C., & Coiffet, P. (2003). *Virtual Reality Technology*, 2nd Edition. Hoboken, NJ, USA: Wiley-IEEE Press.

Casiez, G., Vogel, D., Balakrishnan, R., & Cockburn, A. (2008). The impact of control-display gain on user performance in pointing tasks. *Human-Computer Interaction,* 23(3), 215-250. DOI:10.1080/07370020802278163

Chapman, E., & Tremblay, F. (2015). Tactile suppression. *Scholarpedia,* 10(3), 7953. DOI:10.4249/scholarpedia.7953

Chen, M., Mountford, S.J., & Sellen, A. (1988). A study in interactive 3-D rotation using 2-D control devices. In *ACM SIGGRAPH Computer Graphics,* 22(4), 121-129. Atlanta, GA, USA: ACM. DOI:10.1145/54852.378497

Chitty, N. (2013). *User Fatigue and Eye Controlled Technology.* OCAD University. Toronto, Ontario, Canada.

Cho, J., Kim, Y., Jung, S.H., Shin, H., & Kim, T. (2017). 78-4: Screen door effect mitigation and its quantitative evaluation in VR display. *SID Symposium Digest of Technical Papers*, 48(1), 1154-1156. DOI:10.1002/sdtp.11847

Ciuffreda, K.J. (2014). Eye movements; Vergence. In Aminoff, M.J., & Daroff, R.B. (Eds.), *Encyclopedia of the Neurological Sciences,* 2nd Edition, 258-259. San Diego, CA, USA: Elsevier Academic Press. DOI:10.1016/b978-0-12-385157-4.00131-7

Coles, T.R., Meglan, D., & John, N.W. (2011). The role of haptics in medical training simulators: A survey of the state of the art. *IEEE Transactions on Haptics,* 4(1), 51-66. DOI:10.1109/toh.2010.19

Conti, F., & Khatib, O. (2005). Spanning large workspaces using small haptic devices. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC '05),* 183-188. Pisa, Italy: IEEE. DOI:10.1109/WHC.2005.118

Cortes, N.E., Oñate, J.A., & Morrison, S. (2014). Differential effects of fatigue on movement variability. *Gait & Posture*, 39(3), 888-893. DOI:10.1016/j.gaitpost.2013.11.020

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), 64-72. DOI:10.1145/129888.129892

Dinis, F.M., Guimarães, A.S., Carvalho, B.R., & Martins, J.P.P. (2017). Virtual and augmented reality game-based applications to civil engineering education. In *2017 IEEE Global Engineering Education Conference (EDUCON '17)*, 1683-1688. Athens, Greece: IEEE. DOI:10.1109/educon.2017.7943075

Dominjon, L., Lécuyer, A., Burkhardt, J.-M., Andrade-Barroso, G., & Richir, S. (2005a). The "bubble" technique: Interacting with large virtual environments using haptic devices with limited workspace. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC '05)*, 639-640. Pisa, Italy: IEEE. DOI:10.1109/WHC.2005.126

Dominjon, L., Lécuyer, A., Burkhardt, J.-M, Richard, P., & Richir, S. (2005b). Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *IEEE Proceedings. Virtual Reality (VR 2005)*, 19-25. Bonn, Germany: IEEE. DOI:10.1109/VR.2005.1492749

Dooley, K. (2017). Storytelling with virtual reality in 360-degrees: A new screen grammar. *Studies in Australasian Cinema*, 11(3), 161-171. DOI:10.1080/17503175.2017.1387357

El Saddik, A., Orozco, M., Eid, M., & Cha, J. (2011). Haptics technologies: Bringing touch to multimedia. In Avizzano, C.A., Azorín, J.M. (Eds.), *Springer Series on Touch and Haptic Systems*. Berlin Heidelberg, Germany: Springer-Verlag.

Ernst, M.O., & Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature,* 415(6870), 429-433. DOI:10.1038/415429a

Erolin, C., Reid, L., & McDougall, S. (2019). Using virtual reality to complement and enhance anatomy education. *Journal of Visual Communication in Medicine*, 42(3), 93-101. DOI:10.1080/17453054.2019.1597626

Everett, K.A., Exon, R.E., Rosales, S.H., & Gerling, G.J. (2009). A virtual reality interface to provide point interaction and constriction to the finger. In *2009 Systems and Information Engineering Design Symposium,* 203-207. Charlottesville, VA, USA: IEEE. DOI:10.1109/sieds.2009.5166184

Foley, J.D., Van Dam, A., Feiner, S.K., & Hughes, J.F. (1995). *Computer Graphics: Principles and Practice in C,* 2nd Edition. Boston, MA, USA: Addison-Wesley Professional.

Foulsham, T. (2015). Eye movements and their functions in everyday tasks. *Eye,* 29(2), 196-199. DOI:10.1038/eye.2014.275

Frees, S., & Kessler, G.D. (2005). Precise and rapid interaction through scaled manipulation in immersive virtual environments. In *IEEE Proceedings. Virtual Reality (VR 2005),* 99-106. Bonn, Germany: IEEE. DOI:10.1109/VR.2005.1492759

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*, 345-352. The Hague, Netherlands: ACM. DOI:10.1145/332040.332455

Gallo, L., Minutolo, A., & De Pietro, G. (2010). A user interface for VR-ready 3D medical imaging by off-the-shelf input devices. *Computers in Biology and Medicine,* 40(3), 350-358. DOI:10.1016/j.compbiomed.2010.01.006

Gauldie, D., Wright, M., & Shillito, A.M. (2004). 3D modelling is not for WIMPS part II: Stylus/mouse clicks. In *Proceedings of Eurohaptics*, 182-189.

Geldard, F.A. (1960). Some neglected possibilities of communication. *Science*, 131(3413), 1583-1588. DOI:10.1126/science.131.3413.1583

Goldstein, E.B., & Brockmole, J.R. (2016). *Sensation and Perception,* 10th Edition. Boston, MA, USA: Cengage Learning.

Gross, M.H. (1998). Computer graphics in medicine: from visualization to surgery simulation. In *ACM SIGGRAPH Computer Graphics*, 32(1), 53-56. Orlando, FL, USA: ACM. DOI:10.1145/279389.279462

Guttentag, D.A. (2010). Virtual reality: Applications and implications for tourism. *Tourism Management,* 31(5), 637-651. DOI:10.1016/j.tourman.2009.07.003

Hamam, A. & El Saddik, A. (2015). User force profile of repetitive haptic tasks inducing fatigue. In *2015 Seventh International Workshop on*

Quality of Multimedia Experience (QoMEX '15), 1-6. Pilos, Greece: IEEE. DOI:10.1109/QoMEX.2015.7148127

Harrell, W.E. (2007). Three-dimensional diagnosis & treatment planning: The use of 3D facial imaging and 3D Cone Beam CT in orthodontics and dentistry. *Australasian Dental Practice,* 18(4), 102-113.

Haslwanter, T., & Clarke, A.H. (2010). Eye movement measurement: Electro-oculography and video-oculography. *Handbook of Clinical Neurophysiology*, 9, 61-79. DOI:10.1016/s1567-4231(10)09005-2

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences,* 9(4), 188-194. DOI:10.1016/j.tics.2005.02.009

Hinchet, R., Vechev, V., Shea, H., & Hilliges, O. (2018). DextrES: Wearable haptic feedback for grasping in VR via a thin form-factor electrostatic brake. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18),* 901-912. Berlin, Germany: ACM. DOI:10.1145/3242587.3242657

Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., & Kassell, N. (1997). Usability analysis of 3D rotation techniques. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology (UIST '97),* 1-10. Banff, Alberta, Canada: ACM. DOI:10.1145/263407.263408

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics,* 6(2), 65-70.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*, 45-52. Santa Barbara, CA, USA: ACM. DOI:10.1145/2168556.2168563

Howell, J.N., Conatser, R.R., Williams II, R.L., Burns, J.M., & Eland, D.C. (2008). Palpatory diagnosis training on the virtual haptic back: Performance improvement and user evaluations. *Journal of the American Osteopathic Association,* 108(1), 29-36.

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Frontiers in Psychology*, 10, Article 2246. DOI:10.3389/fpsyg.2019.02246

Issa, T., & Isaias, P. (2015). Usability and human computer interaction (HCI). *Sustainable Design,* 19-36. London, UK: Springer. DOI:10.1007/978-1-4471-6753-2_2

Itkowitz, B., Handley, J., & Zhu, W. (2005). The OpenHaptics™ toolkit: A library for adding 3D Touch™ navigation and haptics to graphics applications. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC '05),* 590-591. Pisa, Italy: IEEE. DOI:10.1109/whc.2005.133

Jacob, R.J.K. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems,* 9(2), 152-169. DOI:10.1145/123078.128728

Jacob, R.J.K. (1995). Eye tracking in advanced interface design. In Barfield, W., & Furness, T.A. (Eds.), *Virtual Environments and Advanced Interface Design,* 258-288. New York, NY, USA: Oxford University Press.

James, D.R.C., Leff, D.R., Orihuela-Espina, F., Kwok, K.-W., Mylonas, G.P., Athanasiou, T., Darzi, A.W., & Yang, G.-Z. (2013). Enhanced frontoparietal network architectures following "gaze-contingent" versus "free-hand" motor learning. *Neuroimage,* 64, 267-276. DOI:10.1016/j.neuroimage.2012.08.056

James, D.R.C., Orihuela-Espina, F., Leff, D.R., Mylonas, G.P., Kwok, K.-W., Darzi, A.W., & Yang, G.-Z. (2010). Cognitive burden estimation for visuomotor learning with fNIRS. In Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (Eds.), *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2010). Lecture Notes in Computer Science,* Volume 6363, 319-326. Berlin Heidelberg, Germany: Springer. DOI:10.1007/978-3-642-15711-0_40

Johnsen, E.G., & Corliss, W.R. (1971). Human factors applications in teleoperator design and operation. In Meister, D. (Eds.), *Wiley series in Human Factors.* New York, NY, USA: Wiley-Interscience.

Jones, T., Moore, T., & Choo, J. (2016). The impact of virtual reality on chronic pain. *PLoS ONE,* 11(12), e0167523. DOI:10.1371/journal.pone.0167523

Jung, T., Tom Dieck, M.C., Lee, H., & Chung, N. (2016). Effects of virtual reality and augmented reality on visitor experiences in museum. In Inversini, A., Schegg, R. (Eds.), *Information and Communication Technologies in Tourism 2016*, 621-635. Cham: Springer DOI:10.1007/978-3-319-28231-2_45

Juravle, G., Binsted, G., & Spence, C. (2017). Tactile suppression in goal-directed movement. *Psychonomic Bulletin & Review,* 24(4), 1060-1076. DOI:10.3758/s13423-016-1203-6

Kaas, J.H. (2012). Somatosensory system. *The Human Nervous System,* 1074-1109. DOI:10.1016/b978-0-12-374236-0.10030-6

Kamińska, D., Sapiński, T., Aitken, N., Rocca, A.D., Barańska, M., & Wietsma, R. (2017). Virtual reality as a new trend in mechanical and electrical engineering education. *Open Physics*, 15(1), 936-941. DOI:10.1515/phys-2017-0114

Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R.E., Avots, E., Helmi, A., Ozcinar, C., & Anbarjafari, G. (2019). Virtual reality and its applications in education: Survey. *Information*, 10(10), 318. DOI:10.3390/info10100318

Kamuro, S., Minamizawa, K., & Tachi, S. (2011). An ungrounded pen-shaped kinesthetic display: Device construction and applications. In *2011 IEEE World Haptics Conference (WHC '11)*, 557-562. Istanbul, Turkey: IEEE. DOI:10.1109/whc.2011.5945546

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2014). Gaze gestures and haptic feedback in mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14),* 435-438. Toronto, Ontario, Canada: ACM. DOI:10.1145/2556288.2557040

Kangas, J., Špakov, O., Isokoski, P., Akkil, D., Rantala, J., & Raisamo, R. (2016). Feedback for smooth pursuit gaze tracking based control. In *Proceedings of the 7th Augmented Human International Conference (AH '16),* Article 6, 1-8. Geneva, Switzerland: ACM. DOI:10.1145/2875194.2875209

Kim, H., & Choi, Y. (2019). Performance comparison of user interface devices for controlling mining software in virtual reality environments. *Applied. Sciences*, 9(13), 2584. DOI:10.3390/app9132584

Kim, H., Kim, M., & Lee, W. (2016). HapThimble: A wearable haptic device towards usable virtual touch screen. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16),* 3694-3705. San Jose, CA, USA: ACM. DOI:10.1145/2858036.2858196

King, F., Jayender, J., Bhagavatula, S.K., Shyn, P.B., Pieper, S., Kapur, T., Lasso, A., & Fichtinger, G. (2016). An immersive virtual reality environment for diagnostic imaging. *Journal of Medical Robotics Research,* 01(01), 1640003. DOI:10.1142/S2424905X16400031

Kinnison, T., Forrest, N.D., Frean, S.P., & Baillie, S. (2009). Teaching bovine abdominal anatomy: Use of a haptic simulator. *Anatomical Sciences Education,* 2(6), 280-285. DOI:10.1002/ase.109

Kula, K., & Ghoneima, A. (2018). *Cephalometry in Orthodontics: 2D and 3D,* 1st Edition. Batavia, IL, USA: Quintessence Publishing.

Kumar, M., Paepcke, A., & Winograd, T. (2007). Eyepoint: Practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, 421-430. San Jose, CA, USA: ACM. DOI:10.1145/1240624.1240692

Kwon, S., Choi, E., & Chung, M.K. (2011). Effect of control-to-display gain and movement direction of information spaces on the usability of navigation on small touch-screen interfaces using tap-n-drag. *International Journal of Industrial Ergonomics,* 41(3), 322-330. DOI:10.1016/j.ergon.2011.02.012

König, W.A., Gerken, J., Dierdorf, S., & Reiterer, H. (2009). Adaptive pointing – design and evaluation of a precision enhancing technique for absolute pointing devices. In Gross, T., Gulliksen, J., Kotzé, P. (Eds.), *IFIP Conference on Human-Computer Interaction (INTERACT 2009). Lecture Notes in Computer Science,* Volume 5726, 658-671. Berlin Heidelberg, Germany: Springer. DOI:10.1007/978-3-642-03655-2_73

LaMotte, R.H. (2000). Softness discrimination with a tool. *Journal of Neurophysiology,* 83(4), 1777-1786. DOI:10.1152/jn.2000.83.4.1777

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception,* 28(11), 1311-1328. DOI:10.1068/p2935

Land, M.F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25-26), 3559-3565. DOI:10.1016/S0042-6989(01)00102-X

Langolf, G.D., Chaffin, D.B., & Foulke, J.A. (1976). An investigation of Fitts' law using a wide range of movement amplitudes. *Journal of Motor Behavior*, 8(2), 113-128. DOI:10.1080/00222895.1976.10735061

Lederman, S.J., & Klatzky, R.L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology,* 19(3), 342-368. DOI:10.1016/0010-0285(87)90008-9

Lederman, S.J., & Klatzky, R.L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, 71(7), 1439-1459. DOI:10.3758/APP.71.7.1439

Lee, O., Lee, K., Oh, C., Kim, K., & Kim, M. (2013). Prototype tactile feedback system for examination by skin touch. *Skin Research and Technology,* 20(3), 307-314. DOI:10.1111/srt.12120

Leff, D.R., James, D.R.C., Orihuela-Espina, F., Kwok, K.-W., Sun, L.W., Mylonas, G.P., Athanasiou, T., Darzi, A.W., & Yang, G.-Z. (2015). The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills. *Frontiers in Human Neuroscience,* 9, Article 526. DOI:10.3389/fnhum.2015.00526

Lele, A. (2011). Virtual reality and its military utility. *Journal of Ambient Intelligence and Humanized Computing*, 4, 17-26. DOI:10.1007/s12652-011-0052-4

Lindner, C., Wang, C.-W., Huang, C.-T., Li, C.-H., Chang, S.-W., & Cootes, T.F. (2016). Fully automatic system for accurate localisation and analysis of cephalometric landmarks in lateral cephalograms. *Scientific Reports,* 6, Article 33581. DOI:10.1038/srep33581

Majaranta, P., Ahola, U.-K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09),* 357-360. Boston, MA, USA: ACM. DOI:10.1145/1518701.1518758

Majaranta, P., Isokoski, P., Rantala, J., Špakov, O., Akkil, D., Kangas, J., & Raisamo, R. (2016). Haptic feedback in eye typing. *Journal of Eye Movement Research,* 9(1), 1-14. DOI:10.16910/jemr.9.1.3

Majaranta, P., & Räihä, K.-J. (2002). Twenty years of eye typing: systems and design issues. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications (ETRA '02)*, 15-22. New Orleans, LA, USA: ACM. DOI:10.1145/507072.507076

Maples-Keller, J.L., Yasinski, C., Manjin, N., & Rothbaum, B.O. (2017). Virtual reality-enhanced extinction of phobias and post-traumatic stress. *Neurotherapeutics,* 14(3), 554-563. DOI:10.1007/s13311-017-0534-y

Mardanbegi, D., Langlotz, T., & Gellersen, H. (2019). Resolving target ambiguity in 3D gaze interaction through VOR depth estimation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19),* Paper 612, 1-12. Glasgow, Scotland, UK: ACM. DOI:10.1145/3290605.3300842

Massie, T.H., & Salisbury, J.K. (1994). The PHANToM haptic interface: A device for probing virtual objects. In *Proceedings of the ASME Dynamic Systems and Control Division*, 55(1), 295-301. Chicago, IL, USA.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin,* 81(12), 899-917. DOI:10.1037/h0037368

Medellín-Castillo, H.I., Govea-Valladares, E.H., Pérez-Guerrero, C.N., Gil-Valladares, J., Lim, T., & Ritchie, J.M. (2016). The evaluation of a novel haptic-enabled virtual reality approach for computer-aided cephalometry. *Computer Methods and Programs in Biomedicine,* 130, 46-53. DOI:10.1016/j.cmpb.2016.03.014

Meyer, C.H., Lasker, A.G., & Robinson, D.A. (1985). The upper limit of human smooth pursuit velocity. *Vision Research,* 25(4), 561-563. DOI:10.1016/0042-6989(85)90160-9

Mujber, T.S., Szecsi, T., & Hashmi, M.S.J. (2004). Virtual reality applications in manufacturing process simulation. *Journal of Materials Processing Technology*, 155-156, 1834–1838. DOI:10.1016/j.jmatprotec.2004.04.401

Mylonas, G.P., Kwok, K.-W., Darzi, A.W., & Yang, G.-Z. (2008). Gaze-contingent motor channelling and haptic constraints for minimally invasive robotic surgery. In Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (Eds.), *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2008). Lecture Notes in Computer Science,* Volume 5242, 676-683. Berlin Heidelberg, Germany: Springer. DOI:10.1007/978-3-540-85990-1_81

Nukarinen, T., Kangas, J., Rantala, J., Koskinen, O., & Raisamo, R. (2018). Evaluating ray casting and two gaze-based pointing techniques for object selection in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18),* Article 86, 1-2. Tokyo, Japan: ACM. DOI:10.1145/3281505.3283382

Obrist, M., Velasco, C., Vi, C.T., Ranasinghe, N., Israr, A., Cheok, A.D., Spence, C., & Gopalakrishnakone, P. (2016). Touch, taste, & smell user interfaces: The future of multisensory HCI. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16),* 3285-3292. San Jose, CA, USA: ACM. DOI:10.1145/2851581.2856462

Okamura, A.M. (2004). Methods for haptic feedback in teleoperated robot-assisted surgery. *Industrial Robot: An International Journal,* 31(6), 499-508. DOI:10.1108/01439910410566362

Ortmaier, T., Deml, B., Kübler, B., Passig, G., Reintsema, D., & Seibold, U. (2007). Robot assisted force feedback surgery. In Ferre, M., Buss, M., Aracil, R., Melchiorri, C., Balaguer, C. (Eds.), *Advances in Telerobotics. Springer Tracts in Advanced Robotics,* Volume 31. Berlin Heidelberg, Germany: Springer. DOI:10.1007/978-3-540-71364-7_22

Ott, R., Gutierrez, M., Thalmann, D., & Vexo, F. (2005). Improving user comfort in haptic virtual environments through gravity compensation. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC '05),* 401-409. Pisa, Italy: IEEE. DOI:10.1109/WHC.2005.78

Oyster, C.W. (1999). *The human eye: Structure and function*, Revised Edition. NC, USA: Sinauer Associates.

Pelz, J.B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research,* 41(25-26), 3587-3596. DOI:10.1016/s0042-6989(01)00245-0

Pfeuffer, K., Alexander, J., Chong, M.K., & Gellersen, H. (2014) Gaze-touch: Combining gaze with multi-touch for interaction on the same surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14),* 509-518. Honolulu, HI, USA: ACM. DOI:10.1145/2642918.2647397

Pfeuffer, K., Alexander, J., Chong, M.K., & Zhang, Y. (2015). Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15),* 373-383. Charlotte, NC, USA: ACM. DOI:10.1145/2807442.2807460

Pfeuffer, K., Mayer, B., Mardanbegi, D., & Gellersen, H. (2017). Gaze+pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17),* 99-108. Brighton, UK: ACM. DOI:10.1145/3131277.3132180

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). The Go-Go interaction technique: Non-linear mapping for direct manipulation in VR. In *Proceedings of the 9th annual ACM symposium on User interface software and technology (UIST '96),* 79-80. Seattle, WA, USA: ACM. DOI:10.1145/237091.237102

Pulijala, Y., Ma, M., Pears, M., Peebles, D., & Ayoub, A. (2018). Effectiveness of immersive virtual reality in surgical training—A randomized control trial. *Journal of Oral and Maxillofacial Surgery*, 76(5), 1065-1072. DOI:10.1016/j.joms.2017.10.002

Purves, D., Augustine, G.J., Fitzpatrick, D., Katz, L.C., LaMantia, A.-S., McNamara, J.O., & Williams, S.M. (2001). *Neuroscience,* 2nd Edition. Sunderland, MA, USA: Sinauer Associates.

Radianti, J., Majchrzak, T.A., Fromm, J., & Wohlgenannt, I. (2019). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, 103778. DOI:10.1016/j.compedu.2019.103778

Randall, D., Metherall, P., Bardhan, K.D., Spencer, P., Gillott, R., De Noronha, R., & Fenner, J.W. (2016). The Oculus Rift virtual colonoscopy: Introducing a new technology and initial impressions. *Journal of Biomedical Graphics and Computing,* 6(1), 34-42. DOI:10.5430/jbgc.v6n1p34

Rantala, J., Majaranta, P., Kangas, J., Isokoski, P., Akkil, D., Špakov, O., & Raisamo, R. (2020). Gaze interaction with vibrotactile feedback: Review and design guidelines. *Human–Computer Interaction,* 35(1), 1-39. DOI:10.1080/07370024.2017.1306444

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin,* 124(3), 372-422. DOI:10.1037/0033-2909.124.3.372

Read, A., Ritchie, J., & Lim, T. (2016). A UNITY sketch based modelling environment for virtual assembly and machining to evaluate DFMA metrics. In *Proceedings of the ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 1B: 36th Computers and Information in Engineering Conference*. Charlotte, NC, USA: ASME. DOI:10.1115/detc2016-59202

Ren, L., & Crawford, J.D. (2009). Coordinate transformations for hand-guided saccades. *Experimental Brain Research,* 195(3), 455-465. DOI:10.1007/s00221-009-1811-8

Ribeiro, M.L., Lederman, H.M., Elias, S., & Nunes, F.L.S. (2016). Techniques and devices used in palpation simulation with haptic feedback. *ACM Computing Surveys,* 49(3), Article 48, 1-28. DOI:10.1145/2962723

Robertson, G.G., Card, S.K., & Mackinlay, J.D. (1993). Three views of virtual reality: Nonimmersive virtual reality. *Computer*, 26(2), 81. DOI:10.1109/2.192002

Ros, M., Trives, J.-V., & Lonjon, N. (2017). From stereoscopic recording to virtual reality headsets: Designing a new way to learn surgery. *Neurochirurgie*, 63(1), 1-5. DOI:10.1016/j.neuchi.2016.08.004

Rönkkö, J., Markkanen, J., Launonen, R., Ferrino, M., Gaia, E., Basso, V., Patel, H., D'Cruz, M., & Laukkanen, S. (2006). Multimodal astronaut virtual training prototype. *International Journal of Human-Computer Studies*, 64(3), 182-191. DOI:10.1016/j.ijhcs.2005.08.004

Samad, M., Gatti, E., Hermes, A., Benko, H., & Parise, C. (2019). Pseudo-haptic weight: Changing the perceived weight of virtual objects by manipulating control-display ratio. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19),* Paper 320, 1-13. Glasgow, Scotland, UK: ACM. DOI:10.1145/3290605.3300550

Satava, R.M., & Robb, R.A. (1997). Virtual endoscopy: Application of 3D visualization to medical diagnosis. *PRESENCE: Teleoperators and Virtual Environments*, 6(2), 179-197. DOI:10.1162/pres.1997.6.2.179

Scali, S., Wright, M., & Shillito, A.M. (2003). 3D Modelling is not for WIMPs. In *Proceedings of HCI International.*

Seth, A., Su, H.-J., & Vance, J.M. (2006). SHARP: A system for haptic assembly and realistic prototyping. In *Proceedings of the ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 3: 26th Computers and Information in Engineering Conference*, 905-912. Philadelphia, PA, USA: ASME. DOI:10.1115/detc2006-99476

Shelstad, W.J., Smith, D.C., & Chaparro, B.S. (2017). Gaming on the Rift: How virtual reality affects game user satisfaction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 2072-2076. DOI:10.1177/1541931213602001

Shen, R., Boulanger, P., & Noga, M. (2008). MedVis: A real-time immersive visualization environment for the exploration of medical volumetric data. In *Proceedings of the 2008 Fifth International Conference BioMedical Visualization: Information Visualization in Medical and Biomedical Informatics*, 63-68. London, UK: IEEE. DOI:10.1109/MediVis.2008.10

Shoemake, K. (1992). ARCBALL: A user interface for specifying three-dimensional orientation using a mouse. In *Proceedings of Graphics Interface (GI '92),* 151-156. Vancouver, British Columbia, Canada. DOI:10.20380/GI1992.18

Sibert, L.E., & Jacob, R.J.K. (2000). Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00),* 281-288. The Hague, Netherlands: ACM. DOI:10.1145/332040.332445

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments,* 6(6), 603-616. DOI:10.1162/pres.1997.6.6.603

Sousa, M., Mendes, D., Paulo, S., Matela, N., Jorge, J., & Lopes, D.S. (2017). VRRRRoom: Virtual reality for radiologists in the reading room. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 4057-4062. Denver, CO, USA: ACM. DOI:10.1145/3025453.3025566

Steinberg, A.D., Bashook, P.G., Drummond, J., Ashrafi, S., & Zefran, M. (2007). Assessment of faculty perception of content validity of PerioSim©, a haptic-3D virtual reality dental training simulator. *Journal of Dental Education,* 71(12), 1574-1582. DOI:10.1002/j.0022-0337.2007.71.12.tb04434.x

Stellmach, S., & Dachselt, R. (2012). Look & touch: Gaze-supported target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, 2981-2990. Austin, TX, USA: ACM. DOI:10.1145/2207676.2208709

Sutherland, I.E. (1968). A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68),* 757-764. San Francisco, CA, USA: ACM. DOI:10.1145/1476589.1476686

Sutton, D. (1993). *Textbook of Radiology and Imaging,* 5th Edition. London, UK: Churchill Livingstone.

Takahashi, A., Tanabe, K., & Kajimoto, H. (2019). Haptic interface using tendon electrical stimulation with consideration of multimodal presentation. *Virtual Reality & Intelligent Hardware*, 1(2), 163-175. DOI:10.3724/SP.J.2096-5796.2019.0011

Taylor, J.L. (2016). Kinesthetic inputs. In Pfaff, D., Volkow, N. (Eds.), *Neuroscience in the 21st Century,* 1055-1089. New York, NY, USA: Springer. DOI:10.1007/978-1-4939-3474-4_31

Teklemariam, H.G., & Das, A.K. (2017). A case study of phantom omni force feedback device for virtual product design. *International Journal on Interactive Design and Manufacturing*, 11(4), 881-892. DOI:10.1007/s12008-015-0274-3

Venson, J.E., Albiero Berni, J.C., Edmilson da Silva Maia, C., Marques da Silva, A.M., Cordeiro d'Ornellas, M., & Maciel, A. (2017). A case-based study with radiologists performing diagnosis tasks in virtual reality. *Studies in Health Technology and Informatics,* 245, 244-248. DOI:10.3233/978-1-61499-830-3-244

Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST '05),* 33-42. Seattle, WA, USA: ACM. DOI:10.1145/1095034.1095041

Webster, R.W., Sassani, J., Shenk, R., & Good, N. (2004). A haptic surgical simulator for the continuous curvilinear capsulorhexis procedure during cataract surgery. *Studies in Health Technology and Informatics*, 98, 404-406. DOI:10.3233/978-1-60750-942-4-404

Webster, R.W., Zimmerman, D.I., Mohler, B.J., Melkonian, M.G., & Haluck, R.S. (2001). A prototype haptic suturing simulator. *Studies in Health Technology and Informatics,* 81, 567-569. DOI:10.3233/978-1-60750-925-7-567

Widmer, A., & Hu, Y. (2012). Difference of perceiving object softness during palpation through single-node and multi-node contacts. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '12),* 936-939. San Diego, CA, USA: IEEE. DOI:10.1109/embc.2012.6346086

Williams II, R.L., Srivastava, M., Howell, J.N., Conatser, R.R., Eland, D.C., Burns, J.M., & Chila, A.G. (2004). The virtual haptic back for palpatory training. In *Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI '04),* 191-197. State College, PA, USA: ACM. DOI:10.1145/1027933.1027966

Winter, S.H., & Bouzit, M. (2007). Use of magnetorheological fluid in a force feedback glove. *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* 15(1), 2-8. DOI:10.1109/tnsre.2007.891401

Wirth, M., Gradl, S., Sembdner, J., Kuhrt, S., & Eskofier, B.M. (2018). Evaluation of interaction techniques for a virtual reality reading room in diagnostic radiology. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*, 867-876. Berlin, Germany: ACM. DOI:10.1145/3242587.3242636

Wobbrock, J.O., Findlater, L., Gergle, D., & Higgins, J.J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on*

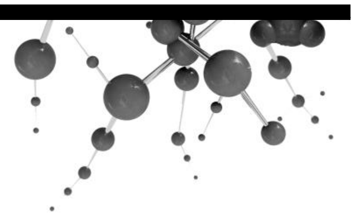
Human Factors in Computing Systems (CHI '11), 143-146. Vancouver, BC, Canada: ACM. DOI:10.1145/1978942.1978963

Wobbrock, J.O., Fogarty, J., Liu, S.-Y., Kimuro, S., & Harada, S. (2009). The angle mouse: Target-agnostic dynamic gain adjustment based on angular deviation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09),* 1401-1410. Boston, MA, USA: ACM. DOI:10.1145/1518701.1518912

Yu, L., Svetachov, P., Isenberg, P., Everts, M.H., & Isenberg, T. (2010). FI3D: Direct-touch interaction for the exploration of 3D scientific visualization spaces. *IEEE Transactions on Visualization and Computer Graphics,* 16(6), 1613-1622. DOI:10.1109/TVCG.2010.157

Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*, 246-253. Pittsburgh, PA, USA: ACM. DOI:10.1145/302979.303053

Zubrycki, I., & Granosik, G. (2017). Novel haptic device using jamming principle for providing kinaesthetic feedback in glove-based control interface. *Journal of Intelligent & Robotic Systems,* 85(3-4), 413-429. DOI:10.1007/s10846-016-0392-6



Paper I

Zhenxing Li, Deepak Akkil and Roope Raisamo (2020). The Impact of Control-Display Gain in Kinesthetic Search. In *Nisky I., Hartcher-O'Brien J., Wiertlewski M., Smeets J. (eds) Haptics: Science, Technology, Applications. EuroHaptics 2020. Lecture Notes in Computer Science*, vol 12272. Springer, Cham. Pages 158-166.
https://doi.org/10.1007/978-3-030-58147-3_18

The Impact of Control-Display Gain in Kinesthetic Search

Zhenxing Li(✉), Deepak Akkil, and Roope Raisamo

Faculty of Information Technology and Communication Sciences,
Tampere University, Tampere, Finland
Zhenxing.li@tuni.fi

Abstract. Kinesthetic interaction typically employs force-feedback devices for providing the kinesthetic input and feedback. However, the length of the mechanical arm limits the space that users can interact with. To overcome this challenge, a large control-display (CD) gain (>1) is often used to transfer a small movement of the arm to a large movement of the onscreen interaction point. Although a large gain is commonly used, its effects on task performance (e.g., task completion time and accuracy) and user experience in kinesthetic interaction remain unclear. In this study, we compared a large CD gain with the unit CD gain as the baseline in a task involving kinesthetic search. Our results showed that the large gain reduced task completion time at the cost of task accuracy. Two gains did not differ in their effects on perceived hand fatigue, naturalness, and pleasantness, but the large gain negatively influenced user confidence of successful task completion.

Keywords: Control-display gain · Force-feedback device · Kinesthetic search

1 Introduction

Kinesthetic interaction as a form of human-computer interaction (HCI) is based on applying force feedback to provide motion sensations in muscles, tendons, and joints [1]. There is an increasing number of kinesthetic applications in different fields, such as education [2], medical training and simulation [3].

Providing realistic force feedback requires dedicated devices such as haptic gloves [4], kinesthetic pens [5] or grounded force-feedback devices (e.g., Geomagic Touch [6]). Among them, force-feedback devices provide a reliable desktop interface with high-resolution forces (up to 1 kHz) [7]. A major limitation of force-feedback devices is that the length of the mechanical arm limits the interaction space [7]. A common solution is to scale a small motion of the mechanical arm to a larger motion of the onscreen haptic interaction point (HIP), i.e., employing a large control-display (CD) gain [8].

The concept of CD gain has been previously studied in the context of pointing devices such as the mouse, touchpad and handheld VR controllers. The results suggest that applying a high CD gain can help reduce task completion time [9, 10]. In the context of kinesthetic interactions, some studies suggested that the visual feedback provided by different CD gains can influence kinesthetic perception and sometimes

I. Nisky et al. (Eds.): EuroHaptics 2020, LNCS 12272, pp. 158–166, 2020.
https://doi.org/10.1007/978-3-030-58147-3_18

even override the perception available through force feedback [11–13]. Further, while using a force-feedback device, applying a large CD gain leads to a mismatch between hand motions and HIP motions, which thus could potentially influence the user's control of the HIP.

Previous studies have used different techniques to enable kinesthetic interactions in large virtual environments without directly using a large CD gain. Dominjon et al. [14] used the bubble technique which adjusts the HIP speed based on the relative positions of the HIP and its bubble to reach objects. Li et al. [15, 16] employed gaze modality to move the HIP for reaching remote targets. Both methods maintained the unit CD gain while touching objects.

Overall, there is an agreement that applying a large gain may influence kinesthetic interactions [14–16]. However, it is still not clear how different CD gains affect task measures such as task completion time, accuracy of interaction and user experience in real-world kinesthetic tasks. In order to fill this gap, we conducted an experiment involving kinesthetic search on a soft tissue. Kinesthetic search is a typical kinesthetic task we perform in the physical world. It requires the users to touch the object and move their fingers along the surface to detect textural and material abnormalities on or under the surface. In a computer-based kinesthetic search task, the user needs to move the HIP while applying appropriate inward force to detect anomalies and the precise control of the HIP is crucial for efficient and accurate interactions.

We evaluated two commonly used CD gains in kinesthetic search: a large CD gain (=3.25) determined by the size of the required virtual space was compared to the baseline unit CD gain (=1). We varied the types of the search area as an independent variable since the effects of the CD gain may be influenced by the interaction area.

We collected objective data (the search time, the number of lumps that the participants missed and the search pattern gathered from the movement data of the HIP) and subjective data (the perceived hand fatigue, naturalness, pleasantness and user confidence in finding all the lumps) to evaluate the two CD gains. The study focused on the below research questions in the context of kinesthetic search:

- Are there differences in the task efficiency and search accuracy using two gains?
- Are there differences in user experience using two gains?

The paper first introduces the experiment, following by the results and discussion.

2 Experiment

2.1 Selection of CD Gains

The explored soft tissue was a cuboid model ($52 \times 32 \times 32$ cm along the x-, y- and z-axes). The model was placed at the center of the virtual space and fully filled the screen of the display. The physical workspace of the force-feedback device used in the experiment was $16 \times 12 \times 12$ cm [6].

The study compared two CD gains (*high* and *default*). The *high* gain was 3.25, determined by the ratio between the tissue size and the device workspace (i.e., 52/16). Thus, a 1 cm arm movement lead to a 3.25 cm HIP movement and the workspace was increased to $52 \times 39 \times 39$ cm which could cover the dimension of the virtual tissue.

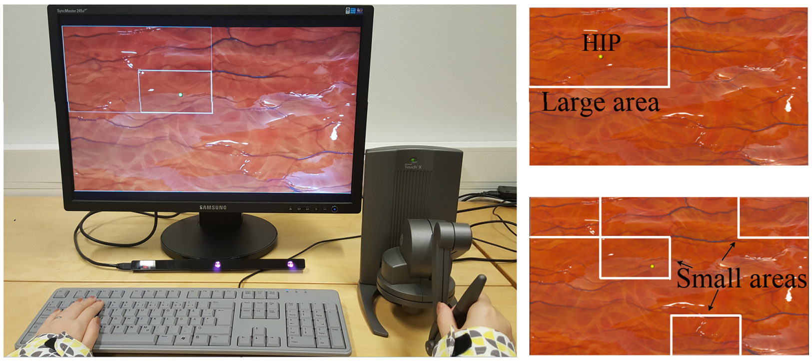
The *default* gain was 1 and thus the workspace was 16 × 12 × 12 cm. To explore the virtual space beyond this workspace, we employed gaze as the section mechanism to relocate the device workspace [16]. The user had to pull the mechanical arm backward to a reset position and gaze at the target area for 500 ms. The workspace would then lock to that area until the user repeated this process. Such a method allowed robust switching of the workspace and ensured that there were no accidental switches during the task. This selected mechanism is not relevant from the perspective of the experiment. All analyses (e.g., search time) pertained only to the period when the user touched the virtual tissue, avoiding any potential influence of this mechanism.

2.2 Experiment Design

A within-subject experiment was designed in a controlled laboratory setting. The task for the participants was to identify the number of lumps underneath a soft tissue.

We manipulated the types of the search area as an independent variable with two levels: four small areas or one large area (Fig. 1(B)). For the large area, the tissue (52 × 32 cm, along the x- and y-axes) was divided into four areas with the size 26 × 16 cm each. The trial of searching the large area included only one area (size = 26 × 16 = 416 cm^2), and four trials as a task group covered the area of the whole tissue. For the small areas, the tissue was divided into 16 small areas with the size 13 × 8 cm each. To make the search size of all trials consistent, one trial of searching the small areas consisted of four randomly selected areas out of the 16 possible options (size = 13 × 8 × 4 = 416 cm^2), and four trials as another task group covered the whole tissue.

The sizes of these areas were selected based on the required search time to avoid a very long experiment. Simultaneously, they were used to examine the effects of the two CD gains in practical applications. The size of each small area was selected so that it could be covered by both workspaces of the two gains. In contrast, the large area could be covered by the workspace using the *high* gain but was beyond the workspace using the *default* gain. The user needed to relocate the workspace four times to fully search the large area (see Fig. 1(A) as an example).



(A) (B)

Fig. 1. (A) shows the experiment environment. The display shows an example of using the *default* CD gain to search a large area. The device workspace with the white boundary switches to the bottom right part of the large area where the user gazes at. (B) shows the area types.

The lump number for each trial was randomly selected from 1 to 4. For each task group with four trials, the total number were 10 (1 + 2 + 3 + 4 = 10). The lumps were sphere models. Since we are interested in examining user control by collecting the movement data of the HIP, all lumps were set as the same radius (0.3 cm) for simplicity. The lumps were randomly distributed (along the x- and y-axes) within the search areas, but placed at the fixed depth (1.5 cm) and were invisible to the participants.

Each participant needed to complete four task groups (2 gains × 2 types of the areas = 4 task groups) with 16 trials (4 task groups × 4 trials per group = 16 trials) and 40 lumps (4 task groups × 10 lumps per group = 40 lumps).

The haptics were developed using H3DAPI with OpenHaptics rendering system [17]. The stiffness of tissue and lumps were implemented by the linear spring law with different stiffness coefficients (tissue: 0.06 and lumps: 0.1) and the friction was implemented by the kinetic friction with the same friction coefficient (both: 0.01). The visual deformation was implemented by the Gauss function, linearly increased following the HIP depth. HIP was visualized as a sphere with 0.3 cm radius.

The participants were asked to input the number of lumps they found using a keyboard after each trial. The system checked and recorded the missing number and the search time. In addition, the system also logged HIP movement data along x-, y- and z-axes during the task. A 7-point Likert scale questionnaire was used to record the subjective data. User confidence were collected after each trial and other subjective data (hand fatigue, naturalness and pleasantness) were collected after each task group.

The participants signed an informed consent form and were asked to complete the tasks as accurately and quickly as possible. They were free to adopt their own strategy (e.g., horizontal and vertical searching) with a maximum of two full searches for each trial. In addition, no extra hand-rest equipment was provided. The order of the CD gains and the area types were counterbalanced among the participants.

2.3 Participants and Apparatus

24 participants were recruited from the local university community (16 women and 8 men), aged between 20 to 35 years (M = 26.17, SD = 4.26). Six participants had used a similar force-feedback device (1–2 times). An MSI GS63VR 7RF laptop was used as the host computer. We used a Samsung 245B monitor as the display, an EyeX [18] to track the gaze, a Touch X device [6] as the kinesthetic interface and a keyboard to input the participants' answer, shown in Fig. 1(A). We employed H3DAPI [17] for haptics and Tobii SDK [18] for accessing the eye tracker.

3 Results

3.1 Objective Data

We first conducted the Shapiro–Wilk Normality test that all data were not normally distributed (all $p < .001$). Thus, we used the 2 × 2 (gains × area types) aligned rank transform (ART) repeated-measures non-parametric ANOVA [19] for the analysis. The Wilcoxon signed-rank test was used for the post hoc analysis. Table 1 shows the overall ART ANOVA results. We focus our analysis on the main effect of CD gains and its significant interaction effect with the area types.

Table 1. Tests of within-subject effects on the objective data (significant values are in bold).

Sources	CD gains			Area types			Interaction effect		
	DF	F	Sig	DF	F	Sig	DF	F	Sig
Search time	1,23	42.07	**<.001**	1,23	5.68	**.026**	1,23	32.53	**<.001**
Missed lumps	1,23	16.33	**.001**	1,23	0.49	.491	1,23	1.89	.183
Covered area	1,23	62.83	**<.001**	1,23	66.58	**<.001**	1,23	78.83	**<.001**
Search depth	1,23	18.95	**<.001**	1,23	0.06	.803	1,23	0.25	.621

Search Time: We calculated the mean search time of four trials in each task group. The results showed that the *high* gain (M = 144.27, SD = 49.48) led to a shorter task completion time than the *default* gain (M = 209.93, SD = 63.86; Z = −4.200, *p* < .001). Figure 2(A) illustrates the interaction effect. In searching the large area, using the *high* gain (M = 112.48, SD = 44.59) led to approximately 47.7% shorter time than using the *default* gain (M = 215.26, SD = 74.22; Z = −4.286, *p* < .001). In searching small areas, using the *high* gain (M = 176.06, SD = 65.26) caused approximately 14.0% shorter time than using the *default* gain (M = 204.60, SD = 67.57; Z = −2.257, *p* = .025).

Missed Lumps: We calculated the sum of the missed lumps for each task group. Figure 2(B) shows that the participants using the *high* gain (M = 1.98, SD = 1.19) missed more lumps than using the *default* gain (M = 0.90, SD = 0.77; Z = −3.426, *p* = .001).

Covered Area: we calculated the proportion of the searched area based on the movement and the radius of HIP. Using the *high* gain (M = 83.29, SD = 4.96) caused searching a smaller area than using the *default* gain (M = 88.39, SD = 4.68; Z = −4.229, *p* < .001). Figure 2(C) shows that using the *high* gain (M = 78.02, SD = 7.37) led to searching a smaller area than using the *default* gain (M = 88.17, SD = 4.91; Z = −4.286, *p* < .001) in searching the large area. There was no difference in searching small areas. A participant's pattern for searching a large area is shown in Fig. 2(E) as an example.

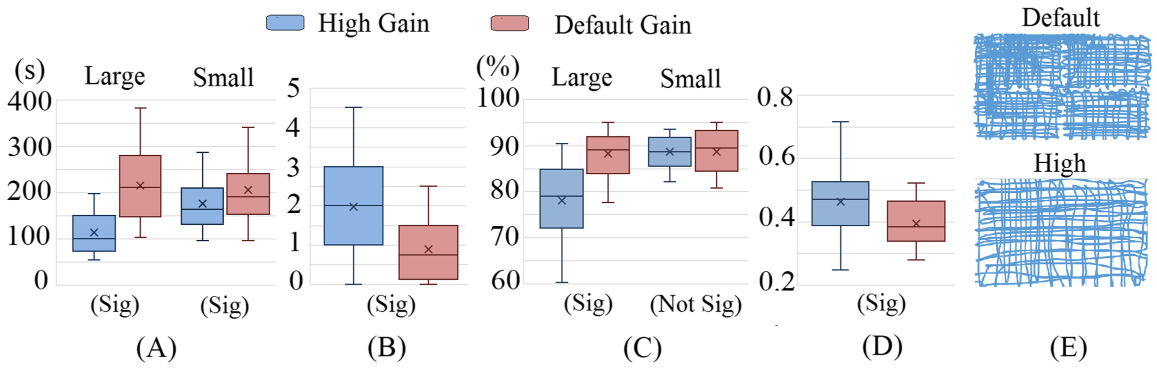


Fig. 2. (A) shows the search times based on the gains and the area types (i.e., large and small); (B) shows the number of the missed lumps based on two gains; (C) shows the area proportion the participants searched based on the gains and the area types; (D) shows the average absolute deviation of the HIP depth based on two gains; (E) shows a participant's pattern for searching a large area. The line in the boxplot is the median value and the cross mark is the mean value.

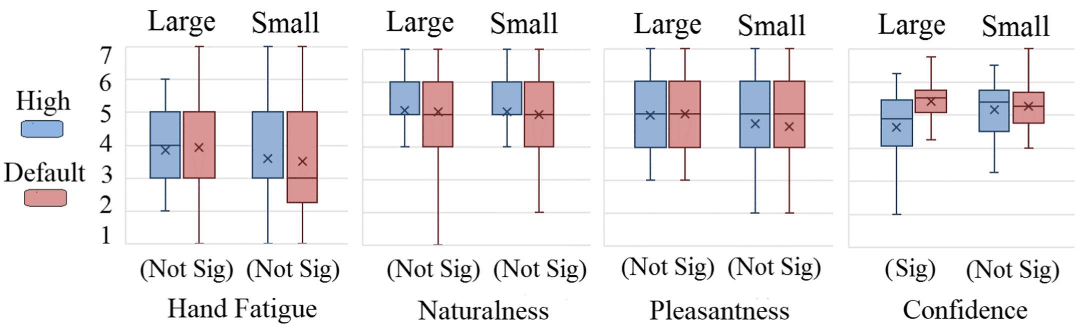


Fig. 3. Subjective results of the study (a higher value is better).

Search Depth: The HIP stability in the search depth may directly affect the search accuracy. To evaluate the stability, we calculated the average absolute deviation value of the HIP data (along z-axis) at the lump depth (1.5 cm) for both gains. Figure 2(D) shows that the *high* gain (M = 0.46, SD = 0.11) caused a lower stability of the HIP than the *default* gain (M = 0.39, SD = 0.07; Z = −3.857, $p < .001$) in the search depth.

3.2 Subjective Data

The data were analyzed with the Wilcoxon signed-rank test (Fig. 3). There were no statistically significant differences between two gains, in terms of perceived hand fatigue, naturalness, pleasantness, as well as user confidence in searching small areas. For the large area, using the *high* gain (M = 4.64, SD = 1.13) led to less confidence than using the *default* gain (M = 5.41, SD = 0.76; Z = −3.312, $p = .001$).

4 Discussion

We experimentally examined the effect of CD gains on kinesthetic search. The results show that CD gains and the area types have significant effects on task performance.

4.1 Differences in Task Completion Time and Search Accuracy

This study focused on the comparison of two different CD gains, where the movement of the device arm led to the different amount of HIP movement (1× and 3.25× respectively). Although the kinesthetic interaction involves complex hand behaviors and interaction feedback, our results show that a large gain increases the movement speed of HIP and thus reduces the task completion time, consistent with the common effect of the CD gain in the pointing tasks using the mouse [9].

However, the search time while using the *high* gain was influenced by the area types (Fig. 2(A)). It can be understood if we consider the search strategy used by our participants. Participants typically adopted a strategy that involved horizontal or vertical sweeping motions (Fig. 2(E)). Searching a large area easily enabled the participants to perform fewer sweeping motions. Searching multiple smaller areas made them perform numerous sweeping motions, potentially leading to longer task times.

While using the *default* gain, irrespective of the area types, participants performed more sweeping motions and thus caused more search time than using the *high* gain.

The results on the search accuracy presents a different picture. Regardless of the area types, using the *high* gain made participants miss more lumps than using the *default* gain (see Fig. 2(B)). There may be two explanations for this phenomenon. First, while using the *high* gain, the participants searched less area than using the *default* gain (Fig. 2(C)). Thus, the participants had a higher probability of missing the lumps using the *high* gain. Second, the lumps were fixed at the same depth inside the tissue. To find the lumps effectively, the participants had to maintain a constant depth of the HIP that could optimally touch the lumps while performing the sweeping motions. A more stable HIP depth presents better probability to find the lumps. Our result (Fig. 2(D)) demonstrated that using the *high* gain causes an increased variability in the HIP depth than using the *default* gain. Previous studies show that hand stability degrades under the stress of the force [20] and fatigue [21]. For the *high* gain, the stability issues may be amplified due to the scaling motion, and thus resulted in lower search accuracy.

4.2 Difference in User Experience

CD gain can potentially affect user experience, such as ease of use and pleasantness, in some HCI applications (e.g., [10]). Surprisingly, we did not find any difference between the two gain conditions in kinesthetic search, in terms of naturalness, pleasantness and hand fatigue. User confidence was influenced by two gains. Participants were generally less confident in finding all lumps while using the *high* gain, specifically while searching a large area. They likely had perceived the limited control over the HIP movement and were aware that they missed many areas. Using the *default* gain made participants more accurate in finding all lumps and subjectively more confident.

4.3 Limitations and Future Studies

This study has a few limitations. First, we examined two commonly used CD gains. Technically, the CD gain values that lie between them are rarely used due to the unsuitable workspace. Two levels (high and low) could sufficiently examine the general effect of the CD gain. However, a very large gain (i.e., the resulted workspace is much larger than the required space size) may cause different user performances (e.g., increase the task completion time, like [10]). Future work may examine this aspect.

Second, we used constant gains along x-, y- and z-axes. Dynamic gains were proposed for the pointing tasks (e.g., [8, 22]). However, their feasibilities for kinesthetic interaction are unknown. Dynamic gains (e.g., velocity-based) may lead to dynamic kinesthetic feedback and affect touch perception. Further, different CD gains could be potentially applied along the different axes. These should be studied further.

Third, the experiment involved a simple cuboid model with a flat surface. Practical applications may include models with irregular shapes and uneven surfaces (e.g., a heart model). The flat surface was a simple model that we could use to examine the effects of CD gains. Future work should test how results differ for complex models.

Fourth, we focused on kinesthetic search, a specific type of kinesthetic interaction. The CD gains may have different effects on different kinesthetic tasks, such as weight perception [11, 13]. Future work could examine CD gains in other kinesthetic tasks.

Fifth, this study included a short-term evaluation with new users. A prolonged usage or recruiting users such as medical professionals who are familiar with kinesthetic search may lead to different results. We propose these for the future research.

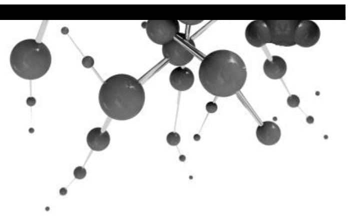
5 Conclusion

This study investigated the effects of CD gains on kinesthetic search. The experiment shows that a large gain improves task efficiency at the cost of user control and thus search accuracy. Our result experimentally demonstrates the significance to maintain the unit CD gain for accurate kinesthetic interaction. In addition, the findings of the study increase theoretical understanding of the CD gain effects on the task performance and user experience, which provide an experimental basis for designing new interaction techniques based on the CD gain for efficient and accurate kinesthetic interaction.

References

1. El Saddik, A., Orozco, M., Eid, M., Cha, J.: Haptics Technologies. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-22658-8
2. Grønbæk, K., Iversen, O.S., Kortbek, K.J., Nielsen, K.R., Aagaard, L.: Interactive floor support for kinesthetic interaction in children learning environments. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 361–375. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74800-7_32
3. Bielser, D., Gross, MH.: Interactive simulation of surgical cuts. In: PCCGA, pp. 116–442 (2000)
4. HaptX. https://www.haptx.com. Accessed 1 Dec 2019
5. Kamuro, S., Minamizawa, K., Tachi, S.: An ungrounded pen-shaped kinesthetic display: device construction and applications. In: WHC, pp. 557–567 (2011)
6. 3D Systems. https://www.3dsystems.com. Accessed 1 Dec 2019
7. Massie, T.H., Salisbury, J.K.: The PHANToM haptic interface. In: DSC, vol. 55, no. 1 (1994)
8. Argelaguet, F., Andújar, C.: A survey of 3D object selection techniques for virtual environments. Comput. Graph. **37**(3), 121–136 (2013)
9. Casiez, G., Vogel, D., Balakrishnan, R., Cockburn, A.: The impact of control-display gain on user performance in pointing tasks. J. CHI **23**(3), 215–250 (2008)
10. Kwon, S., Choi, E., Chung, M.K.: Effect of control-to-display gain and movement direction of information spaces on the usability of navigation on small touch-screen interfaces using tap-n-drag. Ind. Ergon. **41**(3), 322–330 (2011)
11. Dominjon, L., Richard, P., Fre, L.C., Richir, S.: Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In: Proceeding of IEEE VR (2005)
12. Li, M., et al.: Evaluation of pseudo-haptic interactions with soft objects in virtual environments. PLoS One **11**(6), e0157681 (2016)

13. Samad, M., Gatti, E., Hermes, A., Benko, H., Parise, C.: Pseudo-haptic weight: changing the perceived weight of virtual objects by manipulating control-display ratio. In: CHI, Paper 320 (2019)
14. Dominjon, L., Lécuyer, A., Burkhardt, J.M., Barroso, G.A., Richir, S.: The “bubble” technique: interacting with large virtual environments using haptic devices with limited workspace. In: WHC, pp. 639–640 (2005)
15. Li, Z., Akkil, D., Raisamo, R.: Gaze augmented hand-based kinesthetic interaction: what you see is what you feel. IEEE Trans. Haptics **12**(2), 114–127 (2019)
16. Li, Z., Akkil, D., Raisamo, R.: Gaze-based kinaesthetic interaction for virtual reality. Interact. Comput. **32**, 17–32 (2020)
17. H3DAPI. http://www.h3dapi.org. Accessed 1 Dec 2019
18. TOBII. https://www.tobii.com/. Accessed 1 Dec 2019
19. Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J.: The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In: CHI, pp. 143–146 (2011)
20. Borg, G., Sjöberg, H.: The variation of hand steadiness with physical stress. Motor Behav. **13**(2), 110–116 (1981)
21. Gates, D.H., Dingwell, J.B.: The effects of muscle fatigue and movement height on movement stability and variability. Exp. Brain Res. **209**(4), 525–536 (2011). https://doi.org/10.1007/s00221-011-2580-8
22. Wobbrock, J.O., Fogarty, J., Liu, S.Y., Kimuro, S., Harada, S.: The angle mouse: target-agnostic dynamic gain adjustment based on angular deviation. In: CHI, pp. 1401–1410 (2009)



Paper II

Zhenxing Li, Deepak Akkil and Roope Raisamo (2019). Gaze Augmented Hand-Based Kinesthetic Interaction: What You See is What You Feel. *IEEE Transactions on Haptics*, Volume. 12, Issue 2, Pages 114-127.
https://doi.org/10.1109/TOH.2019.2896027

Gaze Augmented Hand-Based Kinesthetic Interaction: What You See Is What You Feel

Zhenxing Li, Deepak Akkil, and Roope Raisamo

Abstract— Kinesthetic interaction between the user and the computer mainly utilizes the hand-based input with force-feedback devices. There are two major shortcomings in hand-based kinesthetic interaction: physical fatigue associated with continuous hand movements and the limited workspace of current force-feedback devices for accurately exploring a large environment. To address these shortcomings, we developed two interaction techniques that use eye gaze as an additional input modality: *HandGazeTouch* and *GazeTouch*. *HandGazeTouch* combines eye gaze and hand motion as the input for kinesthetic interaction, i.e. it uses eye gaze to point and hand motion to touch. *GazeTouch* replaces all hand motions in touch behavior with eye gaze, i.e. it uses eye gaze to point and gaze dwell time to trigger the touch. In both interaction techniques, the user feels the haptic feedback through the force-feedback device. The gaze-based techniques were evaluated in a softness discrimination experiment by comparing them to the traditional kinesthetic interface, *HandTouch*, which only uses the hand-based input. The results indicate that the *HandGazeTouch* technique is not only as accurate, natural, and pleasant as the traditional interface but also more efficient.

Index Terms— Kinesthetic interaction, gaze tracking, hand-eye coordination, force-feedback device, workspace, fatigue.

◆

1 INTRODUCTION

Our everyday interactions in the physical world are inherently multimodal [1]. Touch is one of our most important interaction senses, and it normally utilizes the visual sense in parallel [2]. Humans normally look at their target before a touch-based operation [3], [4] and typical touch manipulation involves close hand-eye coordination. Using the two senses simultaneously to inspect objects can help estimate many important properties. For example, visual cues can be used to know the shape and color of the object and, more importantly, to determine where to touch. Likewise, haptic cues can be employed to identify texture and hardness, for example.

In our typical everyday object manipulation tasks, we tend to look at the object, specifically at the landmark points on the object we will touch, as we initiate our hand movements for reaching. Further, we maintain our gaze on the location as our hands contact the target for touching. The visual feedback provided by our eyes helps guide our hands toward the target and monitor task progress [3], [4], [5]. In this process, the reaching and touching operations are necessary and interlinked because touch is a proximal sense that detects haptic cues from objects close to or in contact with us.

Providing kinesthetic cues in *Human-Computer Interaction* (HCI) follows a similar model to our everyday physical interactions. Virtual objects are modelled by the computer and displayed on the screen, and kinesthetic cues are produced by various force-feedback devices, such as SensAble Phantom [6] and Novint Falcon [7]. These hand-based devices have a mechanical arm that can be moved along three degrees of freedom within the physical workspace [8], and they allow natural hand-based reaching and touching behaviors. For example, a user can control the mechanical arm of the device along the x-y plane to indirectly control the onscreen *Haptic Interaction Point* (HIP) and then push the arm along the z-axis to touch the virtual object and feel the haptic response.

Two major issues limit the usability of current-generation force-feedback devices. First, force-feedback devices have a limited workspace, which practically limits the area of the onscreen objects that can be interacted with. Second, prolonged use of the device is associated with physical fatigue of the hand due to the frequent hand movements required to perform the touch interaction.

Unlike in real-world physical interactions, in HCI it is possible to decouple and replace the reaching and touching operations of hand-based interaction with other input modalities. We naturally look at the object that we are going to physically touch [2]. It is hence possible to use the gaze of the user as a pointing and triggering mechanism to augment traditional kinesthetic interaction.

We developed two interaction techniques that use the gaze of the user as an input. *HandGazeTouch* is an interaction technique that employs gaze input to substitute for the reaching operation. The user simply looks at the point of interest and moves the mechanical arm of the device along the z-axis for touching. *GazeTouch* uses the user's gaze to substitute for both reaching and touching operations. The user looks at a point of interest and stares at it to progressively touch the point. In both techniques, force feedback is felt using the mechanical arm that the user is holding (see Figure 1 for an indicative diagram of the user interaction).

There are multiple motivations for studying gaze in the context of kinesthetic interaction. First, human eyes can move very quickly in comparison to limbs [9]. Using eye movements to substitute for the reaching operation can potentially enable faster kinesthetic interaction, and more importantly it can reduce hand fatigue and the potential for injuries associated with the prolonged operation of force-feedback devices [10], [11].

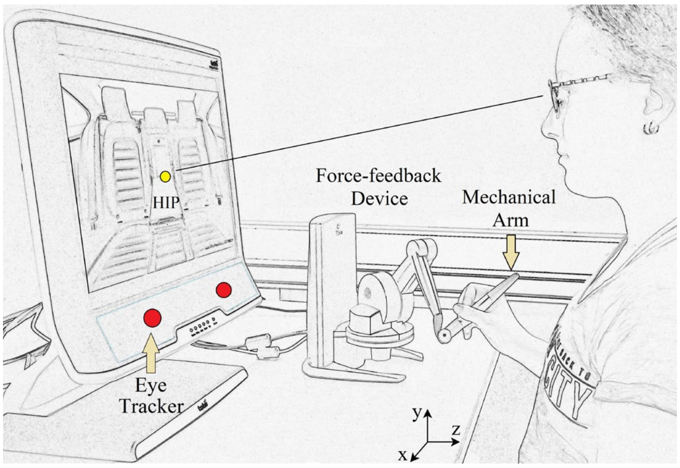


Figure 1: Indicative diagram for gaze-based kinesthetic interaction.

Second, using the eyes as a complementary input modality to control the position of the HIP can help achieve an infinite workspace. Reaching and even touching the target object can be done using the gaze instead of hand movements with the mechanical arm, which overcomes the limited workspace of current kinesthetic interface and has the potential to reduce the build complexity of traditional tabletop force-feedback devices. Third, thanks to technological advancements and a drop in price, eye tracking is no longer a niche technology used in the laboratory; it is widely available in the market [12]. The growing popularity and availability of eye tracking devices makes this a feasible input modality.

Gaze modality has been widely studied in non-haptic applications as a mean to improve user performance for target pointing and selection using a normal 2D display [9], [13], [14], [15] or a Virtual Reality (VR) head-mounted display [16], [17], [18]. In haptic applications involving tactile input devices such as touch screens, eye gaze has been used to improve object acquisition and manipulation, by completely replacing the hand motion [19], [20] or integrating with the hand input [21] to reach targets. Other studies in tactile interaction have investigated user performance while combining gaze input with tactile output [22], [23]. In kinesthetic interaction, gaze modality has been developed as an auxiliary function for solving technical and safety issues in tasks such as robotic surgery [24], [25], [26].

Replacing the reaching component of manual input with eye gaze to improve interaction has been studied for tactile interaction using the touch screen [19], [20]. However, no study extended this concept to kinesthetic interaction and used eye gaze to directly substitute for hand motions of kinesthetic input. Thus, the effects and performance of gaze in this research area have remained unexplored. Our study focused on the following research questions:

- RQ1: Is gaze a feasible input to substitute for hand-based reaching and touching operations in kinesthetic interaction?
- RQ2: How does the combination of gaze and hand modalities influence task performance in terms of efficiency?
- RQ3: Is there an effect on the accuracy of touch feeling, or are there other "side effects" to human perception?

We conducted an experimental study to evaluate two gaze-based interaction techniques, with the conventional hand-based technique (*HandTouch*) as the baseline. We compared the three interaction techniques in a softness discrimination test and focused on objective results, such as the efficiency of task completion and the accuracy of softness detection. In addition, we also recorded the participants' subjective responses, such as naturalness, pleasantness, and physical and mental difficulties, through a questionnaire and interviews.

Our research has the following key novelties:

- The study explored the design space of combining eye gaze and hand motions as the input modality in kinesthetic interaction by developing two gaze-based interaction techniques (*HandGazeTouch* and *GazeTouch*).
- We experimentally compared *HandGazeTouch* and *GazeTouch* to the conventional kinesthetic interface that only uses hand-based input (*HandTouch*).
- Our study provides further theoretical understanding of human kinesthetic perception in identifying softness, and it extends a previous physical study [27] to a virtual environment using a force-feedback device.

The organization of the article is as follows: we first introduce the relevant previous studies in this research area, and then describe the two new interaction techniques (*GazeTouch* and *HandGazeTouch*) in greater detail. This is followed by the details of the experiment and the results. We finish with the discussion and practical implications of our results.

2 BACKGROUND

2.1 Hand-eye Coordination

In terms of hand-eye coordination, the eyes serve two distinct functions: locating the relevant task objects and guiding the appropriate motor actions [3]. The eyes and hands work in close synchrony in our everyday physical tasks. Foulsham [28] has noted that our eye movements are modulated by task characteristics, and the eyes fixate on the relevant objects at critical time points during the task. Bowman et al. [3] have observed that in tasks that require object manipulation, we look at the specific location of touch points before the contact happens. Land et al. [2] have found that human eyes fixate on the target object roughly half a second before its manipulation using the hand.

Similarly, previous studies have noted the systematic coordination between the hand and eyes when using indirect pointing devices such as a mouse [29]. However, few previous studies have leveraged natural hand-eye coordination to improve the kinesthetic interaction, and this is the focus of the current work.

2.2 The Limitations of Current Kinesthetic Interaction

As mentioned before, the mechanical arm of a force-feedback device not only allows the user to freely navigate a 3D environment, but also transfers force and torque to simulate the feeling of touch. However, the workspace of current force-feedback devices is limited by the length of the mechanical arm. Massie and Salisbury [8] have noted that a desktop force-feedback device such as Phantom only has a small wrist-centered workspace and the forearm is allowed only limited movement. Several studies [30], [31] have argued that for tasks requiring accurate positioning in a large environment, reaching the target is physically challenging using the current kinesthetic interface. To overcome the problems caused by the limited workspace of force-feedback devices, Conti and Khatib [30] have proposed the *Workspace Drift Controller*, which progressively centers the physical workspace of the device during the interaction. Dominjon et al. [31] have proposed *Bubble*, which utilizes a spherical area around the HIP with a hybrid position and rate control for accurate and efficient haptic exploration in a large virtual environment.

Physical fatigue is another important issue in interactions that involve repetitive physical actions. Previous studies have shown that multiple factors cause fatigue in the use of force-feedback devices. For example, Ott et al. [10] have found that muscular fatigue could be caused by uncomfortable postures when using force-feedback devices. Hamam and Saddik [11] have demonstrated that repetitive kinesthetic tasks with a larger force or higher distance result in greater user fatigue.

Physical fatigue not only influences the user's experience but also negatively affects performance during interaction. Allen and Proske [32] have shown that muscle fatigue disturbs our sense of position and makes users commit more errors in estimating spatial positions. Cortes et al. [33] have demonstrated that the ability of users to perform a smooth and controlled physical action is influenced by physical fatigue. It is therefore important for designers of kinesthetic interfaces to strive to reduce user fatigue.

Another intuitive method to extend the workspace is to use gaze as an interaction modality. We normally look at the objects we are trying to touch [2]. We can practically achieve an infinite workspace by dynamically redefining the position of the HIP to the location of the gaze, and simultaneously hand fatigue may be reduced by replacing hand motions with eye gaze.

In addition, mental effort is not correlated with muscle movements, so most studies, such as [10], [11], had less concern for mental effort in haptic manipulation. Mental effort is considered in this study. We used both perceived physical and mental difficulty as measures to evaluate the development of the kinesthetic interaction techniques.

2.3 Tactile Interaction with Eye Gaze as An Input Modality

Tactile interaction focuses on applying cutaneous sensation to human-computer interactions [34]. The related studies with eye gaze can be categorized into two research areas: tactile output (e.g. vibration) and tactile input (e.g. touchscreen).

In the studies involving tactile output, some studies have addressed the value of eye gaze with vibrotactile feedback in a variety of devices, such as computers, mobile phones, and smartwatches [35]. Kangas et al. [22] have shown that gaze interaction with vibrotactile feedback increases the efficiency of interaction. Akkil et al. [23] have noted that vibrotactile feedback is a clearer and more noticeable modality for gaze events than visual feedback in small-screen devices such as smartwatches.

With tactile input devices, many have attempted to combine eye gaze with hand gestures onscreen as the input to improve the functions of object manipulation. Pfeuffer et al. [19] have developed *Gaze-touch*: an interaction technique that uses eye gaze for remote selection and touch gestures onscreen for object manipulation. This method decouples the reaching component from the hand-based input and replaces it with eye gaze, so users can remotely manipulate the objects without using their hands to reach for the target. In their following study [20], they developed *Gaze-shifting*: a generic mechanism for switching direct and indirect input modes in touch-based interaction based on the relative position of the gaze location and the touch operation. In addition, Stellmach and Dachselt [21] have developed *Look and Touch*, which integrates eye gaze with manual input instead of completely replacing it. Eye gaze is employed for coarse reaching, and the user selects the target by fine hand-based reaching, similar to the method used in Magic pointing [9].

Eye gaze as an input modality has been widely considered as a fast pointing and selection mechanism [13], [14], [15], and Pfeuffer et al.'s work [19] has demonstrated that using gaze to replace hand motions as the input is promising for tactile interaction. Our study extends this concept to kinesthetic interaction.

2.4 Kinesthetic Interaction with Eye Gaze as An Input Modality

Kinesthetic interaction is another important branch of haptic interaction, concentrating on motion sensations originating in the muscles, tendons, and joints [34]. Force-feedback devices (e.g. the Phantom device) are commonly used as haptic interfaces for kinesthetic interaction. Their mechanical arms often have three or six degrees of freedom [8], which is suitable for hand motions as the kinesthetic input, and they simultaneously transfer the haptic cues of virtual objects as kinesthetic output to the hand. Eye gaze has been considered an auxiliary input channel for kinesthetic interaction, e.g. to enhance safety in critical surgical tasks [24], [25], [26] or foster remote collaboration [36].

Previous studies [24], [25], [26] have employed a technique called *Gaze-Contingent Motor Channeling* (GCMC) for robotic surgery, which uses the gaze to set safety boundaries for the HIP in order to prevent the instrument from inadvertently penetrating the tissue during surgery. The safety boundaries were established by employing a spring force on the HIP toward the eye fixation point, with a magnitude based on the distance between the eye fixation point and the HIP.

Another study combining gaze and force-feedback devices was conducted by Leff et al. [36]. They developed a collaborative system that provides gaze awareness between remote partners during kinesthetic interaction. They used gaze as a channel to foster collaboration.

The focus of our study is to understand the utility of gaze in the reaching and touching operations of kinesthetic interaction. Can gaze be used to replace the reaching operation? Can gaze be used as a mechanism to initiate the touching operation? What are its effects on human kinesthetic perception? The objective of our study is to answer these questions.

3 INTERACTION TECHNIQUES AND RESEARCH

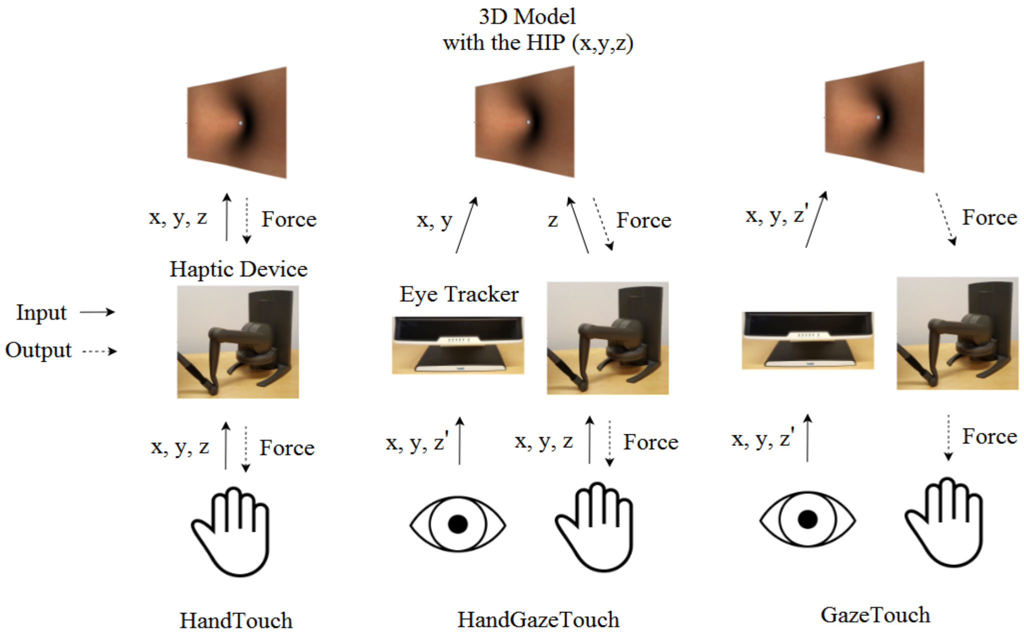


Figure 2: Concept map of the three interaction techniques. The hand movement coordinates respectively are x, y and z on the three axes. The eye gaze coordinates are x and y, and z' is the dwell time of the eye gaze.

TABLE 1.
SPECIFICATION OF INTERACTION TECHNIQUES.

Techniques	Interactive steps from the perspective of the users	Force-feedback device usage	Eye tracker usage	Input operating dimension
HandTouch	Uses the hand to reach and touch the target.	Yes	No	Hand: x, y, z Eyes: -
HandGazeTouch	Uses eye gaze to reach the target and the hand to touch.	Yes	Yes	Hand: z Eyes: x, y
GazeTouch	Uses eye gaze to reach and dwell time to trigger the touch.	Yes	Yes	Hand: - Eyes: x, y, z'

HYPOTHESES

3.1 Interaction Techniques

The prototype system we developed supports three different interaction techniques: *HandTouch* (H), *HandGazeTouch* (HG), and *GazeTouch* (G). (The abbreviations are used in some of the figures and tables.) Figure 2 shows the three interaction techniques. In *HandTouch*, the user controls and manipulates the mechanical arm of the force-feedback device along all three degrees of freedom (x, y, z). In *HandGazeTouch*, the gaze of the user controls the position of the interaction point along the x-y plane. The user controls the touch behavior by manipulating the mechanical arm of the device along the z-axis (see also Figure 1). In *GazeTouch*, the gaze controls the interaction point along the x-y plane and the duration of gaze fixation controls its movement along the z-axis. The details of the three interaction techniques are shown in Table 1.

3.2 Hypotheses

Based on our research questions, we focused on the analysis of the three interaction techniques from five aspects: efficiency, accuracy, naturalness, fatigue, and pleasantness. We formulated the following hypotheses:

H1: *HandGazeTouch* and *GazeTouch* will be faster than *HandTouch*.

- The eyes are faster at reaching the target object than the hand [2]. *GazeTouch* and *HandGazeTouch*, which use eye gaze, may hence be faster than *HandTouch*.

H2: *HandTouch* will be better than both *HandGazeTouch* and *GazeTouch* in the terms of the accuracy of kinesthetic perception.

- Since people are used to employing the hands to touch objects, haptic cues caused by eye gaze may be unfamiliar to users. Users may be worse at discriminating differences in softness using *HandGazeTouch* and *GazeTouch*.

H3: *HandGazeTouch* and *GazeTouch* will cause less tiredness of the hand than *HandTouch*. However, *HandTouch* will be considered more natural and pleasant than the alternatives.

- In both the gaze-based conditions, hand activities are replaced by eye gaze. This may lead to less fatigue of the hand for *HandGazeTouch* and *GazeTouch*. However, since the eyes are naturally used for perception and not for intentional control, *HandGazeTouch* and *GazeTouch* may be considered less natural, less pleasant, and more cognitively demanding by the users.

4 EXPERIMENT

4.1 Method

To test our hypothesis, we designed a controlled lab experiment that followed a within-subject design. Three interaction techniques were examined as the experimental conditions in a softness discrimination task. Softness discrimination was selected as the experimental test because softness is one of the important properties of physical objects and perceiving a change in softness could be an appropriate method to examine the effect of different interaction techniques in kinesthetic perception. In addition, common force-feedback devices such as Phantom only support single-point interaction, which is adequate for accurate softness discrimination [27]. Each experimental condition involved 36 repetitions of the softness discrimination task, and every task involved discriminating the softness between two onscreen square-shaped skin models presented as 2.5D models without considering the thickness. The participant had to touch the two skin models and identify the harder of the two, then communicate the answer using the appropriate arrow key on the keyboard.

Since we are interested in both interaction efficiency and kinesthetic perception, a Fitts's Law [37] type of experiment only measuring pointing efficiency was not suitable for this study. Accurate kinesthetic perception may require participants to touch the same object repeatedly or compare the two models multiple times. Another experimental setup was motivated by the Fitt's Law study setup. The difficulty of touch behavior in softness discrimination relies on two factors: difficulty in reaching the target and difficulty in perceiving the difference in softness. Therefore, we designed the tasks to include two difficulty levels for reach (difficult and easy) and two difficulty levels for perception (difficult and easy).

The difficulty for reach was manipulated by controlling two variables of the skin models: the size of the skin models and the distance between two models displaced on the screen. The two skin models had the same size set at three possible levels: 2.8 cm, 4.8 cm, or 6.8 cm. The distance between the two models had three possible values: 8.1 cm, 16.2 cm, or 24.3 cm (measured from center to center). Each condition involved four occurrences of every combination of the size and distance levels (3 x 3 x 4 = 36 tasks). We used the Index of Difficulty (ID) of Fitts's Law as the tool to categorize the nine unique combinations of size and distance into the two distinct levels of difficulty. The grouping threshold was set at 1.75, with Fitts's ID calculated by the medium size (i.e. 4.8 cm) of the skin models and the small distance (i.e. 8.1 cm) between the models. The combinations with a higher ID belonged to the difficult level for reach, and the combinations with a lower or equal ID belonged to the easy level for reach.

The difficulty for perception was implemented by controlling the softness difference between the two skin models. The softness of each skin model is controlled by manipulating the stiffness coefficient (k) of the Spring Model used to implement the skin models, which is a variable with a range (i.e. 0.065–0.145) used in the software. The difference in values (Δk) of the stiffness coefficient for both the onscreen skin models (k1, k2) was manipulated to result in six different levels with varying difficulties for identifying the softness difference between the two skin models. The lowest value of the difference level was identified such that it is barely perceivable and requires close inspection to identify the harder model (i.e. k1 = 0.09, k2 = 0.12: Δk = 0.03). The highest value of difference level was chosen such that the difference is easy to perceive (i.e. k1 = 0.065, k2 = 0.145: Δk = 0.08). The six difference levels (i.e. 0.03–0.08) occurred six times in each condition (6 x 6 = 36 tasks). The six levels of difference were then categorized into two levels of perception difficulty (i.e. difficult: 0.03–0.05 and easy: 0.06–0.08) for ease of analysis.

For each task, the difficulties for reach and perception were independent of each other and randomly chosen from the list of possible combinations.

In the gaze-based conditions, the position of the HIP was dependent on the point of the user's gaze. Using raw gaze points caused the HIP to be jittery. We hence used a simple recursion-based filter to smoothen the gaze point, before displaying it onscreen, as also used in a previous study [38].

$$y(i) = W * x(i) + (1 - W) * y(i - 1). \quad (1)$$

where $y(i)$ is the ith smoothened gaze position and $x(i)$ is the ith actual gaze position produced by the eye tracker. The percentage weight for the actual gaze position W was selected as 0.1 in the study.

In the *GazeTouch* condition, the position of the HIP along the x- and y-axes was controlled by the gaze position onscreen, and the fixation duration was translated to the z-axis movement of the HIP. The user had to fixate for 1 second to initiate touch behavior and another second to reach the maximum touch force. During this process, the output force was linearly increased with time and continuously transferred to the mechanical arm. After 2 seconds of dwelling, the output force saturated and further dwelling did not lead to any changes. In addition, a large gaze movement (i.e. >1.4 cm) can reset the touch operation in *GazeTouch*, allowing the user to touch the same place again or a different place.

For the *HandTouch* and *HandGazeTouch* conditions, the position of the HIP along the x-y plane was respectively controlled by the hand and eye gaze. The distance to initiate touch in the z-axis for both conditions was set at a maximum of 6 cm, which is variable (normally < 6 cm) depending on how and where the user initially held the mechanical arm.

The study used both objective and subjective measures to understand the strengths and weaknesses of each interaction technique. The objective measures were the task completion time and the number of errors made in softness discrimination. The subjective measures were captured using a custom questionnaire. Table 2 shows the four questions with a 7-point Likert scale based on the subjective assessment questions used in NASA-TLX [39]. In addition, we used a post-test questionnaire in which the participants ranked the three interaction techniques based on the tiredness of their eyes and hand. Furthermore, the participants selected the preferred technique(s) from the three interaction techniques.

TABLE 2.
STATEMENTS IN THE QUESTIONNAIRE.

No.	Description
Q1	This interaction technique is mentally difficult.
Q2	This interaction technique is physically difficult.
Q3	With this interaction technique, it is natural to touch.
Q4	This interaction technique is pleasant.

4.2 Apparatus and Environment

The experiment was conducted on a Dell T3600 Windows 7 desktop computer with an Intel E5-1600 processor, NVIDIA Quadro 4000 graphics, and 8GB 1600MHz of memory. The experiment environment is shown in Figure 3. We used a Phantom Desktop [6] as the force-feedback device and a Tobii T60 [12] as both the display and the eye tracker. The software development kits were the open-source H3DAPI [40] for haptics and Tobii SDK for eye tracking. We also utilized TraQuMe [41], a tool to measure gaze data quality. A keyboard was used to select and record the answer for each task and move to the next task, and headphones were utilized to block out noise.

4.3 Pilot Study

A pilot study was conducted at first, employing six participants (three female and three male) aged between 24 and 42 years (Mean (M) = 31.8, Standard Deviation (SD) = 6.46) who had experience of using eye trackers and/or force-feedback devices. Based on the pilot tests, we further calibrated the system.

- In all conditions, six softness difference levels for the skin model (Δk: 0.03–0.08) were chosen, such that all were perceivable by the participants considering the sensitivity of the force-feedback device and human kinesthetic perception.
- In the *HandTouch* and *HandGazeTouch* conditions, the movement of the mechanical arm was transferred to the movement of the onscreen HIP without any scaling, e.g. in the *HandTouch* condition, a 1 cm lateral movement of the mechanical arm resulted in a 1 cm lateral movement of the HIP onscreen.
- In the *HandGazeTouch* and *GazeTouch* conditions, the weight of the current gaze position in the recursion filter used to smoothen the gaze position was chosen to be 0.1 based on the good balance between jitteriness and the responsiveness of the gaze point.
- In the *GazeTouch* condition, the parameters for translating fixation duration to the z-axis movement of the HIP were selected. The values (1s initiation and 1s dwelling) were selected such that they overcome the *Midas Touch* problem [42] and at the same time do not require staring for too long for touching.

4.4 Procedure

The participants were first introduced to the study and the equipment used. The force-feedback device was placed in position based on the user's dominant hand. All participants signed an informed consent form and then filled in the background questionnaire. Each question in the questionnaire was explained to clarify its meaning, such as the difference between mental difficulty and physical difficulty, before proceeding to the experiment.

Figure 3: The experiment setup involved a Tobii T60 gaze tracker and a Phantom force-feedback device. The two skin models on the screen have the medium size of 4.8 cm (out of the three size levels) and the large distance of 24.3 cm between them (out of the three distance levels).

Before each of the gaze-based conditions, the eye tracker was calibrated using nine-point onscreen calibration. The quality of eye tracking was measured using a nine-point TraQuMe evaluation [41]. We defined an objective criterion for recalibration. If any of the nine points showed more than 2 cm eye tracker offset, the participants were asked to recalibrate. The 2 cm threshold was defined considering the smallest size of the skin model used (i.e. 2.8 cm). If any of the tracking offset values were still beyond this threshold after multiple recalibration, the test was discontinued, and the data were not included in the analysis.

The participants were asked to finish each experimental task as accurately and efficiently as they could. Participants pressed the appropriate arrow key to record their answer for each softness comparison task, after which the system presented the next discrimination task. We did not present the participants with the feedback regarding the accuracy of their discrimination during the experiment.

There were three experimental conditions with 36 tasks in each condition. Before each condition, the participants had up to five minutes to familiarize themselves with the operation of each interaction technique. The order of the experimental conditions was counter-balanced. In the experiment, no hand-rest or elbow-rest equipment was used, and the height of chair/table was adjusted for the participants to make them face the screen and hold the mechanical arm of the force-feedback device horizontally. In addition, the participants were asked to wear headphones to block out the noise generated by the force-feedback device, because the noise level may indicate the magnitude of force and, thus, the softness of the onscreen skin models.

TABLE 3.
TESTS OF WITHIN-SUBJECTS EFFECTS ON COMPLETION TIME.

Sources	df	F	Sig.
Condition	2, 46	10.3	<0.001
Reach	1, 23	31.3	<0.001
Perception	1, 23	42.2	<0.001
Condition & Reach	2, 46	13.9	<0.001
Condition & Perception	2, 46	4.60	0.016
Reach & Perception	1, 23	0.41	0.52
Condition, Reach, & Perception	2, 46	0.60	0.54

4.5 Participants

We recruited 24 participants from our university community (13 female and 11 male) aged between 19 and 42 years (M=26.5, SD=6.13). All participants had normal touch sensitivity. Seven participants had corrected vision and the remainder had normal vision. Only two participants used the left hand as their dominant hand; the remainder were right-dominant. Seven participants had used a similar eye tracker before (≤ two times), and one participant had used the force-feedback device before (one time). Since their experience was limited, we included their data in this study.

The data for one of participants had to be replaced because of issues in gaze tracking. The participant could not pass the gaze tracking accuracy check using the TraQuMe tool, and thus could not complete the test. Another participant was invited to replace the original participant.

The mean gaze tracking accuracy for the 24 participants was 0.56 degrees (SD=0.17 degrees), which translated to 0.62 cm in screen distance (SD=0.19 cm).

5 RESULTS

5.1 Test Completion Time

For each condition with 36 tasks, we calculated the mean value of the task completion time for the different levels of reach difficulty and perception difficulty. The Shapiro–Wilk Normality test was conducted first. Since the data were not normally distributed (p < .001), we analyzed the data using a 3x2x2 Aligned Rank Transform (ART) repeated measures non-parametric ANOVA [43], and the post-hoc analysis was done using the Wilcoxon Signed-Rank test with Holm-modified Bonferroni correction [44] to control for family-wise type-1 error. All the p-values presented are after Holm-modified Bonferroni correction.

Table 3 shows the overall p-values for the completion times based on the conditions (*HandTouch*, *HandGazeTouch*, and *GazeTouch*), difficulty of reach (difficult and easy), and difficulty of perception (difficult and easy) through the repeated measure.

The ART ANOVA test results showed statistically significant main effects for all three factors: condition, reach, and perception. In addition, the ART ANOVA showed a statistically significant interaction effect for condition and reach as well as condition and perception. Other effects were not statistically significant.

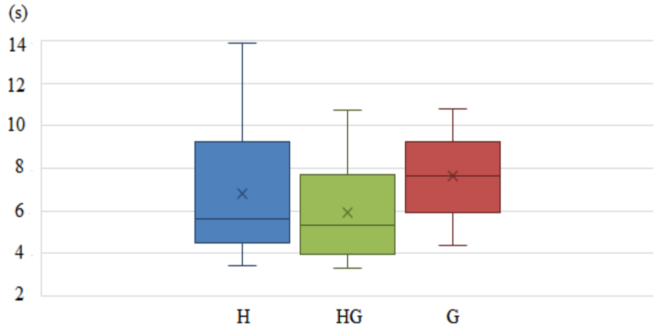


Figure 4: Completion time for the three interaction conditions. The line in the boxplot is the median value and the cross mark is the mean value (the following figures use the same marks).

Figure 4 shows the boxplot for overall completion time for the three interaction techniques. The mean value of task completion time, visualized as the y-axis in the boxplot, for *HandGazeTouch* (M = 5.92, SD = 2.21) was approximately 15% lower than for *HandTouch* (M = 6.82, SD = 2.92) and 29% lower than for *GazeTouch* (M = 7.61, SD = 1.86).

The Wilcoxon Signed-Rank test showed that *HandGazeTouch* was statistically significantly faster than both *HandTouch* (Z = -2.200, p = .048) and *GazeTouch* (Z = -3.743, p < .001). *HandTouch* and *GazeTouch* were not statistically significantly different from each other (Z = -1.914, p = .056), but the difference approached significance.

Unsurprisingly, reach and perception difficulty had a significant effect on the result. When the skin models were further apart or smaller in size, users took more time in movement (Z = -3.514, p < .001). Similarly, when the difference in softness between the two skin models was low, participants took more time in perception (Z = -4.200, p < .001). However, we are more interested in the interaction effect of these factors on the conditions to understand if the task completion times for the conditions were differentially affected by the perception and reach difficulty levels. Next, we analyze the interaction effects in detail.

Condition & Reach: Post-hoc analysis

Figure 5 demonstrates the completion time based on the condition and reach. Table 4 shows the mean completion time for the three interaction techniques based on the difficulty levels in the reach operations.

We analyzed the simple effect of reach difficulty for the three techniques using the Wilcoxon Signed-Rank test.

- *HandTouch* was significantly faster when the reach difficulty was low compared to when the reach difficulty was high (Z = -4.143, p < .001).
- The simple effect of reach difficulty was not statistically significant for the gaze-based conditions *HandGazeTouch* (Z = -1.057, p = .58) and *GazeTouch* (Z = - .029, p = .977).

Condition & perception: Post-hoc analysis

The completion time based on condition and perception is shown

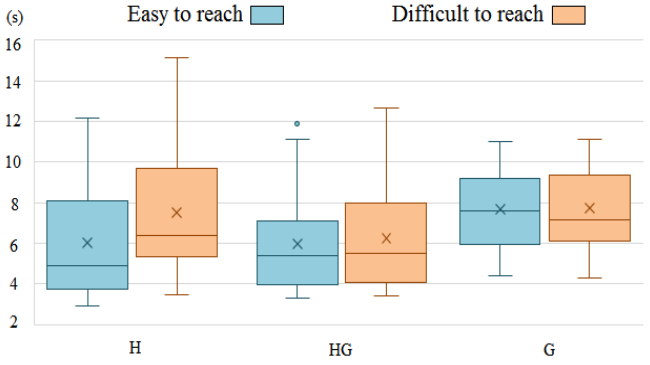


Figure 5: Task completion time of the three interaction techniques, based on reach difficulty levels.

TABLE 4.
MEAN TASK COMPLETION TIME USING THE THREE INTERACTION TECHNIQUES BASED ON REACH DIFFICULTY LEVELS.

Mean time (seconds)	Conditions		
	H	HG	G
Easy to reach	6.01 (SD=2.77)	5.95 (SD=2.45)	7.67 (SD=2.02)
Difficult to reach	7.51 (SD=3.16)	6.26 (SD=2.49)	7.72 (SD=2.01)

in Figure 6. Table 5 gives the mean completion time for each technique based on the perception difficulty levels.

We further analyzed the simple effect of perception difficulty levels for the three interaction techniques using the Wilcoxon Signed-Rank test.

- *HandTouch* was significantly faster when the perception difficulty was low than when the perception difficulty was high (Z = -2.371, p = .018): the mean value of task completion for *HandTouch* with a low perception difficulty was 6.63 seconds. This increased to 7.13 seconds, (a 7.5% increase) when the perception difficulty was high.

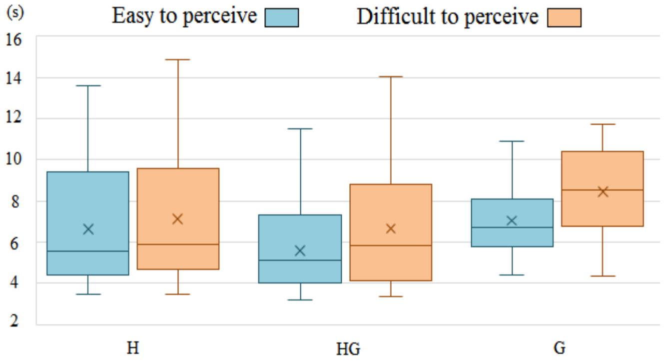


Figure 6: Task completion time of the three interaction techniques, based on perception difficulty levels.

TABLE 5.
MEAN COMPLETION TIME OF EACH TECHNIQUE BASED ON PERCEPTION DIFFICULTY LEVELS.

Mean time (seconds)	Conditions		
	H	HG	G
Easy to perceive	6.63 (SD=2.82)	5.58 (SD=2.20)	7.04 (SD=1.72)
Difficult to perceive	7.13 (SD=3.23)	6.65 (SD=2.85)	8.45 (SD=2.11)

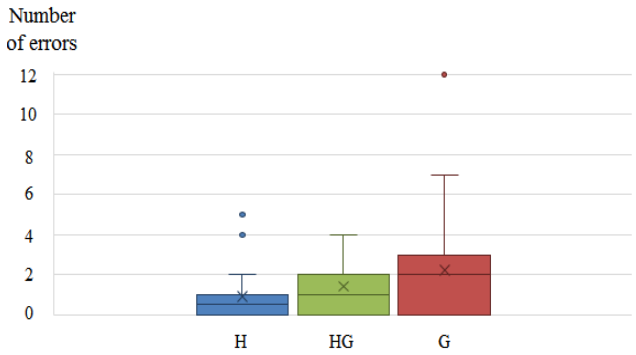


Figure 7: Error distribution of the three interaction techniques.

- The simple effect of perception difficulty was statistically significant for *HandGazeTouch* (Z = -3.600, p < .001): the mean value of the task completion time increased by 19%, from 5.58 seconds for low perception difficulty to 6.65 seconds for high perception difficulty.
- Similarly, the simple effect of perception difficulty was statistically significant for *GazeTouch* as well (Z = -3.943, p < .001): the mean value of the task completion time increased 20%, from 7.04 seconds for low perception difficulty to 8.45 seconds for high perception difficulty.

5.2 Error Analysis

Errors occurred when participants selected the wrong option after comparing the softness of the two skin models. Overall, there were a total of 109 errors, which is an average of 4.54 errors per participant, from a total of 108 tasks per participant (36 tasks per condition x 3 conditions). Of the total 109 errors, 85 errors occurred in tasks with the high perception difficulty. Since the number of errors were so few and most of them occurred in tasks with the high perception difficulty, we concentrated on the analysis of the three conditions with the perception difficulty.

The ART repeated measure 3x2 factorial ANOVA for error analysis showed a significant main effect of interaction techniques on error (F (2, 46) = 8.15, p = .001). The effect of perception difficulty was also statistically significant (F (1, 23) = 47.16, p < .001). There was no significant interaction effect between the perception difficulty and the conditions on the error rates (F (2, 46) = 2.26, p = .116).

Figure 7 provides the error distribution based on each technique. *HandTouch* had the least number of errors (median value below 1), followed by *HandGazeTouch* (median = 1), and *GazeTouch* (median = 2). The Wilcoxon Signed-Rank test showed that the difference between *HandTouch* and *GazeTouch* approached significance (Z = -2.251, p = .061). In addition, *HandGazeTouch* was not different from both *HandTouch* (Z = -1.452, p = .292) and *GazeTouch* (Z = -1.204, p = .229).

5.3 Results of Subjective Data

Subjective data from the questionnaire was evaluated to explore the results, including perceived mental and physical difficulties,

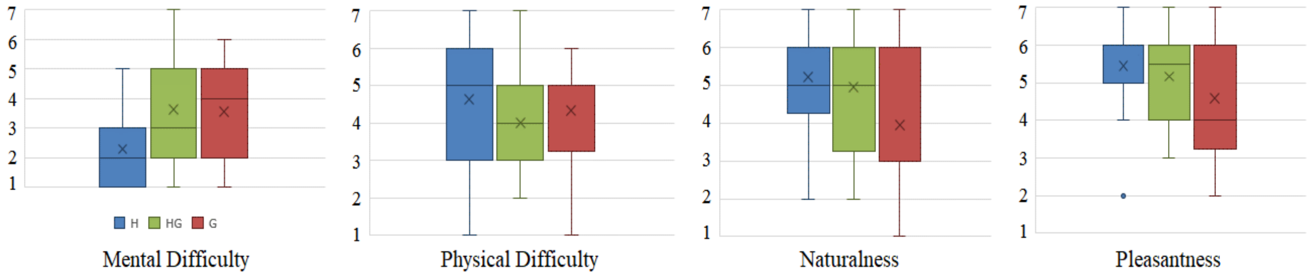


Figure 8: Subjective results of the study.

naturalness, and pleasantness. Figure 8 shows the subjective results of the questionnaire; the data was analyzed with the Wilcoxon Signed-Rank test.

Mental difficulty: *HandTouch* was better in terms of mental difficulty than *HandGazeTouch* (Z = -3.241, p = .003) and *GazeTouch* (Z = -3.002, p = .006). There was no difference between *HandGazeTouch* and *GazeTouch* (Z = -.049, p = .961).

Physical difficulty: There were no differences in terms of physical difficulty among three interaction techniques (*HandTouch-HandGazeTouch*: Z = -1.677, p = .188; *HandTouch-GazeTouch*: Z = -.579, p = .563; *HandGazeTouch-GazeTouch*: Z = -1.064, p = .574).

Naturalness: Both *HandTouch* (Z = -2.485, p = .026) and *HandGazeTouch* (Z = -2.912, p = .012) were considered more natural than *GazeTouch*. There was no difference between *HandTouch* and *HandGazeTouch* (Z = -1.010, p = .313).

Pleasantness: *HandTouch* was considered more pleasant than *GazeTouch* (Z = -2.535, p = .033). However, *HandGazeTouch* was not statistically significantly different from the others (*HandGazeTouch-HandTouch*: Z = -1.311, p = .190; *HandGazeTouch-GazeTouch*: Z = -1.849, p = .130).

5.4 Overall Evaluation of the Interaction Techniques

At the end of the study, all 24 participants ranked the techniques based on the tiredness of the hand and eyes while using the techniques. Overall, *GazeTouch* was ranked to be the least tiring for the hand (21 votes), followed by *HandGazeTouch* (19 votes). *HandTouch* was considered the most tiring for the hand (21 votes).

In terms of eye tiredness, *HandTouch* was voted the least tiring for the eyes (22 votes), followed by *HandGazeTouch* (19 votes). *GazeTouch* was considered the most tiring for the eyes (18 votes).

For overall preference, the participants could select more than one choice if they liked more than one condition equally. Overall, 15 of the 24 participants preferred *HandTouch*, and an almost equal number of participants (14) preferred *HandGazeTouch*. *GazeTouch* was the least preferred interaction technique; it was preferred by only 4 participants.

The users' preferences and the reasons behind them were revealed in the free-form comments provided by the participants:

P1: "I can sense the differences in softness better with it [*HandTouch*]. I make a circular motion on the tissue to better understand the softness." (User preferred *HandTouch.)*

P11: "[I prefer *GazeTouch*] because I did not have to make much physical effort. The *HandTouch* method made my hands ache. The *HandGazeTouch* method felt too complicated because there was too much to do at the same time."

P18: "*HandGazeTouch* was the fastest and most pleasant to use, with a little practice." (User preferred *HandGazeTouch*.)

P20: "[I prefer *HandTouch*] because it is the closest to real life when we touch real objects."

P24: "The *HandGazeTouch* method felt reasonably natural and really fascinating." (User preferred *HandGazeTouch.*)

6 DISCUSSION

This study investigated the use of eye gaze in kinesthetic interaction. It demonstrates that eye gaze is a feasible and beneficial input modality in this context. We will now discuss the findings of the study in relation to our initial research hypotheses and state of the art.

6.1 H1: *HandGazeTouch* and *GazeTouch* will be faster than *HandTouch*

Our study partly supports this hypothesis. With the multimodal input shown in the results, *HandGazeTouch* was significantly faster than both *HandTouch* and *GazeTouch*, which only involved a single input modality.

HandGazeTouch leverages the natural hand-eye coordination, and it uses eye gaze to replace the hand-reaching component in kinesthetic interaction. Similarly, the *GazeTouch* technique uses the gaze to replace both the reaching and touching components. However, this condition had the largest task completion time.

Our task required the user to touch two soft skin models placed apart horizontally. Often, when the difference in softness was small, participants had to touch each model multiple times to evaluate the difference. In *HandGazeTouch*, reaching the soft model was fast and intuitive, as the participants simply had to gaze at the target. The performance improvement in *HandGazeTouch* thus could be because of its improved efficiency in the reach operation. In *GazeTouch*, even though reaching the target was fast, the additional dwell time to cause haptic cues and overcome the *Midas Touch* problem likely slowed down the interaction.

The improved performance of *HandGazeTouch* in the reach operation is evident from our analysis of task completion times for reach difficulty. Our result (Figure 5) shows that the distance between the touch objects and their sizes influenced the task completion times in *HandTouch*. However, these variables did not have a noticeable effect in the gaze-based conditions

(*HandGazeTouch* and *GazeTouch*). Saccadic eye movements, which are responsible for bringing an object of interest to our foveal vision, typically last 30–120 milliseconds, and the effect of the distance of the object on the time it takes for our eyes to focus on it (though almost linearly related) is minimal [45]. This, however, was not the case for the *HandTouch* technique. When the target was further away, using the hand took significantly more time to reach the targets.

Our analysis of task completion times for different perception difficulties shows an interesting result. Overall, participants took more time to complete the task in all conditions when the task perception difficulty was high. This suggests that when it was difficult to perceive the difference in softness, participants using all three techniques had to touch the same tissue multiple times or repeatedly alternate between the two models to clearly gauge the difference in softness. However, the increase in task completion times was less for *HandTouch* (only 7.5%) and substantially higher for the gaze-based conditions (19% and 20% for *HandGazeTouch* and *GazeTouch* respectively). A potential explanation is that even though gaze-based techniques have the advantage in the reach operation due to the saccadic eye movements, participants using *HandGazeTouch* and *GazeTouch* had to repeat the touch activity more times to accurately estimate the difference in softness between the two models. Previous research in touch perception has argued that purposeful hand operations modulate haptic perception [46]. Our results extend this finding to computer-mediated kinesthetic interaction. In tasks that require estimating subtle differences in softness, using gaze as a mechanism to substitute for the reaching or touching operations may lead to an increased amount of time in identifying softness compared to using *HandTouch*.

6.2 H2: *HandTouch* will be better than *HandGazeTouch* and *GazeTouch* in terms of the accuracy of kinesthetic perception.

Our study partially supports this hypothesis. The results give a preliminary indication that *HandTouch* may be better than *GazeTouch* in the accuracy of kinesthetic perception, but *HandGazeTouch* was not statistically significantly different from *HandTouch*.

The experimental task focused on softness discrimination between two skin models. Our participants committed a different number of errors in judging the softness difference based on the three interaction techniques (shown in Figure 7). The difference in the number of errors is interesting because it indicates that the kinesthetic interaction techniques may mediate our kinesthetic perception. Fewer errors in softness discrimination indicates that the generated kinesthetic cues were easier to interpret by the somatosensory system. Similarly, a higher number of errors indicates that the kinesthetic cues were harder to interpret by the somatosensory system.

In *GazeTouch*, participants made noticeably more errors than in *HandTouch*. While the difference only approached statistical significance, we believe our results present preliminary evidence that active haptic interaction using hand motions may be more accurate than passively feeling the kinesthetic cues. Our results are consistent with our understanding of haptic perception in real-world interaction with physical objects. Lederman and Klatzky [47] have argued that to accurately perceive haptic cues provided by an object, both haptic cues generated by the object and the physical motion of the hand made to gain the haptic cues play a decisive role in our cognitive process. *GazeTouch*, which only uses gaze as the input without any hand motion, breaks the link between the human somatosensory system and hand movements, and it might result in lower accuracy in kinesthetic perception compared to *HandTouch*. This is also evident from the performances and comments of the participants. Most participants used specific strategies to sense softness when using *HandTouch* (e.g. moving the hand in a circular motion on the surface of the model or touching it multiple times). Such strategies were difficult to employ with the *GazeTouch* technique. Furthermore, the haptic cues generated from the models using a hand to touch are active haptic information (including both action and reaction forces), but the force feedback using gaze to touch is passive haptic information (including only reaction force). Lamotte [27] has demonstrated that softness discrimination is more accurate for active touch than passive touch.

HandGazeTouch, on the other hand, used gaze to reach but still required hand motion to touch, which could be considered a good compromise for both the fast and accurate exploration of haptic cues. It allowed the participants to easily touch the model multiple times and adjust the gaze position to, e.g. the left or right of the current contact point to simulate the sliding motion of the HIP, thus including both active and passive haptic information. For example, touching objects with a left-right or up-down movement was done by slightly moving the eye gaze, and thus the force feedback is passive; multiple touches on the same part of the model were done using hand movement, and thus the force feedback is active haptic information. This may explain why *HandGazeTouch* had a lower error rate than *GazeTouch* and had a higher error rate than *HandTouch*, but overall had a comparable performance to the other techniques in terms of the interaction accuracy of kinesthetic perception.

6.3 H3: *HandGazeTouch* and *GazeTouch* will cause less tiredness of the hand than *HandTouch*. However, *HandTouch* will be considered more natural and pleasant than the alternatives.

Our results partially support this hypothesis based on the subjective data and overall evaluation of the interaction techniques. Both gaze-based techniques lead to less hand fatigue. However, *HandGazeTouch* was still considered as natural and pleasant as *HandTouch*.

Based on the questionnaire data (Figure 8), there was no difference in overall physical difficulty among the three interaction techniques. However, gaze-based techniques caused less fatigue on the hand than the *HandTouch* technique, as is evident from the overall ranking of the conditions. Repetitive hand-based force-feedback interaction can lead to fatigue of the hand [11]. Our results suggest that using the gaze to replace the hand motion can largely reduce the subjective perception of hand fatigue during the kinesthetic interaction. On the contrary, the use of the eyes as an input modality in gaze-based techniques led to increased fatigue of the eyes. *HandGazeTouch* led to less eye fatigue than *GazeTouch*. This may be because the eyes were used only as a "pointer" in *HandGazeTouch*, while the eyes were used both as a pointer and a mechanism for activation to cause haptic cues in *GazeTouch*. *HandTouch*, obviously, caused the least eye fatigue.

The same reason may also explain the mental difficulty of *GazeTouch*. In *GazeTouch*, the dual use of eye gaze to point and touch might also have induced a greater cognitive load. In *HandGazeTouch*, users needed to combine two senses (eye gaze and hand motion) as the kinesthetic input, which might incur additional cognitive load.

Furthermore, for the naturalness and pleasantness of gaze-based techniques, *GazeTouch* replaced both the reaching and touching operation with eye movements, which is very different from the way we interact with physical objects in the real world. On the other hand, humans are used to looking at an object prior to hand motions in physical tasks [2]. *HandGazeTouch* utilized our natural hand-eye coordination, and it is thus closer to real-world interaction than *GazeTouch*. This may explain our results in terms of the method's naturalness and pleasantness.

6.4 Limitations and Future Research

Our study has several limitations.

- We used a research-quality eye tracker and TraQuMe to ensure high-quality gaze data. We recalibrated the eye tracker when the offset in tracking in any screen area was above the threshold of 2 cm. While most of our participants did not have any problems in tracking accuracy (mean gaze offset value 0.62 cm), we noticed that despite of our best efforts, some of our participants still faced difficulties in interaction using the gaze-based techniques due to the accuracy of tracking. Overall, we think our eye tracking quality was good and may not be representative of the expected quality that could be anticipated in an everyday gaze-tracking scenario outside the lab (e.g. cheaper tracking hardware, frequent movement of the user, fewer user calibrations). Reduced gaze tracking accuracy will introduce additional complexities in the use of both the gaze-based techniques. For example, it would make gaze-based pointing more difficult, especially when the objects are small. Further research is required to better understand how the accuracy of tracking influences the use of the two gaze-based interaction techniques.
- Our experimental task only focused on softness discrimination. Other object properties, such as textural properties, were not part of this study. The performance of the three interaction techniques may turn out to be different in tasks that require discriminating the roughness or smoothness of object surfaces [48]. We propose to examine this in future research.
- Most of our participants had no experience in using gaze-tracking and force-feedback devices to interact with virtual objects. Our study involved a short-term evaluation, and therefore it provides little insight into the long-term use of the three interaction techniques. It is likely that our results on the mental and physical difficulty of the gaze-based interaction techniques may be influenced by the participant selection. Previous work [49] on gaze-based human-computer interaction has shown that novice users suffer from eye fatigue associated with unnatural eye movements (such as staring). However, experienced users do not report any eye fatigue [50]. In addition, it is also likely that the effect of learning was differential for the three interaction techniques and related to the mental difficulty. A longitudinal study is needed to understand how users learn to use these techniques and how the user's opinion of the technique may change with extended use.
- The size of the display we used in the experiment was relatively small. We used a Tobii T60 as both the display and the eye tracker in the study. The width of the T60 screen is only 39.5 cm, which limited the maximum distance between the two skin models used in the experiment. A key question for future research is how the three interaction techniques will fare when the distance between the touch points are larger. Further work using large 2D displays or a VR head-mounted display is required to answer this question.
- The experiment involved predominantly young participants. Ruff and Parker [51] have shown that participants in different age groups have different performances in motor speed and hand-eye coordination. Older users are significantly slower than younger groups. Therefore, further research is required to understand the effect of participant selection in terms of the age-related aspects in our results.
- We studied two gaze-based interaction techniques that replaced different phases of the kinesthetic interaction with eye gaze input. This is by no means a complete exploration of the design space of combining gaze and hand-based input for kinesthetic interaction. For example, another way to augment kinesthetic interaction with gaze would be to use eye gaze for the large movement of the HIP, while allowing fine exploration using hand motions to provide better haptic cues from the object. Future studies should investigate other novel ways of augmenting kinesthetic interaction with gaze input.

We believe a key application area for gaze-augmented kinesthetic interaction would be in the VR environment. VR displays can provide an immersive and large 3D environment. Previous studies [16], [17], [18] have demonstrated the potential benefits of using gaze modality on target pointing and selection in such environments. For kinesthetic interaction, the 3D nature of the VR environment may introduce additional opportunities and challenges for gaze-based input, especially those associated with the depth of objects in space. In our study, the depth of the objects was fixed to one level. When there are multiple close objects of different depths, interacting with the objects becomes more complex, requiring the use of both conventional hand-based kinesthetic interaction and the gaze-augmented kinesthetic interaction that we presented in this paper. Previous research on 3D gaze estimation suggests that it is feasible to estimate the depth of visual focus based on the convergence of the individual's eyes [52]. Thus, gaze-augmented kinesthetic interaction for a 3D environment could potentially utilize the depth of focus as a method to control the position of the HIP along the z-axis. In *HandGazeTouch*, such an approach may enable a more robust and consistent interaction at different object depths. Future work could investigate this aspect.

7 CONCLUSION AND PRACTICAL IMPLICATIONS

This study explored the use of the eyes in kinesthetic interaction. We developed two kinesthetic interaction techniques based on the combination of gaze and hand-based input. Further, we conducted a comprehensive experimental study involving a softness discrimination task and analyzed the multi-faceted effect of the new interaction techniques on the efficiency of the interaction, the accuracy of kinesthetic perception, and the user's experience. Our results suggest that eye gaze as an input channel has both strengths and limitations in improving kinesthetic interaction. Below, we summarize our key findings and the practical implications of our results.

- Gaze-augmented kinesthetic interaction can help overcome two key limitations of conventional kinesthetic interaction: it reduces the fatigue of the hands and infinitely expands the workspace of the current haptic interface. Thus, gaze-augmented methods could be considered in kinesthetic interactions that involve a large interaction environment or involve sustained and repetitive actions.
- Utilizing gaze input as a mechanism to reach objects (*HandGazeTouch*) is better than using it for reaching and touching (*GazeTouch*) in gaze-augmented kinesthetic interaction. *HandGazeTouch* is not only faster than *GazeTouch*, but also more natural and noticeably more accurate in perceiving subtle differences in softness. Despite the limitations, *GazeTouch* is a feasible interaction technique, and may be suitable for specific disabled users.
- *HandGazeTouch* is comparable to conventional hand-only kinesthetic interaction (*HandTouch*) in terms of the accuracy, naturalness, and pleasantness of interaction. In addition, *HandGazeTouch* is more efficient than *HandTouch*, and it may thus be specifically suited to tasks that are time sensitive.
- *HandTouch* is a solid kinesthetic interaction technique that is considered natural, pleasant, and less cognitively demanding. *HandTouch* may be specifically suitable for interactions that are less frequent, non-repetitive, and involve a small haptic interaction space that is easy to navigate using hand motions.
- The suitability of the specific interaction techniques depends on the context of use. For example, for time-sensitive tasks, efficiency may be the key metric and, thus, *HandGazeTouch* may be the most suited of the three techniques. Similarly, for precision tasks, accuracy maybe more important, making *HandTouch* the best method in such contexts. On the other hand, for tasks that involve frequent and prolonged usage, the naturalness and pleasantness provided by both *HandTouch* and *HandGazeTouch* may be key. Our results suggest that it may be best to provide users with the flexibility to choose the interaction technique depending on the specific context of use.

REFERENCES

[1] M. Turk, "Multimodal interaction: A review," J. Pattern Recognition Letters, vol. 36, no. 1, pp.189-195. Jan. 2014. DOI: 10.1016/j.patrec.2013.07.003

[2] M. Land, N. Mennie and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," *J. Perception*, vol. 28, no. 11, pp.1311-1328, Feb. 1999, doi:10.1068/p2935.

[3] M.C. Bowman, R.S. Johannson and J.R. Flanagan, "Eye-hand coordination in a sequential target contact task," *J. Exp. Brain Res.*, vol. 195, issue. 2, pp.273-283, May 2009, doi: 10.1007/s00221-009-1781-x.

[4] M.F. Land and M. Hayhoe, "In what ways do eye movements contribute to everyday activities?" *J. Vision Research*, vol. 41, issues. 25-26, pp.3559-3565, Nov.–Dec. 2001, doi: 10.1016/S0042-6989(01)00102-X.

[5] J.D. Crawford, W.P. Medendorp and J.J. Marotta, "Spatial Transformations for Eye–Hand Coordination," *J. Neurophysiol*, vol. 92, issue. 1, pp. 10–19, Jul. 2004, doi:10.1152/jn.00117.2004.

[6] SensAble Technologies, (n.d.). (Transferred to 3D systems). (2018, Oct 15). Available: https://www.3dsystems.com/

[7] NOVINT, (n.d.). (2018, Oct 15). Available: http://www.novint.com/

[8] T.H. Massie and J.K. Salisbury, "The PHANToM Haptic Interface: A Device for Probing Virtual Objects," *Proc. Of the ASME Dynamic Systems and Control Division*, 1994, vol. 1, DSC-Vol. 55-1. Available: https://alliance.seas.upenn.edu/~medesign/wiki/uploads/Courses/Massie94-DSC-Phantom.pdf

[9] S. Zhai, C. Morimoto and S. Ihde, "Manual and gaze input cascaded (MAGIC) pointing," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '99)*, pp. 246-253, May 1999, doi:10.1145/302979.303053.

[10] R. Ott, M. Gutierrez, D. Thalmann and F. Vexo, "Improving user comfort in haptic virtual environments through gravity compensation," *First Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Envir. and Teleoperator Systems. World Haptics Conf. 2005*, pp.401-409, Mar. 2005, doi:10.1109/WHC.2005.78.

[11] A. Hamam and A.E. Saddik, "User force profile of repetitive haptic tasks inducing fatigue," *2015 Seventh Int. Workshop on Quality of Multimedia Experience (QoMEX)*, May 2015, doi:10.1109/QoMEX.2015.7148127.

[12] TOBII, (n.d). (2018, Oct, 15). Available: https://www.tobii.com/

[13] L.E. Sibert and R.J.K. Jacob, "Evaluation of eye gaze interaction," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '00),* pp.281-288, Apr. 2000, doi:10.1145/332040.332445.

[14] M. Kumar, A. Paepcke and T. Winograd, "EyePoint: practical pointing and selection using gaze and keyboard," *Conf. on Computer-Human Interaction (CHI) 2007*, pp.421-430, Apr/May 2007, doi:10.1145/1240624.1240692.

[15] P. Majaranta and KJ. Räihä, "Twenty years of eye typing: Systems and design issues," *Proc. of the 2002 Symp. on Eye Tracking Research & Applications (ETRA '02)*, pp.15-22, Mar. 2001, doi:10.1145/507072.507076.

[16] V. Tanriverdi and R.J.K. Jacob, "Interacting with eye movements in virtual environments," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '00)*, pp. 265-272, Apr. 2000, doi:10.1145/332040.332443.

[17] T. Piumsomboon, G. Lee, R.W. Lindeman and M. Billinghurst, "Exploring natural eye-gaze-based interaction for immersive virtual reality," *IEEE Symp. on 3D User Interfaces (3DUI)*, pp. 36-39, Mar. 2017, doi:10.1109/3DUI.2017.7893315.

[18] K. Pfeuffer, B. Mayer, D. Mardanbegi and H. Gellersen, "Gaze + pinch interaction in virtual reality," *Proc. of the 5th Symp. on Spatial User Interaction (SUI '17)*, pp. 99-108, Oct. 2017, doi:10.1145/3131277.3132180.

[19] K. Pfeuffer, J. Alexander, M.K. Chong and H. Gellersen, "Gaze-touch: Combining gaze with multi-touch for interaction on the same surface," *Proc. of the 27th Annual Symp. on User Interface Software and Tech. (UIST '14)*, pp. 509-518, Oct. 2014, doi:10.1145/2642918.2647397.

[20] K. Pfeuffer, J. Alexander, M.K. Chong and Y.X. Zhang, "Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze," *Proc. 28th Annual Symp. on User Interface Software and Tech. (UIST '15)*, pp. 373-383, Nov. 2015, doi:10.1145/2807442.2807460.

[21] S. Stellmach and R. Dachselt, "Look & Touch: Gaze-supported target acquisition," *Proc. of the SIGCHI conf. on Human Factors in*

Computing Systems (CHI '12), pp.2981-2990, May 2012, doi:10.1145/2207676.2208709.

[22] J. Kangas, D. Akkil, J. Rantala, P. Isokoski, P. Majaranta and R. Raisamo, "Gaze gestures and haptic feedback in mobile devices," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '14)*, pp.435-438, Apr/May 2014, doi:10.1145/2556288.2557040.

[23] D. Akkil, J. Kangas, J. Rantala, P. Isokoski, O. Spakov and R. Raisamo, "Glance Awareness and Gaze Interaction in Smartwatches," *Proc. of the 33rd Annual ACM Conf. Extended Abstracts on Human Factors in Comput. Systems (CHI EA '15),* pp.1271-1276, Apr. 2015, doi:10.1145/2702613.2732816.

[24] G.P. Mylonas, K.W. Kwok, A. Darzi and G.Z. Yang, "Gaze-contingent motor channelling and haptic constraints for minimally invasive robotic surgery," *Int. Conf. on Med. Image Comput. Comput. Assist. Interv. (MICCAI 2008)*, pp. 676-683, Sep. 2008, doi:10.1007/978-3-540-85990-1_81.

[25] D.R.C. James, F. Orihuela-Espina, D.R. Leff, G.P. Mylonas, K.W. Kwok, A.W. Darzi and G.Z. Yang, "Cognitive burden estimation for visuomotor learning with fNIRS," *Int. Conf. on Med. Image Comput. Comput. Assist. Interv. (MICCAI 2010)*, pp.319-326, Sep. 2010, doi:10.1007/978-3-642-15711-0_40.

[26] D.R.C. James, D.R. Leff, F. Orihuela-Espina, K.W. Kwok, G.P. Mylonas, T. Athanasiou, A.W. Darzi and G.Z. Yang, "Enhanced frontoparietal network architectures following "gaze-contingent" versus "free-hand" motor learning," *J. Neuroimage*, vol. 64, pp. 267-276. Jan. 2013, doi:10.1016/j.neuroimage.2012.08.056.

[27] R.H. LaMotte, "Softness discrimination with a tool," *J. Neurophysiol.*, vol. 83, no. 4, pp.1777-1786, Apr. 2000, doi: 10.1152/jn.2000.83.4.1777.

[28] T. Foulsham, "Eye movements and their functions in everyday tasks," *Cambridge Ophthalmological Symp. 2014*, Eye 29, pp.196-199, Nov. 2014, doi:10.1038/eye.2014.275.

[29] D.J. Liebling and S.T. Dumais, "Gaze and mouse coordination in everyday work," *Proc. of the 2014 ACM Int. Joint Conf. on Pervasive and Ubiquitous Comput.: Adjunct Publication (UbiComp '14 Adjunct)*, pp.1141-1150, Sep. 2014, doi:10.1145/2638728.2641692.

[30] F. Conti. And O. Khatib, "Spanning large workspaces using small haptic devices," *First Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Envir. and Teleoperator Systems. World Haptics Conf. 2005*, pp.183-188, Mar. 2005, doi:10.1109/WHC.2005.118.

[31] L. Dominjon, A. Lécuyer, J.M. Burkhardt, G. Andrade-Barroso, S. Richir, "The "bubble" technique: Interacting with large virtual environments using haptic devices with limited workspace," *First Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Envir. and Teleoperator Systems. World Haptics Conf. 2005*, pp.639–640, Mar. 2005, doi:10.1109/WHC.2005.126.

[32] T.J. Allen and U. Proske, "Effect of muscle fatigue on the sense of limb position and movement," *J. Exp Brain Res*, vol. 170, no. 1, pp.30-38, Mar. 2006, doi:10.1007/s00221-005-0174-z.

[33] N. Cortes, J. Onate and S. Morrison, "Differential effects of fatigue on movement variability," *J. Gait Posture*, vol. 39, no. 3, pp.888-893, Dec. 2013, doi:10.1016/j.gaitpost.2013.11.020.

[34] A.E. Saddik, M. Orozco, M. Eid, and J. Cha, "Haptics Technologies: Bringing Touch to Multimedia," *Springer Series on Touch and Haptic Systems*, M. Ferre, M.O. Ernst and A. Wing eds., Berlin: Springer-Verlag, 2011.

[35] J. Rantala, P. Majaranta, J. Kangas, P. Isokoski, D. Akkil, O. Špakov and R. Raisamo, "Gaze interaction with vibrotactile feedback: review and design guidelines," *J. HCI.*, Mar. 2017, doi:10.1080/07370024.2017.1306444.

[36] D.R. Leff, D.R. James, F. Orihuela-Espina, K.W. Kwok, L.W. Sun, G. Mylonas, T. Athanasiou, A.W. Darzi and G.Z. Yang, "The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills," *J. Front. Hum. Neurosci.*, vol. 9, issue. 526, Oct. 2015, doi:10.3389/fnhum.2015.00526.

[37] P.M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, pp.381–391, Jun. 1954, doi:10.1037/h0055392.

[38] D. Akkil, J.M. James, P. Isokoski and J. Kangas, "GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks," *Proc. of the 2016 CHI Conf. Extended Abstracts on Human Factors in Comput. Systems (CHI EA '16),* pp.1151-1158, May 2016, doi: 10.1145/2851581.2892459.

[39] S.G. Hart and L.E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *J. Advances in Psychology*, vol. 52, Human mental workload, pp.139-183, 1988, doi:10.1016/S0166-4115(08)62386-9.

[40] H3DAPI, (n.d.). (2018, Oct 15). Available: http://www.h3dapi.org/

[41] D. Akkil, P. Isokoski, J. Kangas, J. Rantala and R. Raisamo, "TraQuMe: A Tool for Measuring the Gaze Tracking Quality," *Proc. of the 2014 Symp. on Eye Tracking Research & Applications (ETRA '14)*, pp.327-330, Mar. 2014, doi:10.1145/2578153.2578192.

[42] R.J.K. Jacob, "The use of eye movements in human-computer interaction techniques: what you look at is what you get," *J. ACM Transactions on Information Systems (TOIS) - Special issue on CHI*, vol. 9, issue. 2, pp.152-169, Apr. 1991, doi:10.1145/123078.128728.

[43] J.O. Wobbrock, L. Findlater, D. Gergle and J.J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '11)*, pp.143-146, May 2011, doi:10.1145/1978942.1978963.

[44] S. Holm, "A Simple Sequentially Rejective Multiple Test Procedure," *Scand. J. Statist.*, vol. 6, no. 2, pp. 65–70, Jan. 1979, doi: 10.2307/4615733.

[45] R.J.K. Jacob, "Eye Tracking in Advanced Interface Design," *Virtual environments and advanced interface design*, W. Barfield and T. A. Furness, eds., New York: Oxford Univ. Press, pp.258-288, 1995.

[46] G. Juravle, F. McGlone and C. Spence, "Context-dependent changes in tactile perception during movement execution," *J. Front. Psychol.,* 4: 913, Dec. 2013, doi:10.3389/fpsyg.2013.00913.

[47] S.J. Lederman and R.L. Klatzky, "Hand Movements: A window into haptic object recognition," *J. Cognitive Psychol.*, vol. 19, issue. 3, pp.342-368, Jul. 1987, doi:10.1016/0010-0285(87)90008-9.

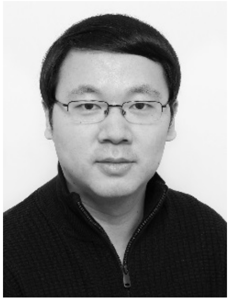
[48] W.M.B. Tiest, "Tactual perception of material properties", *J. Vision Research*, vol. 50, issue 24, pp.2775-2782, Dec. 2010, doi:10.1016/j.visres.2010.10.005.

[49] P. Majaranta, U.K. Ahola and O. Špakov, "Fast Gaze Typing with an Adjustable Dwell Time," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '09)*, pp.357-360, Apr. 2009, doi:10.1145/1518701.1518758.

[50] N. Chitty, "User Fatigue and Eye Controlled Technology," PhD dissertation, OCAD Univ. Toronto, Ontario, Canada, 2013.

[51] R.M. Ruff and S.B. Parker, "Gender- and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the finger tapping and grooved pegboard tests," *J. Percept. Mot. Skills*, 76 (3 Pt 2), pp.1219-1230. Jun. 1993, doi:10.2466/pms.1993.76.3c.1219.

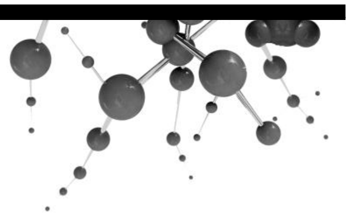
[52] A.T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, Springer International Publishing, pp.77-81, 2017.



Zhenxing Li received his MSc degree in Radio Frequency Electronics from Tampere Univerisy of Technology and MSc degree in User Interface Software Development from Univeristy of Tampere, Finland, in 2009 and 2013 respectively. He currently is a researcher in Tampere Unit for Computer-Human Interaction (TAUCHI) research center at Tampere University and working toward PhD degree in Interaction Technology. His research interests include haptic interaction and rendering, multimodal interaction, gaze tracking, mathematical modelling and algorithms.

Deepak Akkil received his MSc degree in Human-Technology Interaction from University of Tampere in 2014. He is currently pursuing PhD degree in Interaction Technology at Tampere Unit for Computer-Human Interaction (TAUCHI) research center in Tampere University. His research interests include gaze tracking, haptic interaction, computer-supported cooperative work (CSCW) and user experience design.

Roope Raisamo is a professor of computer science in Tampere University, Faculty of Information Technology and Communication Sciences. He received his PhD degree in computer science from the University of Tampere in 1999. He is the head of TAUCHI Research Center (Tampere Unit for Computer-Human Interaction), leading Multimodal Interaction Research Group. His research is focused on multimodal interaction, XR, haptics, gaze, gestures, interaction techniques, and software architectures.



Paper III

Zhenxing Li, Deepak Akkil and Roope Raisamo (2020). Gaze-Based Kinaesthetic Interaction for Virtual Reality. *Interacting with Computers*, Volume 32, Issue 1, Pages 17–32. https://doi.org/10.1093/iwcomp/iwaa002

Advance Access publication on 12 April 2020
doi: 10.1093/iwc/iwaa002

Gaze-based Kinaesthetic Interaction for Virtual Reality

ZHENXING LI*, DEEPAK AKKIL AND ROOPE RAISAMO

Faculty of Information Technology and Communication Sciences, Tampere University, Kalevantie 4, 33100 Tampere, Finland

**Corresponding author: Zhenxing.Li@tuni.fi*

Kinaesthetic interaction using force-feedback devices is promising in virtual reality. However, the devices are currently not suitable for interactions within large virtual spaces because of their limited workspace. We developed a novel gaze-based kinaesthetic interface that employs the user's gaze to relocate the device workspace. The workspace switches to a new location when the user pulls the mechanical arm of the device to its reset position and gazes at the new target. This design enables the robust relocating of device workspace, thus achieving an infinite interaction space, and simultaneously maintains a flexible hand-based kinaesthetic exploration. We compared the new interface with the scaling-based traditional interface in an experiment involving softness and smoothness discrimination. Our results showed that the gaze-based interface performs better than the traditional interface, in terms of efficiency and kinaesthetic perception. It improves the user experience for kinaesthetic interaction in virtual reality without increasing eye strain.

RESEARCH HIGHLIGHTS

- The study proposed a new gaze-enhanced kinaesthetic interface which employs the user' gaze as an input to relocate the workspace of the force-feedback device, for kinaesthetic interaction within a large virtual reality environment.
- The new interface addressed the limitations associated with using the force-feedback device and avoided the issues on usability and applicability existed in the previous gaze-based kinaesthetic interface.
- The study compared the proposed interface with the scaling-based traditional interface in an experiment involving softness and smoothness discrimination and demonstrated its better performances in terms of interaction efficiency, kinaesthetic perception accuracy and user experience.

Keywords: kinaesthetic interaction; gaze tracking; force-feedback device; virtual reality; workspace; fatigue

Handling Editor: Dr. Rod McCall

Received 14 June 2019; Revised 13 November 2019; Accepted 10 January 2020

1. INTRODUCTION

Virtual reality (VR) is becoming increasingly popular in applications such as entertainment (Bates, 1992), professional training (Aggarwal *et al.*, 2006), telepresence, product design, manufacturing (Mujber *et al.*, 2004) and e-commerce. The existing interactions in VR primarily rely on our visual and auditory senses. One of the fundamental ways in which we perceive our physical world is by touch. However, the ability to touch virtual objects, to perceive their geometry, texture and softness, is still difficult in the virtual world.

Kinaesthetic interaction as a form of human–computer interaction (HCI) enables realistic bidirectional touch behaviours. It enables natural hand motions as the kinaesthetic input and simultaneously allows users to perceive touch feeling (Saddik *et al.*, 2011). Multiple existing devices can be used for implementing kinaesthetic interaction in the virtual world, namely ungrounded kinaesthetic pens (Kamuro *et al.*, 2011), wearable haptic gloves (HaptX, 2019; CyberGlove, 2019) and grounded force-feedback devices such as Geomagic Touch (3D Systems, 2019) and Novint Falcon (NOVINT, 2019).

Using grounded force-feedback devices is a promising kinaesthetic interaction setup, because these devices can enable a realistic touch interaction without the need for complex body augmentation. Force-feedback devices (e.g. Geomagic Touch) typically have a mechanical arm with three or six degrees of freedom (Massie and Salisbury, 1994). By holding the arm, the user can use hand motions as the kinaesthetic input for haptic exploration. Further, the device transfers the generated force as the kinaesthetic output to the hand for simulating the feeling of touch.

Using grounded force-feedback devices in VR has two challenges. First, the length of the mechanical arm in these devices limits the interaction space, making it difficult to explore a large VR environment. The common method to avoid this challenge is to use scaling of movement, also called control-display gain (Argelaguet and Andújar, 2013), where a small movement of the mechanical arm results in a large movement of the touch point in the virtual space. This solution helps to explore a large virtual space at the cost of reduced user control (Conti and Khatib, 2005; Dominjon *et al.*, 2005). Second, prolonged interaction can lead to hand fatigue (Hamam and Saddik, 2015). Hand fatigue negatively affects user performance and experience for kinaesthetic interaction (Allen and Proske, 2006; Cortes *et al.*, 2013).

Eye tracking is an emerging hands-free input mechanism for HCI. Numerous studies have demonstrated the feasibility of gaze as an input in various HCI scenarios. For example, gaze can enable easy object pointing and selection (Zhai *et al.*, 1999; Pfeuffer *et al.*, 2017; Nukarinen *et al.*, 2018). Further, the gaze input can assist the interaction with a desktop computer (Kumar *et al.*, 2007), mobile devices (Rantala *et al.*, 2017) and wearables, such as smart glasses (Akkil *et al.*, 2016) and smartwatches (Akkil *et al.*, 2015). Currently, gaze tracking is gaining mainstream significance, and numerous augmented reality and VR devices, such as HTC Vive Pro Eye (HTC Vive, 2019), Microsoft HoloLens 2 (Microsoft, 2019) and Magic Leap (MagicLeap, 2019), include built-in sensors that can track the user's gaze.

Using the gaze as a complementary input mechanism has been proposed to improve the current kinaesthetic interfaces. Tracking the user's gaze has been shown to be a reliable way to determine the touch position in a virtual environment (Cheng *et al.*, 2017). Li *et al.* (2019) directly used the gaze point to determine the point of interaction, and the touch operation was performed by moving the mechanical arm of the force-feedback device along the z-axis. The proposed interface, *HandGazeTouch*, enables fast point-and-touch interactions on a 2D computer screen.

The approach of directly using the gaze point as the point of interaction has several limitations. First, many of our everyday kinaesthetic explorations involve various hand motions other than the simple point-and-touch interactions. For example, when we want to identify the texture and geometry of an object, we commonly slide our fingers back and forth on the surface or along the edges of the object. These behaviours are difficult to perform using the *HandGazeTouch*. Second, this interface requires a consistent use of gaze as the input during kinaesthetic exploration, thus limiting the user's freedom of allocation of visual attention when interacting with an object. Third, as this interaction technique requires an explicit intentional use of gaze for pointing, prolonged interactions can lead to eye strain (Li *et al.*, 2019).

In this study, we present *gaze-switching workspace* (GSW), a new kinaesthetic interface combing the eye gaze with hand motions as the input to address the challenges of using force-feedback devices for VR interactions. The GSW interface uses gaze as a complementary modality to switch the device workspace. When the user pulls the mechanical arm of the device to its reset position (i.e. no longer in contact with any virtual object) and gazes at a new target for a predefined period of time, the device workspace is relocated and locked to the target object. The user can then employ hand motions along the x-, y- and z-axes to control the haptic interaction point (HIP) within the workspace to explore the target.

This gaze-based design infinitely increases the interaction space using the force-feedback devices, potentially relieves hand fatigue, reduces the intentional use of eye gaze and enables natural hand motions for complex kinaesthetic tasks (e.g. for perceiving the shape, texture and material of the virtual object). Figure 1 shows the GSW interaction and an example application.

To evaluate the usability of the GSW interface, we conducted an experimental study comparing the new interface with the traditional interface that employs scaling of hand motions as the baseline. The focus of the study was to understand how the GSW interface compares against the traditional kinaesthetic interface in two different types of kinaesthetic tasks (softness and smoothness discrimination) that inherently involve distinct hand movement patterns. We evaluated the two interfaces based on interaction efficiency, accuracy of kinaesthetic perception and user experience.

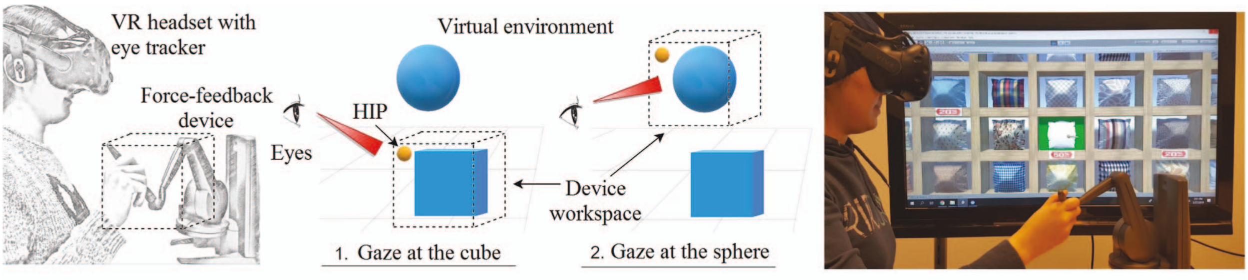
This study is novel in the following ways. We present a new gaze-enhanced kinaesthetic interface that addresses the problems of using the force-feedback device without the limitations existed in the previous work (Li *et al.*, 2019). To our knowledge, this study is the first on the gaze-based kinaesthetic interaction for VR. Moreover, the evaluation is based on an experiment involving both softness and smoothness perception. Softness and smoothness are two important material properties that highly affect our daily kinaesthetic tasks.

The rest of the paper is organized as follows. We first present the relevant previous works. We then describe the design of the gaze-based kinaesthetic interface and the experimental method, followed by the results and discussion.

2. BACKGROUND

2.1. Existing kinaesthetic interfaces for HCI

To provide kinaesthetic feedback to the user, different kinaesthetic hardware have been developed in the past few decades. The kinaesthetic pen (Kamuro *et al.*, 2011) is an ungrounded



(A) (B)

FIGURE 1. (**A**) Indicative diagram of the gaze-switching workspace (GSW) kinaesthetic interaction: in the physical environment, the user wears a VR headset with an integrated eye tracker and holds the mechanical arm of the force-feedback device. In the virtual environment, when the user looks at a virtual object, the workspace of the force-feedback device automatically switches to that object, enabling an efficient and effortless kinaesthetic interaction with the virtual object by moving the mechanical arm. (**B**) An online shopping application using the GSW interface: the online store has a variety of pillows on a large showcase. When the user looks at a specific pillow, the workspace of the force-feedback device locks to that pillow. The user can quickly reach for the pillow and freely feel and compare the texture and softness by moving the mechanical arm.

device that can provide sensations on the user's fingers for simulating the touch feeling with 3D models. However, the force feedback generated by the built-in springs and motors is limited. Haptic gloves use an exoskeleton to provide kinaesthetic feedback on the palmar surface of the hand based on, for example, hydraulic systems (Zubrycki and Granosik, 2017; HaptX, 2019) or electro-mechanical systems (Hinchet *et al.*, 2018; CyberGlove, 2019). They normally provide a limited force or require complex and often bulky hardware.

One of the most popular kinaesthetic interfaces is the grounded force-feedback devices. These devices enable natural hand motions as the kinaesthetic input and thus provide a flexible desktop-based kinaesthetic user interface for different application scenarios. More importantly, they can generate realistic forces with a high resolution (up to 1 kHz) (Massie and Salisbury, 1994). This study employed a force-feedback device as the kinaesthetic interface and augmented it with eye gaze to interact with virtual objects.

2.2. Limitations of grounded force-feedback devices

Kinaesthetically exploring a large virtual environment using the current force-feedback devices is challenging because of the small device workspace (Fischer and Vance, 2003; Conti and Khatib, 2005; Dominjon *et al.*, 2005). A manual clutching technique (Johnsen and Corliss, 1971) was first proposed to address this issue. The user employs the device button on the mechanical arm to declutch the HIP and manually moves the mechanical arm back to the centre position of the device workspace when the workspace limit has been reached. However, this manual method is not practical for kinaesthetic interaction within a large virtual space because of the multiple re-clutching process required to reach a distant target.

Currently, the common solution is to use scaling: to translate a small movement of the mechanical arm to cause a large movement of the HIP in the virtual space (Fischer and Vance, 2003; Argelaguet and Andújar, 2013).

For example, Conti and Khatib (2005) proposed the *Workspace Drift Controller*, which adopts a similar principle of clutching. The technique progressively centres the device workspace during the interaction and uses the scaling method in the movement across the large virtual space. The *Bubble*, proposed by Dominjon *et al.* (2005), utilizes a sphere around the HIP and adjusts the movement speed by scaling based on the relative positions of the HIP and the sphere.

Although scaling the hand motions is practical and easy to implement, it introduces new limitations. First, using the scaling method leads to the mismatch between hand motions and HIP movements, which may negatively affect the user control of the HIP and thus the kinaesthetic task performance.

Another limitation of the current kinaesthetic interfaces is hand fatigue. Previous studies have noted that using the current-generation force-feedback device is usually uncomfortable and could cause hand fatigue, because of the repeated hand movements involved (Ott *et al.*, 2005; Hamam and Saddik, 2015). Hand fatigue not only affects the user experience but also negatively affects user performance in spatial positioning and haptic manipulation (Allen and Proske, 2006; Cortes *et al.*, 2013). Therefore, reducing user fatigue is necessary in the design of kinaesthetic interfaces.

In this study, we present a novel approach to extend the interaction space by using eye gaze to select the location of the workspace instead of relying on scaling hand motions. The new interface has the potential to reduce the hand fatigue associated with kinaesthetic interaction.

2.3. Gaze interaction in HCI

Gaze modality has been used previously as an input channel in numerous HCI scenarios. For example, Zhai *et al.* (1999) combined gaze with the mouse input to improve the efficiency of pointing and to reduce physical effort and fatigue. Majaranta and Räihä (2002) proposed the use of eye gaze to type words on the screen instead of using a physical keyboard. Kumar *et al.* (2007) developed *EyePoint*, an interaction technique that uses the eyes to point and keypress to select for everyday computer usage.

The VR environment provides a large and even infinite interaction space, and gaze has been considered as an efficient input for selecting objects (Tanriverdi and Jacob, 2000; Pfeuffer *et al.*, 2017; Nukarinen *et al.*, 2018). Tanriverdi and Jacob (2000) examined the use of eye gaze to point at objects in a VR environment. Their results showed the strengths of gaze-based selection in terms of interaction efficiency, especially for distant objects. Pfeuffer *et al.* (2017) developed *Gaze + pinch*, an interaction method for manipulating objects in a VR environment, by using gaze to select the object and hand gestures to manipulate it. Nukarinen *et al.* (2018) investigated the combination of gaze and button press from a handheld controller to select objects in VR. The participants perceived the interaction technique as fast, natural and hands-free.

Previous studies have demonstrated the feasibility and value of eye gaze as an input mechanism in VR. In this study, we extend the use of gaze input to the context of VR kinaesthetic interaction.

2.4. Haptic interaction with gaze as the input modality

Haptic interaction in HCI can be categorized into tactile interaction and kinaesthetic interaction (Saddik *et al.*, 2011). Tactile interaction focuses on cutaneous sensation, whereas kinaesthetic interaction focuses on movement-based sensations originating from the muscles, tendons and joints.

For tactile interaction with tactile input devices, Stellmach and Dachselt (2012) proposed *Look and Touch*, which combines gaze with hand motions on the touchscreen as the input. Eye gaze was used for coarse pointing at the target, and fine selection was done by hand touching. The *Gaze-touch* developed by Pfeuffer *et al.* (2014) adopts gaze to control the movement of the cursor and combines it with hand motions on touchscreen for object manipulation. Their subsequent work, *Gaze-shifting* (Pfeuffer *et al.*, 2015), achieved a natural switching mechanism for the functions of gaze for *Gaze-touch*. In terms of tactile output, studies have noted that vibrotactile feedback is feasible and beneficial for gaze-based interaction (e.g. Rantala *et al.*, 2017).

In kinaesthetic interaction using the grounded force-feedback devices, previous studies have explored the use of gaze to augment kinaesthetic interaction, for example, to enhance safety in robotic surgery (Mylonas *et al.*, 2010) or foster a remote haptic collaboration (Leff *et al.*, 2015). Li *et al.* (2019) designed a gaze-based kinaesthetic interface to address the limited workspace. In their design, the HIP movement is controlled simultaneously by eye gaze (along the x- and y-axes) and hand movement (along the z-axis). The proposed interface supports fast point (using the gaze) and touch (using the hand-based input) interactions for softness detection. Other touch behaviours, such as feeling the texture and shape of an object, requires a complex HIP movement along the x- and y-axes and thus the accurate control of eye gaze. However, the human eye is primarily a perceptual organ. The fine control of eye movement is physiologically difficult and even impossible to perform.

The present work can be considered an extension of Li *et al.*'s (2019) study. We propose a new design for gaze-based kinaesthetic interaction that uses eye gaze for switching the device workspace and maintains the hand motions as the input for flexible kinaesthetic explorations.

3. METHOD

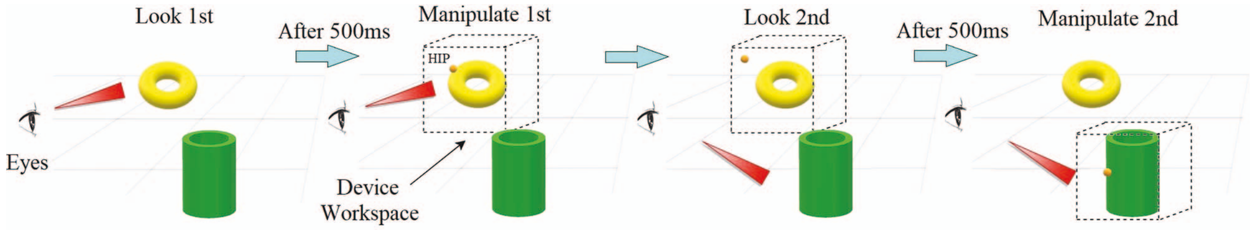
3.1. Design of the prototype system

In the development of the prototype system for the GSW kinaesthetic interface, an eye tracker is used to detect the user's current point of gaze. When the user needs to touch a distant object, he simply pulls the mechanical arm backward to its reset position and gaze at the target area for 500 ms (i.e. as the time interval between two continuous gaze points is fixed, the 500 ms dwell duration was implemented by judging whether there is enough number of gaze points continuously located on the model with an accumulated time of 500 ms). The centre of the workspace then switches to the centre of the target object and remains locked at the target until the user repeats this process. This two-step approach allows for the fast relocating of the device workspace and ensures that no accidental switches will occur during the kinaesthetic exploration due to wandering eyes. Thus, the user has visual freedom to gaze at other objects while interacting with an object.

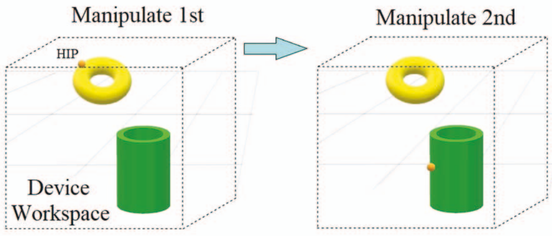
When the workspace switches to a new object, the position of the user's viewpoint and the size of the workspace, 16 cm × 12 cm × 12 cm (3D Systems, 2019), remain unchanged. The user notices the workspace switch by observing that the HIP has moved to the object. Then, the user can control the HIP along the x-, y- and z-axes by hand motions to explore the object.

The implementation of the traditional kinaesthetic interface adopts the scaling method to increase the size of the workspace. The selected scaling factor is 4, that is, a 1-cm physical movement of the mechanical arm results in a 4-cm movement of the HIP along the x-, y- and z-axes, thus increasing the size of the device workspace in the virtual environment by 4 times (i.e. 64 cm × 48 cm × 48 cm). The scaling function is available in the haptic plugin of Unity, the development software used in the study. The interaction processes using two kinaesthetic interfaces are shown in Fig. 2.

For both interfaces, the VR headset is used as the visual display, and the kinaesthetic feedback is transferred to the hand through the mechanical arm of the force-feedback device.



(A) GSW Interface



(B) Traditional Interface

FIGURE 2. (**A**) Interaction process of the gaze-based GSW interface that uses gaze to locate the workspace of the force-feedback device. (**B**) Interaction process of the traditional interface with an extended device workspace using the scaling method.

3.2. Design of the experiment

We designed a within-subject experiment in a controlled laboratory setting to compare the two kinaesthetic interfaces. The experimental task for the participant was to touch and compare two square models with different softness or smoothness characteristics. The widths of the models were the same (8 cm × 8 cm, without considering the thickness). After comparing the material properties, the participants were required to decide which one was harder or rougher and to communicate the answer by pressing the corresponding key on the keyboard (left model: A; right model: D). The participants had no visual feedback for the interaction (i.e. no visual deformation occurred when touching and the same material image was used for all models), because the different visual feedback could influence users' kinaesthetic perception (Samad *et al.*, 2019) and therefore their task performance.

To complete the tasks, the participants had to move the HIP from its resting position to reach the target and to alternately touch the two models to estimate the difference in the material properties. The difficulty to reach the objects was based on the spatial positions of the models. Targets that were far from each other (i.e. their positions along the x- and y-axes) and at a greater depth (i.e. their positions along the z-axis) were likely to be more difficult to reach than those that were close together and at a lower depth. Further, when the difference in material properties was not evident, the participants had to alternate between the targets multiple times to gauge which object was harder or rougher.

The difficulty in reaching the target and in perceiving the difference in the material properties could influence the task performance. Thus, we manipulated the difficulty of reach and the difficulty of perception as the experimental variables.

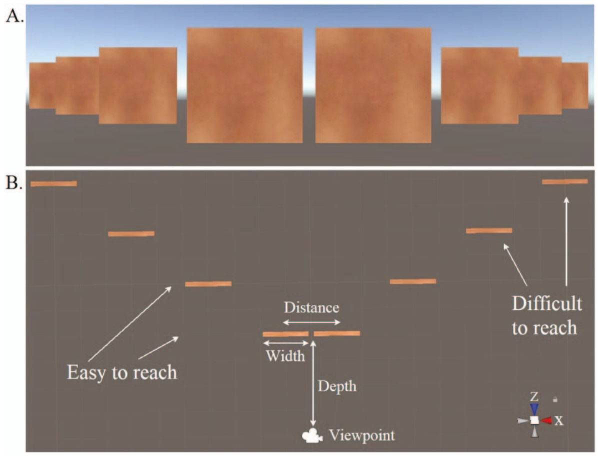


FIGURE 3. (**A**) All possible groups of models for the reaching difficulty from the viewpoint of the participant in the VR environment. (**B**) Top view of all possible groups based on the difficult and easy reaching difficulty levels. Each trial randomly selected one possible group.

The difficulty of reaching had two levels: easy and difficult (Fig. 3). The targets that were close to each other (distance: 9 and 24 cm distance measured from centre to centre) and close to the viewpoint (depth: 18 and 27 cm) were classified as easy to reach. The targets that were farther away from each other (distance: 39 and 54 cm) and farther away from the viewpoint

TABLE 1. Difficulty of perception for the softness and smoothness tasks.

	Softness task		Smoothness task	
Difficult to perceive	Δk: 0.01	k_1: 0.20	$\Delta\mu$: 0.02	μ_1: 0.135
		k_2: 0.21		μ_2: 0.155
	Δk: 0.02	k_1: 0.195	$\Delta\mu$: 0.03	μ_1: 0.13
		k_2: 0.215		μ_2: 0.16
	Δk: 0.03	k_1: 0.19	$\Delta\mu$: 0.04	μ_1: 0.125
		k_2: 0.22		μ_2: 0.165
Easy to perceive	Δk: 0.04	k_1: 0.185	$\Delta\mu$: 0.05	μ_1: 0.12
		k_2: 0.225		μ_2: 0.17
	Δk: 0.05	k_1: 0.18	$\Delta\mu$: 0.06	μ_1: 0.115
		k_2: 0.23		μ: 0.175
	Δk: 0.06	k_1: 0.175	$\Delta\mu$: 0.07	μ_1: 0.11
		k_2: 0.235		μ_2: 0.18

(depth: 36 and 45 cm) were classified as difficult to reach. Thus, based on the distance and depth values as well as the size of the models, the required size of the experimental space was at least 62 cm × 45 cm × 45 cm.

Similarly, the difficulty of perception had two levels: easy and difficult. For the easy level, the softness or smoothness differences between the two models were large and were thus relatively easy for the participants to identify. By contrast, for the difficult level, the differences were small, and identifying the harder or rougher model was more difficult.

The softness and smoothness of each model were implemented separately by the linear spring law ($F = kx$), where k is the stiffness coefficient and x is the penetration depth of the HIP, and by the kinetic friction ($F' = \mu F_n$), where μ is the friction coefficient and $F_n = mg$, the normal force value. Two haptic features were integrated in the haptic plugin of Unity.

We manipulated the softness and smoothness degrees by changing the stiffness coefficient (k) within the range of (0.175–0.235) and the friction coefficient (μ) within the range of (0.11–0.18). The difference in the stiffness coefficient (Δk) and the difference in the friction coefficient ($\Delta\mu$) between two models had six possible values (Δk: 0.01–0.06 in steps of 0.01; $\Delta\mu$: 0.02–0.07 in steps of 0.01). Table 1 presents the details of the difficulty levels of perception.

There were 24 trials for each task using each kinaesthetic interface. The reaching levels were repeated six times in the experiment (4 levels × 6 times = 24 trials), and the perception difficulty levels were repeated four times (6 levels × 4 times = 24 trials). For each trial, the reaching and perception levels were independent and randomly selected from a list of possible values (the selected values were removed from the lists after each selection). Each participant was asked to perform the softness and smoothness tasks using the two interfaces. Thus, each participant completed four rounds of the task (2 tasks × 2 interfaces = 4 task groups), resulting in a total of 96 trials in the experiment.

TABLE 2. Statements in the questionnaire.

Statements	Description
S1	This interaction technique is mentally easy to use.
S2	This interaction technique does not make my hand tired.
S3	This interaction technique does not make my eyes tired.
S4	This interaction technique is natural to use.
S5	This interaction technique is pleasant.

The prototype system recorded the task completion time and the users' answer for each trial as the objective data. Further, we used a seven-point Likert scale questionnaire to record the perceived mental effort, the hand and eye fatigue, and the naturalness and pleasantness of using the interfaces as the subjective data. The questionnaire was motivated by the NASA task load index questions (Hart and Staveland, 1988) to assess the users' subjective feelings. Table 2 presents the statements in the questionnaire.

3.3. Pilot study

Before the experiment, a pilot study was conducted with four participants who had previous experience in gaze or haptic interaction. The purpose of the pilot study was to identify the suitable parameters for the system.

- The size of the models (8 cm × 8 cm) was chosen, so that the users could smoothly perform the touch activities. The selected size was also large enough for the robust model selection by gaze, without the accuracy and precision of gaze tracking affecting the interaction, especially for the tasks in which the model was at a large depth.
- The distance and depth values for the two levels of the reaching difficulty were also chosen. As the spatial positions of objects could be various, we simply placed

FIGURE 4. Experimental environment. The additional 2D display shows one experimental trial with the group of models.

the two models at the same depth and manipulated their depth and the distance between them to design the variable of the reaching difficulty. Two levels (easy and difficult) could sufficiently represent the increasing difficulty in reaching distant objects.

- The values for the softness and smoothness differences between the models were selected based on human touch perception and the sensitivity of the force-feedback device. The high difficulty levels were selected to make them perceivable by the participants but with a more focused comparison.
- The scaling factor (4) for the traditional interface was chosen based on the original workspace size of the force-feedback device and the required virtual space size for the experimental tasks. Thus, the device workspace of the traditional interface (64 cm × 48 cm × 48 cm) could cover the experimental virtual space (62 cm × 45 cm × 45 cm), allowing the users to reach any spatial points required by the experimental tasks.
- We used the raw gaze data returned by the tracker without any additional filtering. Further, the dwell time for the workspace switching was selected as 500 ms. As the interaction involved two steps (i.e. moving the mechanical arm to its reset position and then dwelling), it was robust against unintentional activation. Thus, a short dwell time was selected to increase the feeling of responsiveness of the system and to avoid unnatural long staring while still enabling user control of the gaze-based activations.

4. EXPERIMENT

4.1. Apparatus

We used an MSI GS63VR 7RF Stealth Pro Windows 10 laptop with an Intel i7-7700HQ processor, GeForce GTX 1060 graphics card, and 16 GB of RAM as the host computer. A Geomagic Touch X (3D Systems, 2019) was utilized as the force-feedback device, and an HTC Vive VR headset (HTC Vive, 2019) with an integrated Tobii eye tracker (Tobii, 2019) was used as the display. The headphones on the VR headset were utilized to block out noise, and a keyboard was employed for the participants to record their answers. The test environment is illustrated in Fig. 4.

The experimental development software was Unity with SteamVR, Tobii SDK and haptic plugin for Geomagic OpenHaptics. The haptic plugin was used to connect the force-feedback device with Unity and to control the haptic properties of the virtual objects. The Tobii SDK was used for eye tracking.

4.2. Participants

We recruited 32 participants (21 women and 11 men) from the local university community aged between 19 and 40 years ($M = 24.6$, SD = 4.30). Thirteen participants had normal vision, and the others had corrected vision. Four participants reported they were left-handed, and the others were right-handed. Nine participants had used a similar VR headset one to two times, but none had previous eye-tracking experience. Two participants had used a similar force-feedback device on a 2D display once as a part of coursework. According to the self-reports, all participants had normal touch sensitivity.

4.3. Procedure

We first introduced the study and the equipment to the participants. All the participants signed an informed consent form and filled out the background questionnaire.

Each participant was assigned four task groups, and the order of the task groups was counterbalanced. For each task group, the participants were informed about the task type and introduced to the kinaesthetic interface. They had up to 5 min to familiarize themselves with use of the interface.

Before the experiment started, the eye tracker was calibrated using a five-point calibration system integrated with the Tobii SDK. For both interfaces, the participants used their dominant hand to hold the mechanical arm of the force-feedback device. The mechanical arm was positioned at its maximum distance from the device at the beginning. The participants sat comfortably on the chair and could rest their arm on the table while holding the force-feedback device.

The participants were instructed to complete each experimental trial as accurately and quickly as they could. We did not provide any feedback regarding the task performance. During the experiment, the participants could alternate between touching the two models as many times as they wanted, to identify the difference in softness or smoothness. After the comparison, they moved the mechanical arm back to the start position and used their non-dominant hand to press the corresponding key on the keyboard to record their answer. The prototype system then automatically moved to the next trial. After completing one task group with 24 trials, the participants had a 5-min break before the start of the next task group.

As the level of sound produced by the force-feedback device could indicate the amplitude of the force, constant white noise (fuzzy sound) was played to the participants through the headphones of the VR headset.

5. RESULTS

For simplicity, the results of the softness and smoothness discrimination tasks were analysed separately. We first conducted the Shapiro–Wilk normality test for the data. In each experimental task, the result showed that the task completion times and the task errors were not normally distributed (all $P < .001$). Therefore, we analysed the completion times and task errors separately using a 2 × 2 × 2 (interfaces × difficulty levels of reaching × difficulty levels of perception) aligned rank transform (ART) repeated-measures non-parametric analysis of variance (ANOVA) (Wobbrock *et al.*, 2011) for the softness and smoothness tasks, respectively. The post hoc analysis relied on the Wilcoxon signed-rank test. We used the Holm-modified Bonferroni correction (Holm, 1979) to control the family-wise type 1 error for interaction effects.

In the experiment, each participant was given 24 trials in each task group. We used the mean time to evaluate the interaction efficiency and the total number of errors of each task group to analyse the perception accuracy.

5.1. Softness discrimination task

Table 3 shows the results of the ART ANOVA on the task completion times and the errors for the softness task. As expected, the reaching difficulty and the perception difficulty statistically significantly influenced the task completion times and errors. In the following sections, we focus only on the main effect of the interfaces and the more interesting interaction effects.

5.1.1. *Interaction efficiency.*

For the main effect of interfaces, the post hoc analysis using the Wilcoxon signed-rank test showed that the task completion time when using the GSW interface ($M = 12.68$, SD = 5.41) was statistically significantly shorter than the time when using the traditional interface ($M = 14.5$, SD = 5.05; $Z = -2.749$, $P = .006$).

In addition, a statistically significant interaction effect was found between the interface and the reaching difficulty. Figure 5 shows the task completion time for the two interfaces based on the reaching difficulty.

For the traditional interface, the post hoc analysis showed that the participants spent less time completing the task when the reaching difficulty was low ($M = 13.64$, SD = 5.89) than when it was high ($M = 15.37$, SD = 5.02; $Z = -2.637$, $P = .016$). By contrast, for the GSW interface, the reach difficulty did not affect the task completion times ($M = 12.64$, SD = 5.83 for the low difficulty compared with $M = 12.93$, SD = 5.34 for the high difficulty; $Z = -1.272$; $P = .204$).

5.1.2. *Kinaesthetic perception accuracy.*

During the experiment, the participants committed different numbers of errors in judging the difference in softness while using the two kinaesthetic interfaces. Overall, 402 errors out of the 1536 trials (32 participants × 48 softness trials = 1536 trials), which yielded an overall error rate of 26.2%, were committed.

TABLE 3. Tests of within-subject effects for the softness task.

Sources	Time			Errors		
	df	F	Sig.	df	F	Sig.
Interfaces (I)	1, 31	10.595	**0.003**	1, 31	21.68	<**0.001**
Reaching (R)	1, 31	13.44	**0.001**	1, 31	13.841	**0.001**
Perception (P)	1, 31	8.235	**0.007**	1, 31	60.662	<**0.001**
I and R	1, 31	10.317	**0.003**	1, 31	12.339	**0.001**
I and P	1, 31	0.001	0.972	1, 31	0.818	0.373
R and P	1, 31	2.044	0.163	1, 31	0.067	0.797
I, R, and P	1, 31	1.530	0.225	1, 31	2.201	0.148

All significant P values are in bold

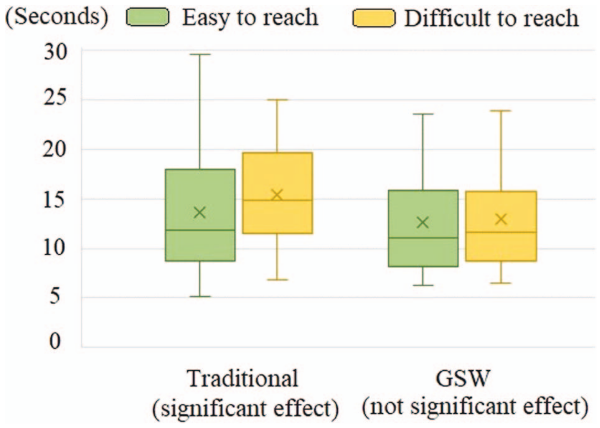


FIGURE 5. Task completion time based on the reaching difficulty for the softness task. (The cross mark is the mean value, and the line is the medium value.)

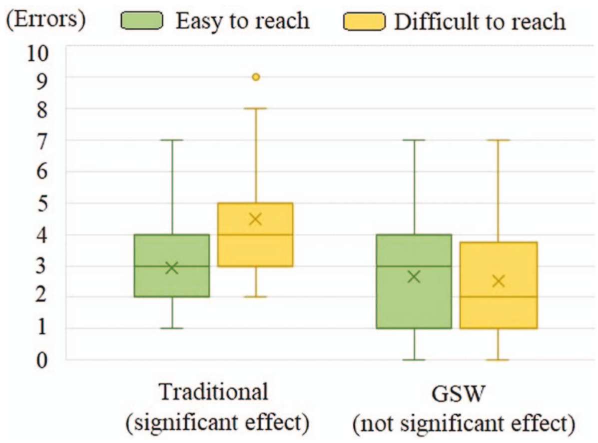


FIGURE 6. Task errors based on the reaching difficulty for the softness task.

For the main effect of the interfaces, the post hoc analysis demonstrated that the participants using the GSW interface ($M = 5.16$, $\mathrm{SD} = 3.02$) made statistically significantly fewer errors than those using the traditional interface ($M = 7.41$, $\mathrm{SD} = 2.33$) for the softness task ($Z = -3.52$, $P < .001$).

The interaction effect between the interfaces and the reaching difficulty was statistically significant. The boxplot is shown in Fig. 6. The post hoc analysis demonstrated that the participants using the traditional interface committed fewer errors when the reaching difficulty was low ($M = 2.94$, $\mathrm{SD} = 1.50$) than when it was high ($M = 4.47$, $\mathrm{SD} = 1.67$; $Z = -3.325$, $P = .002$). However, when using the GSW interface, the difference in the errors based on the reaching difficulty was not statistically significant ($M = 2.5$, $\mathrm{SD} = 1.8$ for the high reaching difficulty compared with $M = 2.66$, $\mathrm{SD} = 1.66$ for the low reaching difficulty; $Z = -0.65$, $P = .516$).

5.2. Smoothness discrimination task

Table 4 shows all the P values for the task completion times and the errors in the smoothness task based on the interfaces, the reaching and the perception difficulties. We focus on the main effect of the interfaces and the interaction effects in the following sections.

5.2.1. *Interaction efficiency.*

For the main effect of the interfaces, the post hoc analysis demonstrated that using the GSW interface ($M = 12.78$, $\mathrm{SD} = 5.28$) led to a statistically significantly shorter task completion time than using the traditional interface ($M = 14.98$, $\mathrm{SD} = 5.50$; $Z = -2.898$, $P = .004$).

Figure 7 shows the boxplot of the task completion times for the statistically significant interaction effect between the interfaces and the perception difficulty. For the traditional interface, no statistically significant difference was found in the task completion time when the smoothness discrimination was easy or difficult ($M = 14.73$, $\mathrm{SD} = 5.36$ compared with $M = 15.69$, $\mathrm{SD} = 6.43$; $Z = -1.589$, $P = .112$). By contrast, using the GSW interface led to a statistically significantly shorter task completion time when the perception difficulty was low ($M = 11.29$, $\mathrm{SD} = 4.08$) than when it was high ($M = 14.63$, $\mathrm{SD} = 6.07$; $Z = -4.394$, $P < .001$).

TABLE 4. Tests of within-subject effects for the smoothness task.

Sources	Time			Errors		
	df	*F*	Sig.	df	*F*	Sig.
Interfaces (C)	1, 31	7.359	**0.011**	1, 31	62.309	**<0.001**
Reaching (R)	1, 31	27.597	**<.001**	1, 31	34.586	**<0.001**
Perception (P)	1, 31	31.39	**<.001**	1, 31	86.856	**<0.001**
I and R	1, 31	0.3	0.588	1, 31	32.809	**<0.001**
I and P	1, 31	7.466	**0.01**	1, 31	4.645	**0.039**
R and P	1, 31	0.978	0.33	1, 31	3.052	0.091
I, R, and P	1, 31	0.521	0.476	1, 31	1.341	0.256

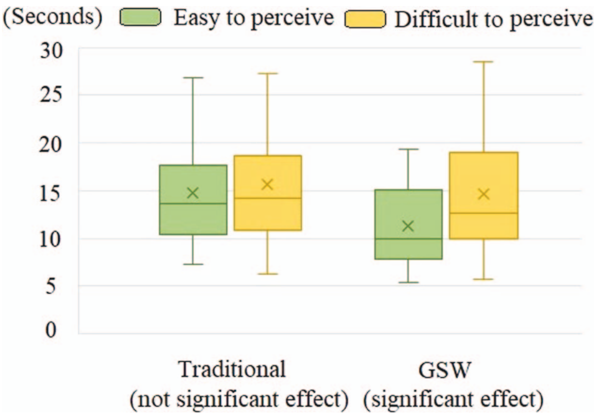


FIGURE 7. Task completion time based on the perception difficulty for the smoothness task.

5.2.2. Kinaesthetic perception accuracy.

In the smoothness task, 238 errors out of the 1536 trials, which yielded an overall error rate of 15.5%, were committed. For the main effect of the interfaces, the post hoc analysis demonstrated that the number of errors committed when the participants used the GSW interface (M = 1.5, SD = 1.5) was statistically significantly lower than the one committed when the participants used the traditional interface (M = 5.94, SD = 3.24; Z = −4.743, P < .001).

Figure 8 shows the boxplot of the task errors based on the interfaces and the reaching difficulty. Using the traditional interface caused fewer errors when the targets were easy to reach (M = 2.06, SD = 2.11) than when they were difficult to reach (M = 4.25, SD = 2.54; Z = −3.895, P < .001). However, when the participants used the GSW interface, no difference in the task errors was found based on the reaching difficulty (low difficulty: M = 0.63, SD = 1.01 compared with high difficulty: M = 0.88, SD = 1.04; Z = −0.989, P = .323).

Figure 9 illustrates the boxplot of the task errors based on the interfaces and the perception difficulty. For the traditional interface, the participants committed fewer errors when the perception difficulty was low (M = 2.5, SD = 2.13) than when it was high (M = 3.81, SD = 2.21; Z = −3.347, P = .002). The results for the GSW interface followed a similar pattern: the participants committed fewer errors when the perception difficulty was low (M = 0.31, SD = 0.54) than when it was high (M = 1.19, SD = 1.33; Z = −3.182, P = .001).

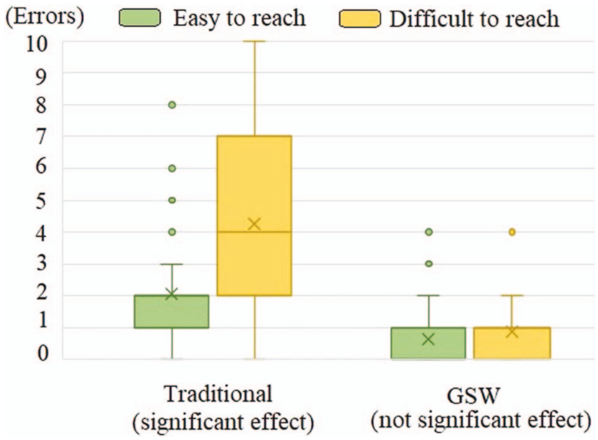


FIGURE 8. Task errors based on the reaching difficulty for the smoothness task.

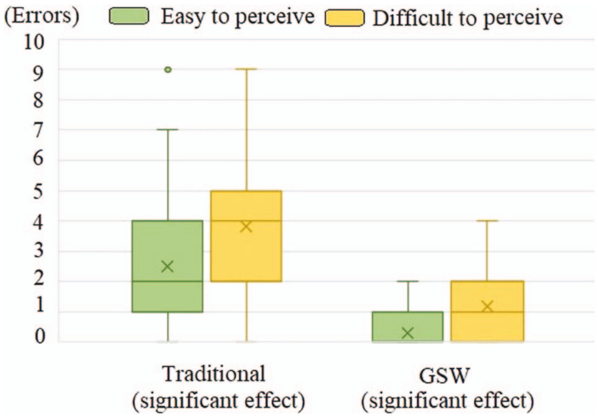


FIGURE 9. Task errors based on the perception difficulty for the smoothness task.

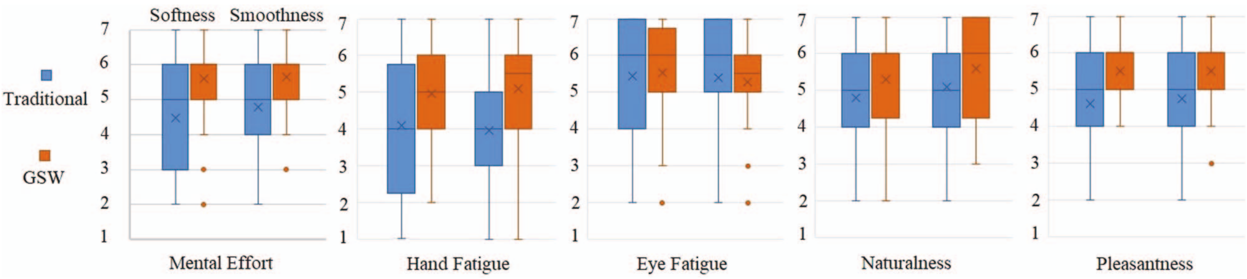


FIGURE 10. Subjective data of the study.

5.3. Subjective data

The subjective data collected from the questionnaire are shown in Fig. 10. We analysed the data using the Wilcoxon signed-rank test.

For the softness and smoothness tasks:

- Mental effort: the GSW interface was considered mentally easier to use than the traditional interface (softness: $Z = -3.224$, $P = .001$ and smoothness: $Z = -2.909$, $P = .004$).
- Hand fatigue: the GSW interface caused less hand fatigue than the traditional interface (softness: $Z = -3.137$, $P = .002$ and smoothness: $Z = -3.166$, $P = .002$).
- Eye fatigue: no difference was found between the two kinaesthetic interfaces (softness: $Z = -0.33$, $P = .741$ and smoothness: $Z = -0.502$, $P = .615$).
- Naturalness: the GSW interface was considered more natural than the traditional interface for the softness task (softness: $Z = -2.062$, $P = .039$), and no difference was found in terms of naturalness for the smoothness task, but the difference approached statistical significance ($Z = -1.952$, $P = .051$).
- Pleasantness: the GSW interface was considered more pleasant than the traditional interface (softness: $Z = -3.256$, $P = .001$ and smoothness: $Z = -3.095$, $P = .002$).

6. DISCUSSION

This study explored the use of eye gaze as an input modality to relocate the workspace of the force-feedback device for VR kinaesthetic interaction. This new design enabled a flexible and robust kinaesthetic exploration within a large virtual space. There were two types of kinaesthetic interaction tasks in the experiment: softness and smoothness discrimination. Softness and smoothness discrimination inherently involve different patterns and complexities of the HIP movement. The study results showed that the GSW interface has benefits for both kinaesthetic tasks. In this section, we discuss our results related to the state of the art.

6.1. Softness discrimination task

The softness discrimination task was characterized by multiple point-and-tap interactions with the objects, by alternating between them several times. Out results showed that the GSW interface generally enabled a shorter task completion time and a better kinaesthetic perception compared with the traditional interface.

6.1.1. *Interaction efficiency.*

The interaction efficiency of two kinaesthetic interfaces was highly influenced by the efficiency to reach the targets because of the type of kinaesthetic interaction the participants performed for softness discrimination. For the traditional interface, the participants took more time to complete the task when the targets were further apart. This was not the case while using the GSW interface.

Using the traditional interface, the participants had to use their hand along the x-, y- and z-axes to reach the targets which required more time when the distances were larger (Sallnäs and Zhai, 2003). Further, researchers have noted that the visual size of an object affects the motor speed in reaching the object (Berthier *et al*., 1996). In VR interactions, the visual size of the target object (the size along the x- and y-axes from the viewpoint of the user) is directly related to its position along the z-axis. For example, when the reaching difficulty was high, the targets were further away from the viewpoint of the participant and thus, appeared smaller. The small visual size of the object could have negatively affected the reaching speed by the hand motions. Our results support the previous results of Sallnäs and Zhai (2003) and Berthier *et al*. (1996) regarding the effects of the distances to the objects and their visual sizes on the reaching speed, which led to the lower efficiency while using the traditional interface.

Conversely, using the GSW interface, participants traversed the virtual space by employing saccadic eye movements that typically last 30–120 ms (Jacob, 1995). This approach eliminated the need for explicit hand motions to reach targets. Eye gaze has been found to be a fast input mechanism for reaching an object in different HCI contexts (Zhai *et al*., 1999; Majaranta and Räihä, 2002; Kumar *et al*., 2007), including kinaesthetic interaction (Li *et al.*, 2019) which uses the gaze point as the touch point.

In this study, we used the gaze to select the object for relocating the device workspace. Schuetz *et al.* (2019) noted that gaze as a selection method is less influenced by the distance between the targets and more influenced by the size of the targets. They suggested that a visual size of above 3 degrees for gaze-based selection could ensure a near-constant selection time, regardless of the distance between the targets. In our case, the smallest visual size of the target (real size: 8 cm) at the highest depth (45 cm) appeared approximately 10 degrees ($2 \times \arctan(8/(45 \times 2)) \approx 10$) in the visual angle. This likely explains why the reach difficulty had no influence on the task completion times and thus the increased interaction efficiency using the GSW interface.

6.1.2. *Kinaesthetic perception accuracy.*

In the experiment, the participants were asked to identify the harder model after comparing the two models. The difference in the number of errors committed by the participants indicates the difference in kinaesthetic perception by our somatosensory system enabled by the two interfaces. The results suggest that the kinaesthetic cues generated from the GSW interface are easier or clearer to interpret with the somatosensory system than those from the traditional interface.

Interestingly, the reaching difficulty differentially influenced the perception accuracy using two interfaces (Fig. 6). The reaching difficulty did not have a statistically significant effect on the number of errors for the GSW interface. However, for the traditional interface, the participants committed more errors when the objects were difficult to reach than when they were easy to reach.

This phenomenon while using the traditional interface has two potential explanations. First, when the objects were difficult to reach, participants had to perform a strenuous hand motion to control the HIP during the exploration (especially with the scaling on the hand motions). This strenuous hand motion could have interfered with hand ability to perceive subtle differences in the material properties, a phenomenon that is widely known as *tactile suppression* (Williams and Chapman, 2002; Chapman and Tremblay, 2015; Juravle *et al.*, 2017). *Tactile suppression* indicates that our somatosensory system naturally suppresses haptic perceptual sensitivity under motor commands (Chapman and Tremblay, 2015). It could explain the higher number of errors committed using the traditional interface when the targets were difficult to reach. A second plausible, although unlikely, explanation could be the effect of haptic memory. When the material properties of two objects are sequentially compared, an increased amount of time between the comparisons could lead to the fading of the haptic representation of the first stimulus in the memory. Previous studies showed that the memory of the haptic representation of an object is short-lived, up to 2 s (Shih *et al.*, 2009) or even 5 s (Metzger and Drewing, 2019). However, in the present case, the time to alternate between targets was much smaller (<1 s) and was unlikely to have had a significant effect.

In comparison, our results showed that the kinaesthetic perception accuracy was not affected by the reaching difficulty while using the GSW interface. Using the GSW interface enabled participants to quickly switch between the target objects without the need for strenuous hand motions, and to easily control the HIP for fine point-and-tap interactions. Similar to the *HandGazeTouch* (Li *et al.*, 2019), the GSW interface maintained a high accuracy in softness perception.

In sum, our results indicate that the GSW interface is a better interaction technique than the traditional interface for VR applications involving perceiving softness in terms of both efficiency and accuracy of perception.

6.2. Smoothness discrimination task

The smoothness discrimination tasks required the participants to switch between the target objects and were characterized by long kinaesthetic interactions with the objects using complex gestures (e.g. sliding back and forth or making a circular motion with the HIP). These touch behaviours are physiologically difficult to perform by eye gaze using the *HandGazeTouch* (Li *et al.*, 2019). However, our results showed that the GSW interface is suitable for the smoothness discrimination task. In the following, we discuss the results of the smoothness task using the GSW interface and the traditional interface.

6.2.1. *Interaction efficiency.*

For the smoothness discrimination task, the GSW interface was also faster than the traditional interface. However, this improved efficiency cannot be fully attributed to the faster reaching of the target enabled by gaze. In fact, the effect of the reaching difficulty did not differentially affect the task completion times for the two interfaces, shown in our results (i.e. no statistically significant interaction effect). Perceiving the smoothness of an object requires complex kinaesthetic interactions on the models, in contrast to the simple point-and-tap style interaction required to detect softness. Thus, the time required to reach the object is likely only a small part of the overall task completion time for the smoothness task.

Another possible reason for the improved overall efficiency of the GSW interface in the smoothness discrimination task could be the improved user control of the HIP. An important difference between the GSW interface and the traditional interface is the different scaling factor applied to the hand motions. As the GSW interface enables effortless relocation of the device workspace using the user's gaze, it overcomes the need for applying scaling on the hand motion to reach distant objects. The HIP movement matches with the hand movement, that is, a 1-cm movement of the mechanical arm moved the HIP by the same amount. By contrast, in the case of the traditional interface, scaling of hand motions must be applied (gain = 4) because of its inherently limited workspace. Consequently, a

TABLE 5. Differences between the HandGazeTouch and GSW.

Advantages	HandGazeTouch	GSW
Suitable for kinaesthetic tasks involving smoothness perception	No	Yes
Flexible for allocation of visual attention during exploration	No	Yes
Reduced intentional and explicit use of the eyes, thus lowering eye strain and mental effort	No	Yes
Simplified the mechanical design for force-feedback devices	Yes	No

1-cm movement of the mechanical arm resulted in four times the amount of movement of the HIP.

The greater control of the HIP using the GSW interface, compared with the traditional interface, could have enabled a more flexible and finer kinaesthetic exploration. With such a kinaesthetic exploration, a better haptic perception could be achieved (Lederman and Klatzky, 1987). Using the GSW interface, the participants could perform flexible hand motions to efficiently detect smoothness. The benefit in the interaction efficiency was more evident in tasks that were easy to perceive (Fig. 7).

6.2.2. *Kinaesthetic perception accuracy.*

Similar to the softness discrimination task, the reaching difficulty modulated the accuracy of the task performance for the traditional interface in the smoothness discrimination task (Fig. 8). Again, this could be attributed to the phenomenon of *tactile suppression* (Chapman and Tremblay, 2015). Conversely, the perception accuracy in the smoothness task using the GSW interface was unaffected by the difficulty of reaching the target.

Another difference between the GSW interface and the traditional interface was evident in the tasks in which the smoothness differences were difficult to perceive. Using the traditional interface, the range of the hand motion to detect the smoothness kinaesthetic cues (i.e. by sliding motions along the x- and y-axes) was limited by the size of the object and, more importantly, by the scaling applied. Thus, the traditional interface was practically not conducive for fine-level kinaesthetic explorations that were required for smoothness discrimination. This could have further led to more errors while using the traditional interface for the smoothness task with the high perception difficulty (Fig. 9). By contrast, the number of errors committed using the GSW interface was only marginally higher when the perception difficulty was high than when the perception difficulty was low.

6.3. User experience

This study compared two kinaesthetic interfaces in terms of a subset of user experience factors, such as perceived mental effort, hand and eye fatigue, and the naturalness and pleasantness of the interaction. Overall, the GSW interface was perceived to cause less mental effort and hand fatigue and considered to be more natural and pleasant to use (Fig. 10).

Multiple factors could have contributed to the improved rating of the GSW interface compared with the traditional interface. The GSW interface reduced the extent of the hand motion required to move between the target objects and thus led to less strain on the hands. Further, the GSW interface enabled the better control of the HIP and allowed a more realistic, natural and fine-level kinaesthetic exploration.

The participants did not report increased eye strain while using the GSW interface. Typically, performing unnatural eye movements, for example, long unnatural staring at objects (Majaranta *et al*., 2009) or unnatural gaze gestures (Chitty, 2013), can cause eye strain for novice users. The design of the GSW interface leverages the natural hand-eye coordination that exists in everyday physical tasks without the need for explicit unnatural eye movements.

At a general level, it is interesting to compare *HandGazeTouch* (Li *et al*., 2019) with this current work. Table 5 summarizes the differences between the two interfaces. Clearly, both the interaction techniques have their own advantages. *HandGazeTouch* can lead to a simplified mechanical design of the force-feedback device because the physical motion of the mechanical arm is required only along the z-axis. However, the GSW interface is better suited for flexible and user-friendly kinaesthetic explorations.

6.4. Limitations and future work

This study has several limitations.

First, the experimental scenario for examining the two kinaesthetic interfaces was simple. The models were large, flat in shape and unblocked in terms of visibility. The relatively large size enabled the easy selection with the eyes, and gaze data quality had little effect on our results. It is likely that when objects are small, placed very deeply or densely distributed, selecting them with the eyes would be highly susceptible to the gaze tracking accuracy and precision. Previous studies have started to address this issue. For example, Mardanbegi *et al.* (2019) used the depth estimation to resolve target ambiguity (e.g. due to partial occlusion) within a 3D environment and to make gaze selection more robust in complex VR scenarios. Future research is required to understand the use of the GSW interface in complex VR environments and study new methods to make gaze-based selection more robust in such environments.

Second, we examined the GSW interface with relatively small virtual objects compared with the size of the device workspace. For very large objects, the current implementation of the GSW interface may be hindered by the limits of the workspace because we always relocate the workspace to the centre of the object. Another possible approach for interacting with very large objects is to design multiple workspace relocation points on the same object based on eye gaze. We propose this idea for the future research.

Third, the study experiment focused on the comparison between the GSW interface and the traditional interface with the fundamental scaling technique that applies the scaling factor on the hand motions for reaching and manipulating objects. Future studies can compare the task performances using the GSW interface with those using the interfaces employing other techniques, such as *Workspace Drift Controller* (Conti and Khatib, 2005) and *Bubble* (Dominjon *et al.*, 2005), which applied the scaling factor on the hand motion only for reaching, within the large virtual environment provided by the VR headset.

Fourth, the participants were predominantly young students. The GSW interface is a novel interaction technique that requires a higher level of hand and eye coordination. Older participants are known to have less hand and eye coordination (Ruff and Parker, 1993) and may perform differently while using the GSW interface. Therefore, further research is required to understand the usability of the GSW interface for different participant groups.

Fifth, the force-feedback device we used supported only single-point interactions. For the softness task, using this device is adequate for an accurate softness discrimination (Lamotte, 2000). For the smoothness task, the participants easily completed the experimental task using this interaction method without any practical problems. However, we normally use multi-point touch (e.g. the palmar surfaces of the fingers) to detect the smoothness difference in the physical world. Conceptually, the GSW interface can be applied to multi-point touch interactions. Future research using an advanced force-feedback device is required to examine the effects of a multi-point interaction on the smoothness task while using the GSW interface.

Sixth, this study mainly focused on the virtual space in front of the user. Conceptually, the GSW interface can be applied to kinaesthetic interactions within the 360 degrees of the virtual space presented by the VR headset. Using the GSW interface, the device workspace is locked at the gazing object, and the direction of kinaesthetic input can be adjusted along with the user's view angle by orienting the workspace. This method makes the kinaesthetic interaction for the extended virtual space around the user possible. Future studies can investigate the performance of the GSW interface for this aspect.

Seventh, the study experiment involved only static objects. Typical VR scenes may have a multitude of moving stimuli (e.g. planets in motion). Interacting with moving objects using the traditional kinaesthetic interface may be particularly difficult. We believe the GSW interface to be well-suited for such interaction tasks because the workspace can remain locked to the object that has current visual attention. Previous research has shown that gaze is particularly suited for selecting moving targets (Shishkin *et al.*, 2018). However, future research is required to understand the cost and benefits of the GSW interface in such scenarios.

7. CONCLUSION

In this study, we presented the GSW interface, a gaze-enhanced kinaesthetic interaction technique for VR. The design of the GSW interface addressed the limited workspace and hand fatigue associated with using a force-feedback device. Eye gaze was used for locating the device workspace while still maintaining hand motions for various kinaesthetic interactions. This design avoided the overuse of the eye gaze and enlarged the applicability of the gaze-based kinaesthetic interface. The results of the experiment showed that the GSW interface is better than the traditional interface in terms of interaction efficiency, perception accuracy and user experience. It could be a compelling choice of interaction technique for future VR kinaesthetic interactions.

REFERENCES

3D Systems. https://www.3dsystems.com/ (retrieved February 20, 2019)

Aggarwal, R., Black, S.A., Hance, J.R., Darzi, A., and Cheshire, N.J.W. (2006) Virtual reality simulation training can improve inexperienced surgeons' endovascular skills. *Eur. J. Vasc. Endovasc Surg.*, 31:6, 588–593. https://doi.org/10.1016/j.ejvs.2005.11.009

Akkil, D., Kangas, J., Rantala, J., Isokoski, P., Špakov, O. and Raisamo, R. (2015) Glance awareness and gaze interaction in smartwatches. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computer Systems (CHI EA '15)*, pp. 1271–1276, Seoul, Republic of Korea. http://doi.org/10.1145/2702613.2732816.

Akkil, D., Lucero, A., Kangas, J., Jokela, T., Salmimaa, M., and Raisamo, R. (2016) User Expectations of Everyday Gaze Interaction on Smartglasses. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*, Gothenburg, Sweden, Article No. 24, pp. 1–10. http://doi.org/10.1145/2971485.2971496

Allen, T. J. and Proske, U. (2006) Effect of muscle fatigue on the sense of limb position and movement. *Exp. Brain Res.*, 170:1, 30–38. https://doi.org/10.1007/s00221-005-0174-z

Argelaguet, F. and Andújar, C. (2013) A survey of 3D object selection techniques for virtual environments. *Comput. Graph.*, 37:3, 121–136. https://doi.org/10.1016/j.cag.2012.12.003

Bates, J. (1992) Virtual reality, art, and entertainment. *Presence: Teleoperators Virt. Environ.*, 1:1, 133–138. https://doi.org/10.1162/pres.1992.1.1.133

Berthier, N.E., Clifton, R.K., Gullapalli, V., McCall, D.D., and Robin, D.J. (1996) Visual information and object size in the control of reaching. *J. Mot. Behav.*, 28: 3, 187–197. https://doi.org/10.1080/00222895.1996.9941744

Chapman, E. and Tremblay, F. (2015) Tactile suppression. *Scholarpedia*, 10:3, 7953. https://doi.org/10.4249/scholarpedia.7953

Cheng, L.-P., Ofek, E., Holz, C., Benko, H., and Wilson, A.D. (2017) Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*, Denver, Colorado, USA, pp. 3718–3728. https://doi.org/10.1145/3025453.3025753

Chitty, N. (2013) *User Fatigue and Eye Controlled Technology*. OCAD University, Toronto, Ontario, Canada.

Conti, F. and Khatib, O. (2005) Spanning large workspaces using small haptic devices. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 2005 IEEE World Haptics Conference*, Pisa, Italy, pp. 183–188. https://doi.org/10.1109/WHC.2005.118

Cortes, N.E., Oñate, J.A., and Morrison, S. (2013) Differential effects of fatigue on movement variability. *Gait & Posture*, 39: 3, 888–893. https://doi.org/10.1016/j.gaitpost.2013.11.020

CyberGlove. http://www.cyberglovesystems.com/ (retrieved February 20, 2019)

Dominjon, L., Lécuyer, A., Burkhardt, J.-M., Andrade-Barroso, G., and Richir, S. (2005) The "bubble" technique: interacting with large virtual environments using haptic devices with limited workspace. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 2005 IEEE World Haptics Conference*, Pisa, Italy, pp. 639–640. https://doi.org/10.1109/WHC.2005.126

Fischer, A. G. and Vance, J. M. (2003) PHANToM haptic device implemented in a projection screen virtual environment. In *Proceedings of the Workshop on Virtual Environments (EGVE '03)*, pp. 225–229. Zurich, Switzerland. https://doi.org/10.1145/769953.769979.

Hamam, A. and Saddik, A. E. (2015) User force profile of repetitive haptic tasks inducing fatigue. In *In the Seventh International Workshop on Quality of Multimedia Experience (QoMEX '15)*, Pylos-Nestoras, Greece. https://doi.org/10.1109/QoMEX.2015.7148127.

HaptX. https://www.haptx.com/ (retrieved February 20, 2019)

Hart, S. G. and Staveland, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances Psychol.*, 52, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hinchet, R., Vechev, V., Shea, H. and Hilliges, O. (2018) DextrES: wearable haptic feedback for grasping in VR via a thin form-factor electrostatic brake. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*, pp. 901–912, Berlin, Germany. https://doi.org/10.1145/3242587.3242657.

Holm, S. (1979) A simple sequentially rejective multiple test procedure. *Scand. J. Statist.*, 6: 2, 65–70. HTC Vive. https://www.vive.com/ (retrieved February 20, 2019).

Jacob, R. J. K. (1995) Eye tracking in advanced interface design. In Barfield, W., Furness, T. A. (eds), *Virtual Environments and Advanced Interface Design*, pp. 258–288. Oxford University Press, New York.

Johnsen, E. G. and Corliss, W. R. (1971) Human factors applications in teleoperator design and operation. In Meister, D. (ed), *Wiley series in Human Factors*. Wiley-Interscience, New York.

Juravle, G., Binsted, G., and Spence, C. (2017) Tactile suppression in goal-directed movement. *Psychon. Bull. Rev.*, 24: 4, 1060–1076. https://doi.org/10.3758/s13423-016-1203-6

Kamuro, S., Minamizawa, K., and Tachi, S. (2011) An Ungrounded Pen-Shaped Kinesthetic Display: Device Construction and Applications. In *2011 IEEE World Haptics Conference*, Istanbul, Turkey. https://doi.org/10.1109/whc.2011.5945546

Kumar, M., Paepcke, A., and Winograd, T. (2007) EyePoint: practical pointing and selection using gaze and keyboard. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, San Jose, California, USA, pp. 421–430. http://doi.org/10.1145/1240624.1240692

Lamotte, R.H. (2000) Softness discrimination with a tool. *J. Neurophysiol.*, 83:4, 1777–1786. https://doi.org/10.1152/jn.2000.83.4.1777

Lederman, S.J. and Klatzky, R.L. (1987) Hand movements: A window into haptic object recognition. *Cognitive Psychol.*, 19:3, 342–368. https://doi.org/10.1016/0010-0285(87)90008-9

Leff, D. R., James, D. R. C., Orihuela-Espina, F., Kwok, K. W., Sun, L. W., Mylonas, G., Athanasiou, T., Darzi, A. W. and Yang, G. Z. (2015) The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills. *Front. Hum. Neurosci.*, 9, 526. https://doi.org/10.3389/fnhum.2015.00526.

Li, Z., Akkil, D., and Raisamo, R. (2019) Gaze augmented hand-based kinesthetic interaction: What you see is what you feel. *IEEE Trans. Haptics*, 12: 2, 114–127. https://doi.org/10.1109/TOH.2019.2896027

MagicLeap. https://www.magicleap.com/ (retrieved February 20, 2019).

Majaranta, P., Ahola, U.-K., and Špakov, O. (2009) Fast gaze typing with an adjustable dwell time. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, Boston, MA, USA, pp. 357–360. http://doi.org/10.1145/1518701.1518758

Majaranta, P. and Räihä, K.-J. (2002) Twenty years of eye typing: systems and design issues. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications (ETRA '02)*, pp. 15–22, New Orleans, Louisiana. http://doi.org/10.1145/507072.507076.

Mardanbegi, D, Langlotz, T., and Gellersen, H. (2019) Resolving Target Ambiguity in 3D Gaze Interaction through VOR Depth Estimation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*, Glasgow, Scotland UK, Paper No. 612. https://doi.org/10.1145/3290605.3300842

Massie, T.H. and Salisbury, J.K. (1994) The PHANToM haptic interface: A device for probing virtual objects. *In Proceedings of the ASME Dynamic Systems and Control Division*. 55:1, 295–301.

Metzger, A. and Drewing, K. (2019) Effects of stimulus exploration length and time on the integration of information in haptic softness discrimination. *IEEE Trans. Haptics*, 12:4, 451–460. https://doi.org/10.1109/TOH.2019.2899298

Microsoft. https://www.microsoft.com/en-us/hololens/ (retrieved February 20, 2019).

Mujber, T. S., Szécsi, T. and Hashmi, M. S. J. (2004) Virtual reality applications in manufacturing process simulation. *J. Materials Process. Tech.*, 155-156, 1834–1838. https://doi.org/10.1016/j.jmatprotec.2004.04.401.

Mylonas, G.P., Kwok, K.W, James, D.R.C., Leff, D.R., Orihuela-Espina, F., Darzi, A.W., and Yang, G.Z. (2010) Gaze- contingent motor channeling, haptic constraints and associated cognitive demand for robotic MIS. *Med. Image. Anal.*, 16:3, 612–631. https://doi.org/10.1016/j.media.2010.07.007

NOVINT. http://www.hapticshouse.com/ (retrieved February 20, 2019)

Nukarinen, T., Kangas, J., Rantala, J., Koskinen, O., and Raisamo, R. (2018) Evaluating Ray Casting and Two Gaze-Based Pointing Techniques for Object Selection in Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18)*, Tokyo, Japan, Article No. 86. https://doi.org/10.1145/3281505.3283382

Ott, R., Gutierrez, M., Thalmann, D., and Vexo, F. (2005) Improving user comfort in haptic virtual environments through gravity compensation. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 2005 IEEE World Haptics Conference*, Pisa, Italy, pp. 401–409. https://doi.org/10.1109/WHC.2005.78

Pfeuffer, K., Alexander, J., Chong, M. K. and Gellersen, H. (2014) Gaze-touch: Combining gaze with multi-touch for interaction on the same surface. In *Proceedings of the 27th Annual Symposium on User Interface Software and Technology (UIST '14)*, pp. 509–518, Honolulu, HI, USA. https://doi.org/10.1145/2642918.2647397.

Pfeuffer, K., Alexander, J., Chong, M. K., Zhang, Y. and Gellersen, H. (2015) Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. In *Proceedings of the 28th Annual Symposium on User Interface Software and Technology (UIST '15)*, pp. 373–383, Charlotte, NC, USA.https://doi.org/10.1145/2807442.2807460.

Pfeuffer, K., Mayer, B., Mardanbegi, D. and Gellersen, H. (2017) Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*, pp. 99–108, Brighton, United Kingdom. https://doi.org/10.1145/3131277.3132180.

Rantala, J., Majaranta, P., Kangas, J., Isokoski, P., Akkil, D., Špakov, O. and Raisamo, R. (2017) Gaze interaction with vibrotactile feedback: Review and design guidelines. *J. HCI*, 1–39. https://doi.org/10.1080/07370024.2017.1306444.

Ruff, R.M. and Parker, S.B. (1993) Gender- and age-specific changes in motor speed and eye-hand coordination in adults: Normative values for the finger tapping and grooved pegboard tests. *Percept. Mot. Skills*, 76:3 (Pt 2), 1219–1230. https://doi.org/10.2466/pms.1993.76.3c.1219

Saddik, A. E., Orozco, M., Eid, M. and Cha, J. (2011) Haptics technologies: Bringing touch to multimedia. In Avizzano, C. A., Azorín, J. M. (eds), *Springer Series on Touch and Haptic Systems*. Springer-Verlag, Berlin Heidelberg.

Sallnäs, E.-L. and Zhai, S. (2003) Collaboration meets Fitts' law: Passing virtual objects with and without haptic force feedback. *In Proceedings of INTERACT 2003: IFIP Conference on Human-Computer Interaction*, Amsterdam, Netherland, pp. 97–104.

Samad, M., Gatti, E., Hermes, A., Benko, H., and Parise, C. (2019) Pseudo-haptic weight: Changing the perceived weight of virtual objects by manipulating control-display ratio. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*, Glasgow, Scotland UK, Paper No. 320. https://doi.org/10.1145/3290605.3300550

Schuetz, I., Murdison, T.S., MacKenzie, K.J., and Zannoli, M. (2019). An explanation of Fitts' law-like performance in gaze-based selection tasks using a psychophysics approach. *In Proceedings the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*, Glasgow, Scotland UK, Paper No. 535. http://doi.org/10.1145/3290605.3300765

Shih, R., Dubrowski, A., and Carnahan, H. (2009) Evidence for haptic memory. *In 3rd Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. 2009 IEEE World Haptics Conference*, Salt Lake City, UT, USA, pp. 145–149. https://doi.org/10.1109/WHC.2009.4810867

Shishkin, S.L., Velichkovsky, B.M., Melnichuk, E.V., Dubynin, I.A., Zhao, D.G., and Isachenko, A.V. (2018). The Pursuing Gaze Beats Mouse in Non-Pop-Out Target Selection. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Miyazaki, Japan. http://doi.org/10.1109/smc.2018.00595

Stellmach, S. and Dachselt, R. (2012) Look & Touch: Gaze-supported target acquisition. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, Austin, Texas, USA, pp. 2981–2990. http://doi.org/10.1145/2207676.2208709

Tanriverdi, V. and Jacob, R.J.K. (2000) Interacting with eye movements in virtual environments. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*, Hague, Netherlands, pp. 265–272. https://doi.org/10.1145/332040.332443

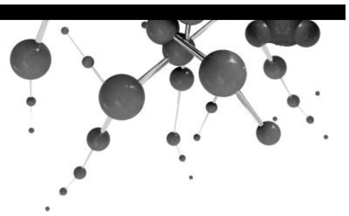
Tobii. https://www.tobii.com/ (retrieved February 20, 2019)

Williams, S.R. and Chapman, C.E. (2002) Time course and magnitude of movement-related gating of tactile detection in humans. III. Effect of motor tasks. *J. Neurophysiol.*, 88, 4: 1968–1979. https://doi.org/10.1152/jn.2002.88.4.1968

Wobbrock, J. O., Findlater, L., Gergle, D. and Higgins, J. J. (2011) The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, pp. 143–146, Vancouver, BC, Canada. http://doi.org/10.1145/1978942.1978963.

Zhai, S., Morimoto, C. and Ihde, S. (1999) Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*, pp. 246–253. Pittsburgh, Pennsylvania, USA. http://doi.org/10.1145/302979.303053.

Zubrycki, I. and Granosik, G. (2017) Novel haptic device using jamming principle for providing kinaesthetic feedback in glove-based control interface. *J. Intelligent and Robotic Systems*, 85: 3–4, 413–429. http://doi.org/10.1007/s10846-016-0392-6



Paper IV

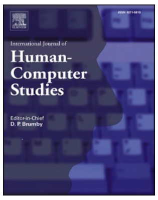
Zhenxing Li, Maria Kiiveri, Jussi Rantala and Roope Raisamo (2021). Evaluation of Haptic Virtual Reality User Interfaces for Medical Marking on 3D Models. *International Journal of Human Computer Studies*, Volume 147, 102561.
https://doi.org/10.1016/j.ijhcs.2020.102561

Int. J. Human–Computer Studies 147 (2021) 102561

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs



Evaluation of haptic virtual reality user interfaces for medical marking on 3D models

Zhenxing Li [*], Maria Kiiveri, Jussi Rantala, Roope Raisamo

Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland

ARTICLE INFO

Keywords:
3D visualization
Virtual reality
Haptic feedback
Force-feedback device
Medical marking

ABSTRACT

Three-dimensional (3D) visualization has been widely used in computer-aided medical diagnosis and planning. To interact with 3D models, current user interfaces in medical systems mainly rely on the traditional 2D interaction techniques by employing a mouse and a 2D display. There are promising haptic virtual reality (VR) interfaces which can enable intuitive and realistic 3D interaction by using VR equipment and haptic devices. However, the practical usability of the haptic VR interfaces in this medical field remains unexplored. In this study, we propose two haptic VR interfaces, a vibrotactile VR interface and a kinesthetic VR interface, for medical diagnosis and planning on volumetric medical images. The vibrotactile VR interface used a head-mounted VR display as the visual output channel and a VR controller with vibrotactile feedback as the manipulation tool. Similarly, the kinesthetic VR interface used a head-mounted VR display as the visual output channel and a kinesthetic force-feedback device as the manipulation tool. We evaluated these two VR interfaces in an experiment involving medical marking on 3D models, by comparing them with the present state-of-the-art 2D interface as the baseline. The results showed that the kinesthetic VR interface performed the best in terms of marking accuracy, whereas the vibrotactile VR interface performed the best in terms of task completion time. Overall, the participants preferred to use the kinesthetic VR interface for the medical task.

1. Introduction

Three-dimensional (3D) visualization has been widely used in many professional fields, such as medicine (Preim and Bartz, 2007), architecture and industrial manufacturing (Bouchlaghem et al., 2005). In medicine, 3D visualizations of the human skeleton, organs and other anatomic structures are implemented based on radiological imaging, such as computed tomography (CT) and magnetic resonance imaging (MRI) scans (Sutton, 1993). 3D visualization technique offers numerous benefits. For example, volumetric medical images can enable medical students to better understand the spatial anatomy of body organs (Silén et al., 2008), improve the accuracy of medical diagnoses (Satava and Robb, 1997) and help surgeons plan and simulate surgical procedures (Gross, 1998).

Highlighting relevant points on volumetric medical images of CT and MRI scans is an important 3D manipulation task performed by medical practitioners in computer-aided medical diagnosis and planning. Medical practitioners manipulate the models (i.e., rotate, pan and zoom) and mark critical points for later inspection, measurement and analysis of skeletal relationships (Kula and Ghoneima, 2018), treatment planning (Harrell, 2007) and as a tool for discussing and developing treatment consensus (Reinschluessel et al., 2019). For example, during cephalometric tracing, medical practitioners select and mark a point on the skeleton model or surrounding soft tissue as a point of reference for operations related to positioning, measurement and orientation. The task difficulty depends on the marking locations and the structure complexity of the virtual models (Medellín-Castillo et al., 2016). The accuracy of the markers directly influences the results of the medical analyses, and thus, the overall quality of the medical services (Lindner et al., 2016).

Despite the recent advances in 3D visualization technology, the tools used to present and interact with these volumetric images have not changed in the field of medicine. A conventional 2D display is still the main visual channel to present volumetric data from CT and MRI scans, which provides the user with a fixed screen-based viewing perspective. Further, it is still a common practice to use a mouse with the rotate-pan-zoom technique to indirectly manipulate 3D models (Jankowski and Hachet, 2013). However, previous studies (Hinckley et al., 1997;

---

[*] Corresponding author
E-mail address: Zhenxing.li@tuni.fi (Z. Li).

https://doi.org/10.1016/j.ijhcs.2020.102561
Received 31 July 2019; Received in revised form 24 September 2020; Accepted 20 October 2020

Bowman et al., 2004) have argued that using the mouse-based interface for 3D manipulation is difficult. Some researchers have investigated the mouse-based rotation techniques to understand their issues for 3D manipulation (Bade et al., 2005). Other researchers have conducted comparative studies to examine the usability of other user interfaces such as the tangible interface (Besançon et al., 2017) and the touchscreen-based interface (Yu et al., 2010). However, these interaction methods either did not exceed the performance of the mouse (Yu et al., 2010) or had a limited application area (Besançon et al., 2017).

Following the technical advances in virtual reality (VR), VR equipment (e.g., Oculus Rift (Oculus, 2020) and HTC Vive (VIVE, 2020)) has been developed. A combination of a VR headset and a handheld VR controller can provide the user with an intuitive and immersive interaction environment. In this VR interface, 3D models are presented to the user through the head-mounted display and the models can be manipulated by the user using the VR controller with six degrees of freedom (Oculus, 2020; VIVE, 2020). Compared with the traditional 2D interface, VR devices offer a flexible 3D view based on the position and orientation of the user's head and allow using 3D hand gestures to manipulate the objects. In addition, the VR controller can provide vibrotactile feedback to the user's hand and enable tactile interaction. Vibrotactile feedback as an augmentative sensory channel has many medical applications, such as, robot-assisted teleoperation (Peddamatham et al., 2008), minimally invasive surgery (Schoonmaker and Cao, 2006) and rehabilitation medicine (Shing et al., 2003). Because of the flexible viewing perspective and the natural hand-based input, the VR interface has been proposed to use in the field of medicine. For example, it has been employed to interact with skeleton and organ models for anatomy learning (Fahmi et al., 2019) and treatment planning (Reinschluessel et al., 2019). Multiple companies have employed it to develop software for medical diagnosis services (e.g., Surgical Theater, 2020; Adesante, 2020). However, the potentially beneficial vibrotactile feedback generated from the VR controller was not used in their interactive VR systems.

Further, force-feedback devices, such as the Geomagic Touch (3D systems, 2020) and the Novint Falcon (Novint, 2020), have been proposed as another beneficial interaction device for medical services (Ribeiro et al., 2016). These devices can support bidirectional kinesthetic exploration. The mechanical arm of the devices not only allows hand-based motions with six degrees of freedom for object manipulation but also transfers the generated kinesthetic feedback to the hand, to simulate the feeling of touch (Massie and Salisbury, 1994). Force-feedback devices have been used with a 2D display for, for example, anatomy education (Kinnison et al., 2009), surgery training (Steinberg et al., 2007; Webster et al., 2004) and medical analysis (Medellín-Castillo et al., 2016). The only study that has combined the force-feedback device with the VR headset, to the best of our knowledge, is the work by Saad et al. (2018). They have technically investigated the feasibility to connect these devices.

We combined a VR headset with a VR controller and a force-feedback device to create haptic VR interfaces which provide the user with a flexible viewing perspective, a natural hand-based input and haptic feedback simultaneously. These VR interfaces can enable intuitive and realistic 3D interaction, thus promising for the tasks involving 3D manipulation (Bowman et al., 2004) such as medical diagnosis and planning tasks. However, the usability of the haptic VR interfaces for these medical tasks, covering effectiveness, efficiency and satisfaction (Issa and Isaias, 2015), has not been explored. Furthermore, these two VR interfaces are based on similar interaction models but employ different interaction devices with different types of haptic feedback. Their difference in usability remains unclear in the context of medical diagnosis and planning. A comparison of two VR interfaces can help better understand the suitability of their interaction methods for 3D manipulation and reveal the effects of different types of haptic feedback for these high-standard medical tasks. More importantly, 2D interaction method using a mouse and a 2D display is still a powerful user interface and dominant in the field of medicine. A comparative study with the 2D interaction technique is necessary to explore the potential of the haptic VR interfaces to improve current medical diagnosis and planning work.

In the present study, we examined the two haptic VR interfaces, the kinesthetic VR interface using a force-feedback device and the vibrotactile VR interface using a VR controller, in an experiment involving medical marking on 3D models. To examine their practical usability, we compared two VR interfaces with the traditional 2D interface that uses a mouse and a 2D display as the baseline. In the experiment, because the structural complexity of the models and the marking locations can influence users' performance in the medical marking task, we employed three human anatomic structures as the experimental models with two different difficulty levels for the marking positions. To evaluate the three user interfaces, we collected both objective and subjective data. The objective data included task completion time and marking accuracy, and the subjective data included rating data for the perceived mental effort, hand fatigue, naturalness, immersiveness and user preference. The aim of the study was to answer the following questions in the context of medical marking:

- What are the differences between the kinesthetic VR interface and the vibrotactile VR interface, in terms of task completion time, marking accuracy and user experience? How do the marking locations affect users' performance with the two VR interfaces?
- What are the differences between the two VR interfaces and the traditional 2D interface? How do the marking locations affect users' performance with the traditional 2D interface?

This study makes the following contributions: We proposed two haptic VR interfaces to interact with volumetric medical images for computer-aided medical diagnosis and planning. The vibrotactile VR interface and the kinesthetic VR interface were evaluated based on a medical diagnosis and planning task on virtual models of the human skeleton and organ. The results revealed the strengths and weaknesses of two VR interfaces associated with current popular VR equipment and haptic device, which simultaneously provided empirical understanding for developing efficient and user-friendly interactive VR systems. In addition, through comparing with the 2D interaction technique, the better performances of two VR interfaces, in terms of marking accuracy (the kinesthetic VR interface) and task completion time (the vibrotactile VR interface), demonstrated their potential to replace the traditional 2D interface for these medical tasks.

The paper is organized as follows. Relevant previous studies are introduced, and then the prototype system and the experiment are described. The results are presented in detail, followed by the discussion of the main findings and the conclusion.

2. Background

2.1. 2D and 3D visualization in the field of medicine

In current medical imaging systems, the most common information visualization is based on 2D slices. Slice-by-slice views support accurate exploration and diagnosis of medical imaging data (Tietjen et al., 2006). At the same time, 3D volume-rendering visualization has become a valuable technique in the diagnosis and planning phases. It helps medical staff in understanding 3D spatial relations and an overview of the model structure, as well as facilitates diagnostic analysis (Tietjen et al., 2006). For example, cephalometry analysis is an important tool in orthodontics (Kula and Ghoneima, 2018). Traditional cephalometry analysis on 2D slices suffers from visual distortion of skull structures and inaccurate marking locations, due to the overlap of skull structures in the 2D view (Lindner et al., 2016; Bholsithi et al., 2009). Some studies (Olszewski et al., 2007; Katsumata et al., 2005; Troulis et al., 2002) showed that cephalometry analysis on 3D models can improve the precision of the diagnoses. However, other studies (Van Vlijmen et al., 2010; Swennen and Schutyser, 2007) argued that compared to 2D-cephalometry, marking on the models during 3D-cephalometry analysis is difficult and time-consuming. These studies as examples indicated the importance of accurate medical marking on 3D models, but all were conducted using the traditional 2D interface based on a mouse and a 2D display.

In the present study, we contributed to this line of research by introducing two haptic VR interfaces to address the issues of medical marking on 3D models. Because the 3D models and the marking locations in medical diagnosis vary depending on the specific medical purpose, the experiment of this study involved marking tasks using several models of different parts of the human body with multiple selected marking positions. The aim was to examine the general usability of the VR interfaces for high-standard medical diagnosis and planning.

2.2. *The dominant 2D interactive system*

In traditional human-computer interaction systems, a 2D display is commonly used as the visual output channel, while the input may be provided with devices such as mice, trackballs (Imagine Media, 1996), joysticks and touch pads. Among the interaction devices, a mouse with a 2D display constitutes the most popular interactive system used in the field of medicine. To visually present 3D models using a 2D screen, a 3D projection technique is needed. By transforming and mapping 3D objects onto a 2D plane, the projection provides the visual effect of 3D models on the 2D screen as realistic as the human visual view in the physical world (Foley et al., 1995). However, the viewing angle of interacting with the models is often fixed, due to the nature of the 2D screen.

To manipulate 3D models, the rotate-pan-zoom technique is commonly used with mouse-based interfaces (Jankowski and Hachet, 2013). For example, to rotate the models, rotation techniques, such as Virtual Sphere (Chen et al., 1988) and ArcBall (Shoemake, 1992), are widely used. Both techniques adopt a virtual ball around the manipulating object and calculate the rotation axis and angle by utilizing the projection of the mouse location onto the sphere (Jankowski and Hachet, 2013). The panning operation is typically implemented by using the mouse to point at the object and dragging the mouse along the x-y plane while pressing a button of the mouse. The zooming function often adopts the method of rolling the mouse wheel with a discrete zooming step. The mouse-based user interface relies on these techniques to indirectly manipulate 3D models.

However, using 2D interaction devices such as the mouse to manipulate 3D models can be difficult (Bowman et al., 2004). For example, a previous study demonstrated that using a multidimensional interaction device could achieve more efficient interaction than using the mouse-based interface in a 3D manipulation task (Hinckley et al., 1997). Bade et al. (2005) have compared the existing mouse-based 3D rotation techniques and provided their design principles to address this issue. In addition, multiple studies have examined other interaction methods for 3D manipulation tasks. Yu et al. (2010) presented a touchscreen-based data exploration technique for 3D manipulation. The interaction method was easy to learn and use, but its task performance could not exceed the performance using the mouse-based interface. Besançon et al. (2017) compared a tangible interface and a touchscreen interface with a mouse-based interface. The results showed that using the tangible interface could lead to a shorter task completion time. However, it had a limited application area compared to the mouse. There are advanced 3D interactive systems implemented by using VR equipment and force-feedback devices. In this study, we experimentally compared these interactive VR systems with the mouse-based interface and investigated their strengths and weaknesses in medical marking.

2.3. *The interactive VR system using a VR headset and a VR controller*

VR equipment offers a new interactive experience with virtual objects in the field of medicine. Previous studies have used the CAVE environment for visualizing medical imaging data (Shen et al., 2008; Al-Khalifah et al., 2006). Currently the head-mounted VR display presents a more flexible and natural 3D view based on the position and orientation of the user's head. Many studies have employed the VR headset to visually present 2D slices (Wirth et al., 2018; King et al., 2016) and volumetric imaging data (Venson et al., 2017; Sousa et al., 2017; Randall et al., 2016) for radiologists. Other studies proposed to use the VR headset for the medical purposes such as treating chronic pain (Jones et al., 2016), anxiety disorders and phobias (Maples-Keller et al., 2017).

The VR controller allows 3D hand gestures to manipulate objects with optional vibrotactile feedback which can enable tactile interaction (Oculus, 2020; VIVE, 2020). Tactile interaction, as one branch of haptic interaction, concentrates on touch interaction that stimulates mechanoreceptors in the human skin (El Saddik et al., 2011). This technique has been widely used in many medical fields. For example, vibrotactile feedback was used to enrich interaction channels while using a surgical robot (Peddamatham et al., 2008), improve surgeons' performance in minimally invasive surgery (Schoonmaker and Cao, 2006) and help users in using a hand rehabilitation system (Shing et al., 2003).

Because of the natural interaction method, the VR interface using a VR headset and a VR controller has been used in an anatomy learning system (Fahmi et al., 2019). Reinschluessel et al. (2019) have proposed to use it to interact with 3D organ models for treatment planning. Some products such as Precision VR by Surgical Theater (2020) and SurgeryVision by Adesante (2020) have employed the VR interface to manipulate 3D models of human anatomic structures for medical analysis and other services. However, the vibrotactile feedback generated from the controller was missing in their interactive VR systems.

While the VR interface has been widely used, it remains unclear how well it compares with the traditional 2D interface using a mouse and a 2D display. It is likely that the benefit and cost are dependent on the context of the interaction such as the type of the task and the complexity of the model. For example, a previous study in the field of 3D geological modelling, comparing a VR controller with a mouse in manipulating an industrial software, showed that the controller was more difficult to use and caused more fatigue on the user's hand (Kim and Choi, 2019). In the present study, we evaluated the efficiency and user experience of this VR interface with vibrotactile feedback to interact with 3D objects in the context of medical marking.

2.4. *The interactive VR system using a VR headset and a force-feedback device*

Using force-feedback devices enables kinesthetic interaction which is another type of haptic interaction (El Saddik et al., 2011). The kinesthetic technique focuses on movement sensations originating in the muscles, tendons and joints (El Saddik et al., 2011). It enables bidirectional touch exploration closely as realistically as in the physical world and has been used for many medical practices, such as medical education (Kinnison et al., 2009), surgery operation simulation and training (Alaraj et al., 2015; Steinberg et al., 2007; Webster et al., 2004; Bielser and Gross, 2000), robot-assisted surgery (Okamura, 2004) and medical analysis (Medellín-Castillo et al., 2016).

For medical education, a kinesthetic simulator (Kinnison et al., 2009) has been demonstrated to be an engaging and efficient method for teaching students the human anatomy, which addresses current challenges in teaching using real human tissues (e.g., anxiety and fear). Other studies have suggested that kinesthetic simulation is an efficient and repeatable method for surgery training without wasting real surgical samples. Kinesthetic simulators have previously been used, for instance, in dental (Steinberg et al., 2007) and eye (Webster et al., 2004) surgery simulations, basic cutting (Bielser and Gross, 2000) and aneurysm clipping simulation (Alaraj et al., 2015). In addition, the kinesthetic technique brings benefits in the context of robot-assisted surgery. For instance, a previous study (Okamura, 2004) has argued that offering kinesthetic feedback contributes significantly to the safe performance of the surgical procedures. In medical analysis, kinesthetic feedback was used to enhance medical marking in 2D, 2.5D and 3D computer-aided cephalometry analysis (Medellín-Castillo et al., 2016). The results indicated that haptic-enabled 3D cephalometry marking is better than the haptic-enabled 2D or 2.5D cephalometry marking, in terms of task accuracy.

Combining a force-feedback device with a VR headset as a new VR interface to interact with 3D models is feasible (Saad et al., 2018). In the present study, we evaluated two haptic VR interfaces (i.e., using either a VR controller or a force-feedback device) and examined the different performances of vibrotactile and kinesthetic feedback in medical marking. To demonstrate the practical usability, we compared these two interfaces with the traditional 2D interface.

3. Method

3.1. Design of the prototype system

The experimental prototype system included three user interfaces: vibrotactile VR interface (V), kinesthetic VR interface (K) and traditional 2D interface (T). We use the abbreviations (V, K, T) for the three interfaces in figures.

The vibrotactile VR interface employed a VR headset as the visual output channel and a VR controller as the manipulation tool with vibrotactile feedback. The VR headset provided realistic visual feedback with a flexible head movement-based viewing perspective. While pressing the trigger button of the controller with the index finger (see Fig. 1A), the user could manipulate the model by rotating (rotation along the x-, y- and z-axes), panning (displacement along the x- and y-axes) and zooming (displacement along the z-axis), following the movement of the whole arm (i.e., rotation was done by rotating the wrist, and panning and zooming were done by moving the arm along the x-, y- and z-axes respectively). In addition, the user could feel vibrotactile feedback from the controller while the cursor touched the model. The cursor was visible as a yellow sphere in the virtual environment with a 0.2 cm radius hanging on the tip of the controller. The controller was invisible to the user in the virtual environment. The duration of the haptic pulse for the vibrotactile feedback was set to 80 ms.

The kinesthetic VR interface used the same VR headset as the visual output channel, but a force-feedback device was used as the manipulation tool. Consistent with the vibrotactile VR interface, the user could rotate, pan and zoom the model with the movement of the whole arm while pressing the button on the device with the index finger (see Fig. 1B). When the user touched the model with the cursor (visible as a yellow sphere with a 0.2 cm radius hanging on the tip of the device, and the device was invisible in virtual environment), he or she could feel kinesthetic feedback based on the surface properties of the model. The kinesthetic feedback involved the stiffness implemented by the linear spring law ($F = kx$), where k is the stiffness coefficient and x is the penetration depth of the interaction point, as well as the friction ($F' = \mu F_n$), where μ is the friction coefficient and $F_n = mg$, the normal force value.

With both haptic VR interfaces, the user could move the cursor to reach and touch any position on the medical model as the selected marking point for the experimental task.

The traditional 2D interface employed a standard 2D display as the visual output channel and a mouse as the manipulation tool. The 2D display also provided realistic visual feedback with a screen-based viewing angle. To rotate 3D models, the user could move the mouse along the x- and y-axes on the screen while holding down the left mouse button (see Fig. 1C). The rotation was based on the ArcBall technique with a selected rotation speed parameter of 3.5. Panning of the model (along the x- and y- axes) was implemented by dragging the model using the mouse while pressing down the right mouse button. Zooming of the model was done by rolling the mouse wheel with a selected zooming speed factor of 6. Marking on the model was based on the ray-casting technique (Poupyrev et al., 1998). This technique utilized the pointer position of the mouse on the 2D screen as the starting point of a virtual beam that traversed the virtual space along the viewing direction of the 2D screen. When the beam intersected with the 3D model, the intersecting point would be the selected marking point (Gallo et al., 2010). In addition, the mouse-based interface had no haptic feedback.

For all three user interfaces, the space key on a keyboard was used to create markers (visible as green nodes with a 0.2 cm radius) on the

Table 1
Specification of three user interfaces.

Interfaces	Equipment	Hand-based operation	Visual channel	Haptic feedback
Vibrotactile VR interface (V)	VR headset and controller	3D gesture along the x-, y- and z- axes	Head movement-based view	Vibrotactile
Kinesthetic VR interface (K)	VR headset and force-feedback device	3D gesture along the x-, y- and z- axes	Head movement-based view	Kinesthetic
Traditional 2D interface (T)	2D display and mouse	2D gesture along the x- and y- axes	Screen-based view	None

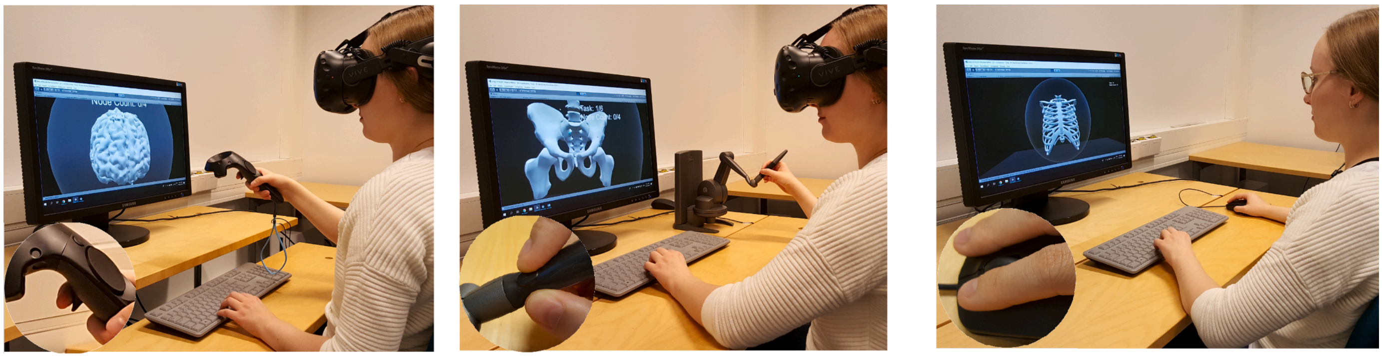

(A) (B) (C)

Fig. 1. (A) The vibrotactile VR interface using a Vive controller and a Vive VR headset. A close-up image of the trigger button of the controller is shown at the bottom left corner. The additional Samsung LCD 2D display (not visible to the user) shows the brain model. (B) The kinesthetic VR interface using a Geomagic Touch X force-feedback device and a Vive VR headset. A close-up image of the device button is shown at the bottom left corner. The additional 2D display shows the hipbone model. (C) The traditional 2D interface using a standard mouse with the Samsung LCD 2D display. A close-up image of the mouse buttons is shown at the bottom left corner. The screen shows the sternum model. With all three user interfaces, the space key on a standard keyboard was used to create the markers on the marking positions.

selected marking points. The three interfaces with the device buttons are shown in Fig. 1 and the specification of the user interfaces are listed in Table 1.

3.2. *Experiment design*

A within-subject experiment was conducted in a controlled laboratory setting with three experimental conditions: the vibrotactile VR interface, the kinesthetic VR interface and the traditional 2D interface. The experimental task was medical marking on 3D models, and there were three models involved in the experiment: the sternum, hipbone and brain (see Fig. 1). All models were inside a 30 cm diameter sphere. This ensured that the models were of a similar size. The selected stiffness coefficients (sternum and hipbone: 1.0; brain: 0.4) and friction coefficient (all: 0.7) of the models were set for kinesthetic feedback.

The task-related marking positions (visible as red dots with a 0.1 cm radius) were predefined on the models before the experiment. To mark these positions, the participants employed the interaction tools to manipulate the model for visually searching the marking positions (through the visual display) and then reach (using the VR controller or the force-feedback device) or point (using the mouse) at the positions on the model for determining the positions the participants wanted to mark. Markers were created on the positions by pressing the space key on the keyboard.

We varied the marking positions as an independent variable in the experiment. Because of the complex characteristic and structure of human anatomic models, we generally categorized the marking positions into two difficulty levels (easy and difficult). We placed the model away from the viewpoint (45 cm from the viewpoint to the centre of the model along z-axis) to make the model fully visible to the user and then determined the difficulty level of the marking position based on its visibility from the user's view. If the marking position could be visually blocked by other parts of the model surrounding it while rotating the model (excluding the situation of rotating to other faces of the model which made the marking surface completely invisible), we categorized the position as difficult. In this case, the user had to manipulate the model carefully and find an appropriate angle to mark the position. In contrast, if the marking position could not be visually covered by other parts of the model surrounding it while rotating the model, the user was able to spend less effort manipulating the model and marking the position and thus we categorized the position as easy (Fig. 2 shows examples of two marking difficulty levels). We predefined 24 possible marking positions on the surface of each model (12 for easy to mark and 12 for difficult to mark) for the experiment task.

The participants performed six trials with each interface: three trials were easy to mark and three trials were difficult to mark. Each trial involved only one 3D model. The order of the three models (i.e., the sternum, hipbone and brain) was randomly assigned for each difficulty level, and thus, the models were used twice, once for each difficulty level. In each trial, the participants needed to mark 4 positions randomly selected from 12 possible positions based on the marking difficulty level of the trial. Therefore, for each participant, there were a total of 18 trials (3 interfaces × 6 trials = 18 trials) with 72 marking positions (18 trials × 4 positions = 72 positions).

The prototype system recorded the task completion times and the positions of the participants' markers as objective data. In addition, a 7-point Likert scale questionnaire was used to record subjective data, including perceived mental effort, hand fatigue, naturalness and immersiveness (1: strongly disagree to 7: strongly agree). The questionnaire was inspired by the NASA Task Load Index (Hart and Staveland, 1988). The statements used in the questionnaire are shown in Table 2. Moreover, we used a post-test questionnaire in which the participants ranked the three user interfaces based on personal preference and gave comments.

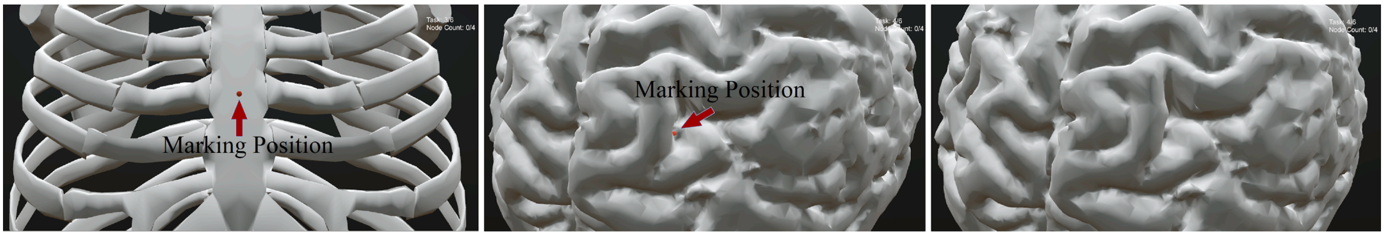
3.3. *Pilot study*

We recruited four participants who had previous experience using the VR equipment and the force-feedback device to conduct a pilot study. The participants were asked to use the prototype system and comment which parameters were the most suitable.

- For the kinesthetic VR interface, we did not scale the movement (Fischer and Vance, 2003), that is, the movement of the hand holding the mechanical arm resulted in the same amount of movement of the cursor in the virtual space along the x-, y- and z-axes. This enabled more accurate control of the cursor (Li et al., 2020).
- For the traditional 2D interface, the parameter for the rotation (speed factor 3.5) was selected, so that users could quickly and accurately rotate the models. The zooming factor (6) was selected based on the distance between the viewpoint and the target model. With maximum zooming, the viewpoint could closely reach the model surface. By applying these parameters, users could smoothly use the mouse to manipulate the models for the medical task without any operational difficulty.
- For the vibrotactile VR interface, the duration of the vibrotactile feedback (80 ms) was selected. With this duration, the vibrotactile feedback could be easily perceived when the user touched the model, but it was not too disruptive to influence users' hand posture and stability.
- The sizes of the three experimental models within the sphere (30 cm diameter) were chosen so that they were suitable for the size of the

Table 2
Statements in the questionnaire.

Statements	Description
S1	This user interface is mentally easy to use.
S2	This user interface does not make the hand tired.
S3	This user interface is natural to use.
S4	This user interface is immersive.



(A) (B) (C)

Fig. 2. (A) shows an example of the easy-to-mark task in which the marking position (red dot) was distinct on the model surface. (B) shows an example of the difficult-to-mark task in which the marking position was only visible from some viewing angles. (C) shows that the marking position could not be visible from some other viewing angles.

workspace of the force-feedback device and the area of hand movement using the controller.
- The size of the interaction point (a sphere with a 0.2 cm radius) of two haptic interaction devices was chosen. It ensures that the interaction point was easy to control while simultaneously minimizing its influence on the marking accuracy. The size of the task-required marking positions (dots with a 0.1 cm radius) was selected so that it maintained a high-level requirement on the marking accuracy to examine the three user interfaces. At the same time, the size of the marking positions was big enough and thus clearly visible to the user.
- For the kinesthetic VR interface, stiffness coefficients (sternum and hipbone: 1.0 and brain: 0.4) and friction coefficient (all: 0.7) were chosen so that touching the virtual objects would resemble touching real objects as closely as possible.

The above parameter selections were made for the prototype system and the experiment, based on the users' performance and comments. We next introduce the experiment apparatus, participants and procedure.

4. Experiment

4.1. Apparatus and environment

The host computer in the experiment was an MSI GS63VR 7RF Stealth Pro laptop with an Intel i7-7700HQ processor, a GeForce GTX 1060 graphics card and a 16GB RAM. We used a Vive VR headset (VIVE, 2020) and a Samsung 245B Plus 24" LCD display as the visual displays in the experiment. A standard computer mouse, a Vive controller and a Geomagic Touch X force-feedback device (3D Systems, 2020) were utilized as the manipulation tools for the three user interfaces. A standard keyboard was used for the participants to mark the positions. The experimental setup and environment are shown in Fig. 1. The experimental system was developed on Unity (2020), along with SteamVR (Steam, 2020) and Geomagic OpenHaptics plugin (Unity haptic plugin, 2020).

4.2. Participants

Twenty-four participants were recruited from the local university community (9 women and 15 men). Their ages varied between 21 and 43 years (M = 28.33, SD = 7.65). No participants had previous experience with using a mouse and a 2D display to do medical marking or similar tasks. Seven participants had used a similar VR headset and a controller one or two times. Two participants had used a similar force-feedback device with a 2D display once, but they had no use experience with a VR headset. All participants reported that they were right-handed.

4.3. Experiment procedure

The participants were first introduced to the experimental task and the apparatus used. They signed an informed consent form and filled in the background information in the questionnaire before the experiment.

The order of the experimental conditions and the marking difficulty levels was counterbalanced among the participants. For each condition, the participants were introduced to the user interface and had up to 5 minutes to familiarize themselves with the use of the related devices. They sat comfortably on a chair and looked at the 2D display or wore the VR headset, and then used their dominant hand to hold the manipulation tool and their non-dominant hand to press the space key of keyboard to mark the positions. For the two VR interfaces, head movement to change their viewing perspective was allowed. We did not provide any extra wrist-rest or arm-rest equipment for any of the interfaces.

The total experiment time for each participant was approximately one hour. The participants were informed that they needed to finish the experimental task as accurately and quickly as possible and the accuracy was the top priority. When each trial started, the model was placed at 45 cm away from the viewpoint along z-axis and its orientation was semi-random. We predefined a list of orientation values which could respectively make one of the six faces of the model to face the user (i.e., by considering the model as a cube) and the system randomly selected one of them as the starting orientation value for the model. For each marking position, the participants had only one chance to mark the position. In other words, they could not remove the marker and place it again. After the participants marked the four marking positions, the system automatically recorded the data and proceeded to the next trial. After finishing the six trials of one condition, the participants were asked to fill out the questionnaire, after which the next condition was executed. After completing all conditions, the participants commented the three interfaces and ranked them based on personal preference.

5. Results

We collected two types of objective data from the experiment: task completion times and the positions of the participants' markers. We used the task completion times to evaluate the interaction speed and calculated the error distances between the participants' markers and the marking positions required by the tasks to evaluate the marking accuracy. The Shapiro–Wilk Normality test showed that the data were not normally distributed (all $p < .001$). Thus, we analyzed the data using the 3 × 2 (interfaces × marking difficulty levels) aligned rank transform (ART) repeated-measures non-parametric ANOVA (Wobbrock et al., 2011) separately for the task completion times and the error distances. We used the Wilcoxon signed-rank test for post-hoc analysis of objective and subjective data and performed the Holm-modified Bonferroni correction (Holm, 1979) to control the family-wise type-1 error.

5.1. Task completion time

To evaluate the interaction speed of the three interfaces, we calculated the mean value of the task completion times of the six trials for each interface. Table 3 shows the p values of the ART ANOVA result for the task completion times. There were statistically significant main effects for both the interfaces and the marking difficulty levels. In addition, there was a statistically significant interaction effect between the interfaces and the marking difficulty levels.

Fig. 3 shows a boxplot of the task completion times with the three interfaces. The Wilcoxon signed-rank test showed that participants spent statistically significantly less time to complete the task while using the vibrotactile VR interface (M = 103.52, SD = 27.82) than while using the kinesthetic VR interface (M = 131.01, SD = 37.00; Z = − 3.514, $p < .001$). The task completion time using the traditional 2D interface (M = 120.62, SD = 38.08) was not statistically significantly different from that using the vibrotactile VR interface (Z = − 1.771, $p = .152$) and that using the kinesthetic VR interface (Z = − 1.657, $p = .097$).

Not surprisingly, the main effect of the marking difficulty was statistically significant. Completing difficult marking trials (M = 173.44, SD = 50.76) took statistically significantly more time than completing easy marking trials (M = 63.33, SD = 12.49; Z = − 4.286, $p < .001$). The more interesting part from the perspective of this study is the interaction effect between the interfaces and the marking difficulty levels.

Fig. 4 illustrates the task completion times of the three user interfaces

Table 3
Tests of within-subjects effects on task completion times.

Sources	df	F-value	Sig.
Interfaces	2, 46	11.928	<0.001
Marking difficulty levels	1, 23	438.442	<0.001
Interfaces and marking difficulty levels	2, 46	8.963	0.001

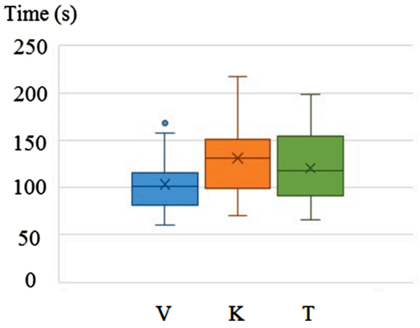


Fig. 3. Overall task completion times for the three interfaces: the vibrotactile VR interface (V), the kinesthetic VR interface (K) and the traditional 2D interface (T). The cross mark shows the mean value and the line in the boxplot shows the median value (The same abbreviations and marks are used in the following figures).

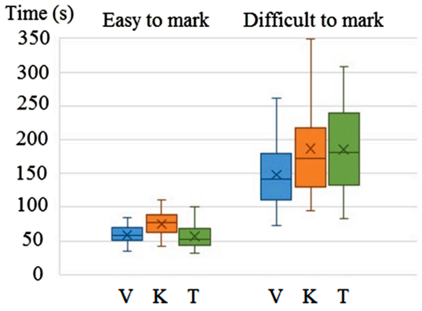


Fig. 4. Task completion times using the three user interfaces, based on the marking difficulty levels.

based on the marking difficulty levels. According to the Wilcoxon signed-rank test, when the positions were easy to mark, the participants spent statistically significantly longer time using the kinesthetic VR interface (M = 74.83, SD = 17.77) than using the vibrotactile VR interface (M = 58.81, SD = 13.45; Z = – 3.200, *p* = .002) and using the traditional 2D interface (M = 56.35, SD = 18.61; Z = – 3.400, *p* = .003). There was no statistically significant difference in task completion times between the vibrotactile VR interface and the traditional 2D interface (Z = – 0.800, *p* = .424). When the positions were difficult to mark, using the vibrotactile VR interface (M = 148.23, SD = 50.44) resulted in shorter task completion times than using the kinesthetic VR interface (M = 187.18, SD = 65.80; Z = – 3.114, *p* = .006) and using the traditional 2D interface (M = 184.90, SD = 64.39; Z = – 2.239, *p* = .05). There was no statistically significant difference in task completion times between the kinesthetic VR interface and the traditional 2D interface (Z = – 0.514, *p* = .607).

5.2. *Marking accuracy*

To evaluate the marking accuracy of the three interfaces, we calculated the mean error distance for each trial consisting of four marking positions. Table 4 shows the results of the ART ANOVA for the distance data: the main effects of the interfaces and the marking difficulty levels, as well as their interaction effect, were statistically significant.

Table 4
Tests of within-subjects effects on distance data.

Sources	df	F-value	Sig.
Interfaces	2, 46	38.893	<0.001
Marking difficulty levels	1, 23	124.652	<0.001
Interfaces and marking difficulty levels	2, 46	39.728	<0.001

Fig. 5 shows a boxplot of the marking accuracy using the three interfaces. According to the Wilcoxon signed-rank test, the participants were statistically significantly more accurate in marking when using the kinesthetic VR interface (M = 0.13, SD = 0.04) than when using the vibrotactile VR interface (M = 0.35, SD = 0.20; Z = – 4.286, *p* < .001) and the traditional 2D interface (M = 0.61, SD = 0.68; Z = – 4.257, *p* < .001). There was no statistically significant difference in the marking accuracy between the vibrotactile VR interface and the traditional 2D interface (Z = – 0.743, *p* =.458). For the main effect of the marking difficulty, the participants were more accurate in marking when the positions were easy to mark (M = 0.15, SD = 0.04) than when the positions were difficult to mark (M = 0.58, SD = 0.46; Z = – 4.286, *p* < .001).

Fig. 6 shows the marking accuracy using the three interfaces based on the marking difficulty levels. The Wilcoxon signed-rank test showed that when the positions were easy to mark, using the vibrotactile VR interface (M = 0.25, SD = 0.06) was statistically significantly less accurate than using the kinesthetic VR interface (M = 0.1, SD = 0.03; Z = – 4.286, *p* < .001) and using the traditional 2D interface (M = 0.09, SD = 0.06; Z = – 4.257, *p* < .001). There was no statistically significant difference between the kinesthetic VR interface and the traditional 2D interface (Z = – 1.029, *p* = .304). When positions were difficult to mark, participants were statistically significantly more accurate in marking when using the kinesthetic VR interface (M = 0.17, SD = 0.06) than when using the vibrotactile VR interface (M = 0.44, SD = 0.38; Z = – 4.286, *p* < .001) and using the traditional 2D interface (M = 1.13, SD = 1.36; Z = – 4.257, *p* < .001). Using the vibrotactile VR interface was statistically significantly more accurate than using the traditional 2D interface (Z = – 2.057, *p* = .04).

5.3. *Subjective data*

Fig. 7 shows the subjective data for perceived mental effort, hand tiredness, naturalness and immersiveness.

Mental effort: There were no statistically significant differences among the three interfaces in terms of perceived metal effort (the vibrotactile VR interface and the kinesthetic VR interface: Z = – 0.461, *p* = .645; the kinesthetic VR interface and the traditional 2D interface: Z = – 1.685, *p* = .184; the vibrotactile VR interface and the traditional 2D interface: Z = – 2.039, *p* =.123).

Hand tiredness: The participants experienced statistically significantly more hand tiredness using the vibrotactile VR interface than using the kinesthetic VR interface (Z = – 2.913, *p* = .008) and using the traditional 2D interface (Z = – 3.725, *p* < .001). Furthermore, the participants experienced more hand tiredness using the kinesthetic VR interface than using the traditional 2D interface (Z = – 2.345, *p* =.019).

Naturalness: The participants perceived the traditional 2D interface to be statistically significantly less natural than the kinesthetic VR interface (Z = – 2.721, *p* = .014) and the vibrotactile VR interface (Z = – 2.994, *p* = .009). There was no difference between the kinesthetic VR

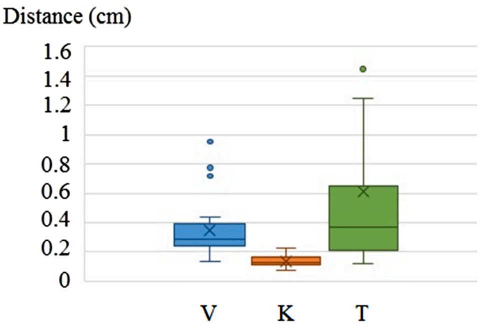


Fig. 5. Error distances between the marking positions and the participants' markers using the three interfaces.

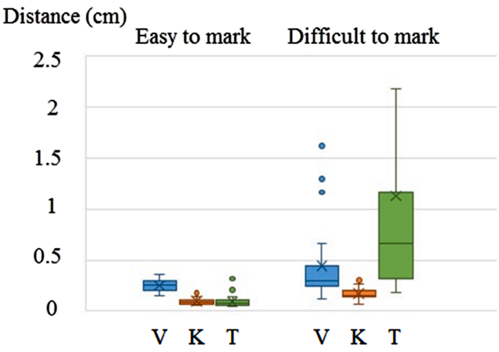


Fig. 6. Error distances between the marking positions and the participants' markers using the three interfaces, based on the marking difficulty levels.

interface and the vibrotactile VR interface in terms of the naturalness ($Z = -0.661$, $p = .509$).

Immersiveness: The participants perceived the traditional 2D interface to be statistically significantly less immersive than the kinesthetic VR interface ($Z = -4.310$, $p < .001$) and the vibrotactile VR interface ($Z = -4.233$, $p < .001$). There was no difference between the kinesthetic VR interface and the vibrotactile VR interface ($Z = -0.966$, $p = .334$).

In addition, a post-test questionnaire was employed to collect data about user preference for the three interfaces. The participants could select multiple interfaces if they liked them equally. The results showed that the kinesthetic VR interface was voted the most preferred interface for the experimental task (19 votes), followed by the vibrotactile VR interface (6 votes) and the traditional 2D interface (5 votes). The participants also provided free form comments as follows.

> P4 said, "*Controller excelled in rotation zoom and movement, but with no physical (kinesthetic) support, it felt very inaccurate for precision tasks.*" [preferred the kinesthetic VR interface and the traditional 2D interface].
>
> P8 said, "*Finding the nodes with VR was way more efficient than 2D screen. The VR display felt natural and best suited for this task.*" [preferred the kinesthetic VR interface and the vibrotactile VR interface].
>
> P22 said, "*Haptic (kinesthetic) feedback was more accurate in difficult positions. With the mouse it was almost impossible to landmark accurately. With the controller, the accuracy was also bad, especially towards the end. Haptic (kinesthetic) was much better in uneven/rough positions.*" [preferred the kinesthetic VR interface].
>
> P24 said, "*The haptic (force-feedback) device gave the most realistic impression of the object, providing depth and 3D understanding.*" [preferred the kinesthetic VR interface].

The above participants' comments as examples gave insights into their reason for preferring the interfaces. We next discuss the experiment results with the three user interfaces.

6. Discussion

This study experimentally compared two haptic VR interfaces (vibrotactile and kinesthetic VR interfaces) with the traditional 2D interface. The interfaces varied in terms of the visual display as well as the manipulation tool. They were evaluated in the experiment of 3D medical marking task and the results showed that the different user interfaces influenced the task performance in terms of interaction speed and accuracy. In addition, the difference in the marking locations modulated the task performance. We now discuss the main findings in relation to the research questions.

6.1. Differences between the kinesthetic VR interface and the vibrotactile VR interface

The kinesthetic VR interface and the vibrotactile VR interface used the VR headset as the visual display and employed a force-feedback device and a handheld VR controller as the manipulation tool respectively. The results showed a speed-accuracy trade-off between the two interfaces. Generally, the vibrotactile VR interface was faster, but less accurate, than the kinesthetic VR interface (see Fig. 3 and Fig. 5).

6.1.1. Task completion time

Regardless of the marking difficulty, using the vibrotactile VR interface led to a shorter task completion time than using the kinesthetic VR interface (see Fig. 4). The two VR interfaces used the same VR headset as the visual display and allowed for head movement-based viewing perspective changes. The difference between two interfaces was in terms of the manipulation tool.

The vibrotactile VR interface using the controller allowed free hand-based input to reach the target and provided vibrotactile feedback as confirmation of contact with the 3D model. Further, natural 3D hand gestures allowed for efficient pan, rotation and zoom of the 3D models. All the marking tasks, especially the ones with high marking difficulty, required extensive use of these complex interactions to acquire the marking locations. In contrast, the mechanical arm of the force-feedback device and its length limited the flexibility of the gestures that can be used and the area that can be interacted with in the case of the kinesthetic VR interface (Massie and Salisbury, 1994). This often meant that the rotating, panning and zooming interactions were neither as intuitive nor as straightforward to perform. For example, when the user had to perform a large pan or rotation of the 3D model which was beyond the mechanical limits of the force-feedback device, the user had to break it down into multiple smaller pan movements or rotations. This may have contributed to the lower efficiency while using the kinesthetic VR interface.

6.1.2. Marking accuracy

The kinesthetic VR interface was more accurate in medical marking

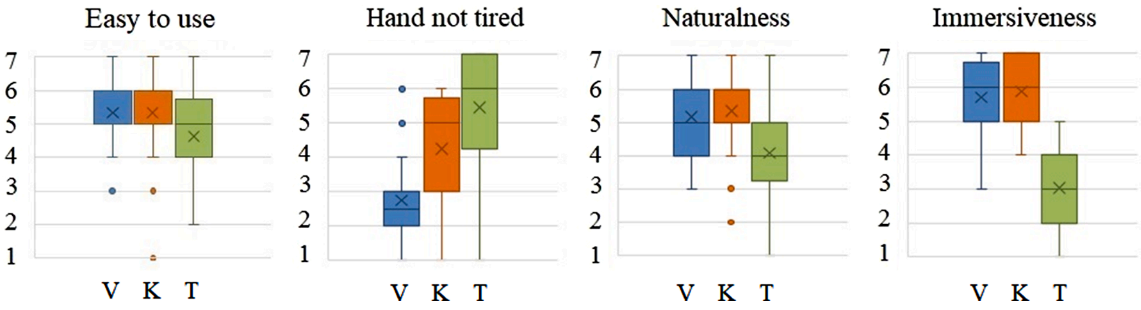


Fig. 7. Subjective results of the study.

than the vibrotactile VR interface regardless of the marking difficulty (see Fig. 6). The kinesthetic and vibrotactile VR interfaces adopted a similar interaction mechanism, requiring the participants to reach and touch the model to determine the marking position. The different performance in marking accuracy is likely because of the different types of haptic feedback. Compared with the vibrotactile feedback which provided a simple vibration for touching, the bidirectional kinesthetic interaction allowed the user to understand the 3D structure of the model and the marking locations. This meant that when the surface near the marking location was uneven (e.g., the ridges and grooves on the brain surface), the user could accurately find the target location by relying on the kinesthetic feedback. Such feedback for the 3D structure was not available in the case of the vibrotactile VR interface.

Another important difference between the kinesthetic and vibrotactile VR interfaces was regarding the interaction boundary. The force-feedback device provided kinesthetic cues about the surface that was in contact. Once in contact with a surface, if the user applied further pressure inward, the device would produce a reaction force according to the material properties of the surface. It required a considerable amount of force for the kinesthetic interaction point to penetrate the surface of the 3D model. In the case of the vibrotactile VR interface, although vibrotactile feedback was provided when the user was in contact with the surface of the 3D model, any slight inward hand movement during the marking caused the controller to pass through the surface of the model leading to a depth error in the marking.

Further, the participants commonly performed mid-air hand gestures to touch the model for marking while using the VR controller. The stability of mid-air gestures is known to show considerable individual variability (Ángyán et al., 2007) and degrade with hand fatigue (Gates and Dingwell, 2011), which might have negatively affected the marking accuracy while using the vibrotactile VR interface. Although the participants also performed mid-air hand gestures while holding the arm of the force-feedback device, the kinesthetic feedback could help maintain the hand stability during the touching process. This might further improve the marking accuracy while using the kinesthetic VR interface.

Previous studies have shown that kinesthetic feedback is a valuable complementary output modality in the fields of, for example, palpation simulation (Ribeiro et al., 2016) and robot-assisted surgery (Okamura, 2004). In the context of medical diagnosis and planning, Medellín-Castillo et al. (2016) have applied kinesthetic feedback for the 2D, 2.5D and 3D cephalometry marking and examined their differences. Our result showed that using kinesthetic feedback which allowed the user to realistically feel the model was better than using vibrotactile feedback that provided a simple tactile signal as confirmation of surface contact, in terms of marking accuracy. This finding directly demonstrated the significance of kinesthetic feedback for accurate medical marking while using hand gestures to touch the model to determine the marking position.

6.1.3. *User experience*

Based on the subjective data collected in the study (see Fig. 7), there were no statistically significant differences between the kinesthetic and vibrotactile VR interfaces, in terms of mental effort, naturalness and immersiveness. Both VR interfaces employed a VR headset, enabled changes in the head movement-based viewing perspective and allowed hand gestures to manipulate the models. It is not surprising that the participants rated both interfaces in a similar way. However, the participants reported more tiredness of the hand while using the vibrotactile VR interface compared with the kinesthetic VR interface.

The vibrotactile VR interface required the user to perform mid-air gestures, which is known to cause tiredness and heaviness of the hand especially when the interaction lasts for a long duration of time, a condition often referred as the gorilla arm effect (Hincapié-Ramos et al., 2014). The kinesthetic VR interface caused less fatigue, because the participants could rest their hand on the table while manipulating the models and on the models during the touching process.

6.2. *Differences between the two VR interfaces and the traditional 2D interface*

The traditional 2D interface that uses a 2D display and a mouse is still the most popular technique used to interact with volumetric images in the field of medicine. For the two haptic VR interfaces to be practically useful, they must perform well in comparison with the conventional 2D interface. The traditional 2D interface provided a fixed screen-based view, used a mouse-based rotate-pan-zoom technique to manipulate 3D models and used the ray-casting technique to determine the marking locations. In contrast, the two haptic VR interfaces employed a VR display for the visual output, which changed the scene perspective based on the user's head movement. Further, they relied on 3D hand gestures to manipulate the 3D models and determined the marking positions by reaching and touching the model.

6.2.1. *Task completion time*

Overall, there were no statistically significant differences between using the traditional 2D interface and using the two haptic VR interfaces, in terms of task completion time (see Fig. 3). However, a closer analysis based on the difficulty of the task revealed differences among the three interfaces. For the tasks with the low marking difficulty, the traditional 2D interface was faster than the kinesthetic VR interface. For the tasks with the high marking difficulty, the vibrotactile VR interface performed better than the traditional 2D interface and there was no difference between the kinesthetic VR interface and the traditional 2D interface (see Fig. 4).

The haptic VR interfaces and the traditional 2D interface have their own strengths and weaknesses for medical marking. On one hand, the marking method based on the ray-casting technique employed in the traditional 2D interface might lead to an improved interaction speed, due to the fast pointing capability of the mouse (along the x- and y- axes) on the 2D screen (Kim and Choi, 2019). The two VR interfaces required the user to reach and touch the model to determine the marking position. It is known that reaching an object in the 3D environment (along the x-, y- and z-axes) is difficult, and the speed of interaction is significantly influenced by the size and position of the target (Berthier et al., 1996). This especially provided an advantage for the traditional 2D interface when the marking targets were clearly presented on the surface of the 3D model without the need for any complex manipulation of the model.

On the other hand, the two haptic VR interfaces adopted hand-based 3D gestures which could implement rotating, panning and zooming simultaneously for manipulating the models, whereas the traditional mouse-based rotate-pan-zoom technique required the user to separately use three buttons for each manipulation function (i.e., the left mouse button for rotating, the right mouse button for panning and the mouse wheel for zooming). These functions could not be performed simultaneously, which might have slowed down the interaction while using the traditional 2D interface. Further, the viewing perspective change following the user's head movement, provided by the VR headset, was more flexible than the fixed viewing angle presented by the standard 2D display. The participants could rotate the 3D model while also subtly changing the viewing angle with their head movement. This likely facilitated the visual search of marking positions. Therefore, it is likely that these benefits provided a distinct advantage to the VR interfaces, especially in the tasks involving the high marking difficulty.

Previous studies have explored to use the tangible interface (Besançon et al., 2017) and the touchscreen-based interface (Yu et al., 2010) for 3D manipulation tasks instead of using a mouse-based interface. We contributed to this research area by proposing two haptic VR interfaces and examining them with the traditional 2D interface. Through comparing with the traditional 2D interface in the medical marking task, the competitive performance of the two haptic VR interfaces in task completion time, especially the better performance using the vibrotactile VR interface in the difficult tasks, demonstrated their practical usability for 3D manipulation tasks.

6.2.2. *Marking accuracy*

Unlike the two haptic VR interfaces which required the user to reach the model to mark the positions, the traditional 2D interface used ray-casting to estimate the point of marking (Gallo et al., 2010). During the marking process, the participants commonly made displacement errors between the intended marking position and the actual marking position on the 2D screen while using the traditional 2D interface. When the marking difficulty was low, the surfaces of the marking positions were already parallel to the 2D screen or could be easily rotated parallel to the screen plane, and thus, a small displacement error on the 2D screen resulted in the same amount of displacement error of the user's marker on the 3D model. When the marking difficulty was high, the marking locations were easily visually blocked by other parts of the model and visible only from certain visual angles. The surfaces of the marking positions were difficult to rotate parallel to the screen plane while keeping the marking locations visible to the user. Thus, due to the angle that resulted between the marking surface and the screen plane, a small displacement error of the mouse pointer on the 2D screen would lead to a large displacement error of the user's marker on the 3D model, which negatively affected the marking accuracy. As expected, in comparison with the vibrotactile VR interface, using the traditional 2D interface resulted in an improved accuracy when the marking difficulty was low. In contrary, the traditional 2D interface had a lower marking accuracy when the marking difficulty was high (see Fig. 6).

Overall, the results suggested that the task performance of the traditional 2D interface was more affected by the marking positions in terms of marking accuracy. The results also indicated that, compared with the traditional 2D interface, using natural hand 3D gestures with haptic feedback to touch the model for medical marking could achieve more stable task performance regarding to the marking positions. Thus, it supported the potential of the two haptic VR interfaces, specifically the kinesthetic VR interface, for accurate medical marking.

6.2.3. *User experience*

Based on the subjective data (see Fig. 7), the two haptic VR interfaces were perceived to be similarly cognitively demanding and more natural and immersive compared with the traditional 2D interface. The VR interfaces adopted hand gestures to manipulate the 3D objects with haptic feedback and presented head movement-based viewing angles. These interaction experiences are similar to everyday interactions in the physical world, which might have made the participants feel easy to use and contributed to the improved perceived naturalness and immersiveness.

However, an advantage of the traditional 2D interface is that it caused the least hand fatigue. The mouse is a very familiar device to computer users, and users can easily rest their hands on the table while using the mouse. Using the vibrotactile VR interface led to the highest level of perceived hand fatigue, followed by the kinesthetic VR interface. In practical applications, extra arm-rest equipment may be useful to relieve the fatigue induced by the prolonged operation of the VR controller.

6.3. *Limitations and future studies*

This study has a few limitations. First, the experiment of the study involved medical marking on 3D models. For other scenarios in medical diagnosis and planning, such as highlighting a large area or manipulating multiple models, haptic VR interfaces may have different task performances. Further studies can employ the VR interfaces in different medical diagnosis and planning scenarios and explore their usability.

Second, we did not test the interfaces with medical professionals. It is possible that the results would have been different with a participant group who have previous experience in medical marking task using current medical marking systems with a mouse and a 2D display. They may have developed different strategies for interacting with complex 3D models following their years of usage experience. Thus, it is likely that the conventional 2D interface may have performed better if we had evaluated the system with medical professionals. However, for the sake of experimental design, it was a conscious choice to select participants for whom the experimental task and the interaction methods were novel. Future work can evaluate the effect of medical background on the results.

Third, for each medical marking trial, our participants performed a series of sub-tasks. For example, while searching and marking the target positions, the participants performed rotating, panning and zooming operations to manipulate the model. It is possible that the different interfaces we studied performed differently for each of the sub-tasks involved in terms of interaction speed. However, for the medical marking task, the participants often rotated, panned and zoomed the model simultaneously while using the controller and the force-feedback device. Future studies can investigate the performances of the VR interfaces in the sub-task level.

Fourth, this study included a short-term evaluation and the experiment task was new to the participants. It is likely that the task performance and user experience may change after more practice while using the three user interfaces, especially the haptic VR interfaces. The users are familiar with the traditional 2D interface but have only used the haptic VR interfaces for a short time. More training may further improve user performance while using the two VR interfaces. Thus, a long-term study can be conducted in the future to examine the learning effects of the use of the VR interfaces and the 2D interface.

Fifth, the two input manipulating tools used in the experiment for the haptic VR interfaces supported only single-point haptic interactions. Our everyday haptic interactions in the physical world commonly are multiple-point interactions. Multiple-point interaction method may largely improve haptic VR interfaces for manipulating 3D objects. To explore this, sophisticated and reliable haptic devices which support multi-point tactile and kinesthetic interaction are needed. This will require development of haptic technologies before such an experiment is feasible.

At last, we compared two haptic VR interfaces with a baseline of the traditional 2D interface. Such a comparison was necessary since, for any new interface to be practically useful and potentially be adopted by the medical practitioners, it is important for it to exceed the performance of the currently established methods. In addition, this comparative study contributes to the understanding of the strengths and weaknesses of the two VR interfaces. The VR headset with the head movement-based viewing perspective has contributed to the experimental task but was not fully explored in this study. To further explore the potential of the VR headset, a different comparative study is needed, which requires to have the type of displays as an independent variable and employ the same manipulation tool for the experimental interfaces. The mouse may not be an appropriate device for the VR environment and the VR controller is also specially designed for VR and not for a 2D display. One possible experimental setup to fully explore the effects of the VR headset in medical marking would be to use the force-feedback device as the manipulation tool with a 2D display and a VR headset respectively. This comparative study is different from our study but can be a follow-up study to continue from the basis of the present work. We hope our promising results will encourage more future studies in this research area.

7. Conclusion

Three-dimensional visualization has been widely used in computer-aided medical services such as highly accurate diagnosis and planning. The user interfaces for interacting with the 3D models are still largely based on the traditional 2D interaction method with a mouse and a 2D display. In this study, we proposed haptic VR interfaces to manipulate 3D models for the medical diagnosis and planning tasks and implemented a prototype system including two types of haptic VR interfaces (kinesthetic and vibrotactile VR interfaces). We conducted an experiment to compare the two VR interfaces with the traditional 2D interface for medical marking. The haptic VR interfaces showed promise in terms of interaction speed and accuracy. When the tasks involved complex 3D manipulation, using the vibrotactile VR interface led to the shortest task completion time and using the kinesthetic VR interface resulted with the best marking accuracy. The results demonstrated the potential of haptic VR interfaces to interact with volumetric medical images for medical diagnosis and planning. Our future work will investigate their usability in other medical diagnosis and planning scenarios and address how these haptic VR interfaces can be integrated as a part of medical workflow.

CRediT authorship contribution statement

Zhenxing Li: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Maria Kiiveri:** Methodology, Software, Investigation, Writing - review & editing. **Jussi Rantala:** Validation, Writing - review & editing. **Roope Raisamo:** Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

There is no conflict of interest regarding the publication of this article.

References

3D systems 2020 3D systems, 2020. Geomagic haptic device. https://www.3dsystems.com/ (accessed 15 September 2020).

Adesante, 2020. VR solution for Surgeons. https://www.surgeryvision.com/ (accessed 15 September 2020).

Alaraj, A., Luciano, C.J., Bailey, D.P., Elsenousi, A., Roitberg, B.Z., Bernardo, A., Banerjee, P.P., Charbel, F.T., 2015. Virtual Reality Cerebral Aneurysm Clipping Simulation with Real-Time Haptic Feedback. Operative Neurosurgery 11 (1), 52–58. https://doi.org/10.1227/NEU.0000000000000583.

Al-Khalifah, A., McCrindle, R., Sharkey, P.M., Alexandrov, V.N., 2006. Using virtual reality for medical diagnosis, training and education. Int. J. Dis. Human Dev. 5 (2), 187–193. https://doi.org/10.1515/IJDHD.2006.5.2.187.

Bade, R., Ritter, F., Preim, B., 2005. Usability Comparison of Mouse-Based Interaction Techniques for Predictable 3d Rotation. In: International Symposium on Smart Graphics. SG 2005, pp. 138–150. https://doi.org/10.1007/11536482_12.

Berthier, N.E., Clifton, R.K., Gullapalli, V., McCall, D.D., Robin, D.J., 1996. Visual information and object size in the control of Reaching. J. Mot. Behav. 28 (3), 187–197. https://doi.org/10.1080/00222895.1996.9941744.

Besançon, L., Issartel, P., Ammi, M., Isenberg, T., 2017. Mouse, Tactile, and Tangible Input for 3D Manipulation. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17, pp. 4727–4740. https://doi.org/10.1145/3025453.3025863.

Bholsithi, W., Tharanon, W., Chintakanon, K., Komolpis, R., Sinthanayothin, C., 2009. 3D vs. 2D cephalometric analysis comparisons with repeated measurements from 20 Thai males and 20 Thai females. Biomed. Imaging Interv. J. 5 (4), e21. https://doi.org/10.2349/biij.5.4.e21.

Bielser, D., Gross, M.H., 2000. Interactive simulation of surgical cuts. In: Proceedings the Eighth Pacific Conference on Computer Graphics and Applications. https://doi.org/10.1109/PCCGA.2000.883933.

Bouchlaghem, D., Shang, H., Whyte, J., Ganah, A., 2005. Visualisation in architecture, engineering and construction (AEC). J. Automation in Construction. 14 (3), 287–295. https://doi.org/10.1016/j.autcon.2004.08.012.

Bowman, D.A., Kruijff, E., LaViola Jr., J.J., Poupyrev, I., 2004. 3D User interfaces: theory and practice, 1st Ed. Addison-Wesley Professional Boston.

Chen, M., Mountford, S.J., Sellen, A., 1988. A study in interactive 3-D rotation using 2-D control devices. ACM SIGGRAPH Computer Graphics 22 (4), 121–129. https://doi.org/10.1145/54852.378497.

El Saddik, A., Orozco, M., Eid, M., Cha, J., 2011. Haptics Technologies: Bringing Touch to Multimedia. In: Avizzano, C.A., Azorín, J.M. (Eds.), Springer Series on Touch and Haptic Systems. Springer-Verlag, Berlin Heidelberg.

Fahmi, F., Nainggolan, F., Siregar, B., 2019. 3D anatomy learning system using Virtual Reality and VR Controller. Journal of Physics: Conference Series 1235, 012043. https://doi.org/10.1088/1742-6596/1235/1/012043.

Fischer, A., Vance, J.M., 2003. PHANToM haptic device implemented in a projection screen virtual environment. In: Proceedings of the Workshop on Virtual Environments. EGVE '03, pp. 225–229. https://doi.org/10.1145/769953.769979.

Foley, J.D., Andries, D., Feiner, S.K., Hughes, J.F., 1995. Computer Graphics: Principles and Practice in C, 2nd Ed. Addison-Wesley Professional, Boston.

Gallo, L., Minutolo, A., De Pietro, G., 2010. A user interface for VR-ready 3D medical imaging by off-the-shelf input devices. Comput. Biol. Med. 40 (3), 350–358. https://doi.org/10.1016/j.compbiomed.2010.01.006.

Gates, D.H., Dingwell, J.B., 2011. The effects of muscle fatigue and movement height on movement stability and variability. Exp. Brain Res. 209 (4), 525–536. https://doi.org/10.1007/s00221-011-2580-8.

Gross, M.H., 1998. Computer graphics in medicine: from visualization to surgery simulation. ACM SIGGRAPH Computer Graphics 32 (1), 53–56. https://doi.org/10.1145/279389.279462.

Harrell, W.E., 2007. Three-dimensional diagnosis & treatment planning: The use of 3D facial imaging & ConeBeam CT (CBCT) in orthodontics & dentistry. Australasian Dental Practice. 102–113. July/August 2007.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances Psychol 52, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hincapié-Ramos, J.D., Guo, X., Moghadasian, P., Irani, P., 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14, pp. 1063–1072. https://doi.org/10.1145/2556288.2557130.

Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell, N., 1997. Usability analysis of 3D rotation techniques. In: Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology. UIST '97, pp. 1–10. https://doi.org/10.1145/263407.263408.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Statist. 6 (2), 65–70. https://www.jstor.org/stable/4615733.

Imagine Media, 1996. The Next Generation 1996 Lexicon A to Z: TrackBall. Next Generation 15, 42. https://archive.org/details/nextgen-issue-015.

Issa, T., Isaias, P., 2015. Usability and Human Computer Interaction (HCI). Sustainable Design. Springer, London. https://doi.org/10.1007/978-1-4471-6753-2_2.

Jankowski, J., Hachet, M., 2013. A survey of interaction techniques for interactive 3d environments. In: Eurographics Conference, EG 2013 - STARs. https://doi.org/10.2312/conf/EG2013/stars/065-093.

Jones, T., Moore, T., Choo, J., 2016. The Impact of Virtual Reality on Chronic Pain. PLoS ONE 11 (12), e0167523. https://doi.org/10.1371/journal.pone.0167523.

Katsumata, A., Fujishita, M., Maeda, M., Ariji, Y., Ariji, E., Langlais, R.P., 2005. 3D-CT evaluation of facial asymmetry. J. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. 99 (2), 212–220. https://doi.org/10.1016/j.tripleo.2004.06.072.

Kim, H., Choi, Y., 2019. Performance Comparison of User Interface Devices for Controlling Mining Software in Virtual Reality Environments. Appl. Sci. 9 (13), 2584. https://doi.org/10.3390/app9132584.

King, F., Jayender, J., Bhagavatula, S.K., Shyn, P.B., Pieper, S., Kapur, T., Lasso, A., Fichtinger, G., 2016. An Immersive Virtual Reality Environment for Diagnostic Imaging. J. Medical Robotics Research. 1 (1), 1640003 https://doi.org/10.1142/S2424905X16400031.

Kinnison, T., Forrest, N.D., Frean, S.P., Baillie, S., 2009. Teaching bovine abdominal anatomy: Use of a haptic simulator. J. Anat. Sci. Educ. 2 (6), 280–285. https://doi.org/10.1002/ase.109.

Kula, K., Ghoneima, A., 2018. Cephalometry in Orthodontics: 2D and 3D, 1st ed. Quintessence Publishing, USA.

Li, Z., Akkil, D., Raisamo, R., 2020. The Impact of Control-Display Gain in Kinesthetic Search. In: Nisky, I., Hartcher-O'Brien, J., Wiertlewski, M., Smeets, J. (Eds.), Haptics: Science, Technology, Applications. EuroHaptics 2020, 12272. Springer, Cham. https://doi.org/10.1007/978-3-030-58147-3_18. Lecture Notes in Computer Science.

Lindner, C., Wang, C.W., Huang, C.T., Li, C.H., Chang, S.W., Cootes, T.F., 2016. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. Scientific Reports 6, 33581. https://doi.org/10.1038/srep33581.

Maples-Keller, J.L., Yasinski, C., Manjin, N., Rothbaum, B.O., 2017. Virtual reality-enhanced extinction of phobias and post-traumatic stress. Neurotherapeutics 14 (3), 554–563. https://doi.org/10.1007/s13311-017-0534-y.

Massie, T.H., Salisbury, J.K., 1994. The PHANToM haptic interface: a device for probing virtual objects. In: Proceedings of the ASME International Mechanical Engineering Congress and Exhibition, Dynamic Systems and Control, 55, pp. 295–301.

Medellín-Castillo, H.I., Govea-Valladares, E.H., Pérez-Guerrero, C.N., Gil-Valladares, J., Lim, T., Ritchie, J.M., 2016. The evaluation of a novel haptic-enabled virtual reality approach for computer-aided cephalometry. Comput. Methods Programs Biomed. 130, 46–53. https://doi.org/10.1016/j.cmpb.2016.03.014.

Novint, 2020. Novint Falcon haptic device. https://www.vrs.org.uk/virtual-reality-gear/haptic/novint-falcon.html (accessed 15 September 2020).

Oculus, 2020. Oculus Rift VR. https://www.oculus.com/ (accessed 15 September 2020).

Okamura, A.M., 2004. Methods for haptic feedback in teleoperated robot-assisted surgery. Int. J. Industrial Robot. 31 (6), 499–508. https://doi.org/10.1108/01439910410566362.

Olszewski, R., Zech, F., Cosnard, G., Nicolas, V., Macq, B., Reychler, H., 2007. Three-dimensional computed tomography cephalometric craniofacial analysis, experimental validation in vitro. Int. J. Oral Maxillofac. Surg. 36 (9), 828–833. https://doi.org/10.1016/j.ijom.2007.05.022.

Peddamatham, S., Peine, W., Tan, H.Z., 2008. Assessment of Vibrotactile Feedback in a Needle-Insertion Task using a Surgical Robot. In: 2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. https://doi.org/10.1109/haptics.2008.4479920.

Poupyrev, I., Ichikawa, T., Weghorst, S., Billinghurst, M., 1998. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. Computer Graphics Forum 17 (3), 41–52. https://doi.org/10.1111/1467-8659.00252.

Preim, B., Bartz, D., 2007. Visualization in Medicine: Theory, Algorithms, and Applications, 1st ed. Morgan Kaufmann, Burlington, USA.

Randall, D., Metherall, P., Bardhan, K.D., Spencer, P., Gillott, R., Noronha, R., Fenner, J. W., 2016. The Oculus Rift virtual colonoscopy: introducing a new technology and initial impressions. J. Biomedical Graphics and Computing. 6 (1), 34–42. https://doi.org/10.5430/jbgc.v6n1p34.

Reinschluessel, A.V., Muender, T., Uslar, V.N., Weyhe, D., Schenk, A., Malaka, R., 2019. Tangible Organs: Introducing 3D Printed Organ Models with VR to Interact with Medical 3D Models. In: Proceedings of CHI Conference on Human Factors in Computing Systems Extended Abstracts. CHI'19 Extended Abstracts. https://doi.org/10.1145/3290607.3313029.

Ribeiro, M.L., Lederman, H.M., Elias, S., Nunes, F.L.S., 2016. Techniques and Devices Used in Palpation Simulation with Haptic Feedback. J. ACM Computing Surveys. 49 (3) https://doi.org/10.1145/2962723. Article No. 48.

Saad, E., Funnell, W.R.J., Kry, P.G., Ventura, N.M., 2018. A Virtual-Reality System for Interacting with Three-Dimensional Models Using a Haptic Device and a Head-Mounted Display. In: 2018 IEEE Life Sciences Conference, pp. 191–194. https://doi.org/10.1109/LSC.2018.8572120.

Satava, R.M., Robb, R.A., 1997. Virtual Endoscopy: Application of 3D Visualization to Medical Diagnosis. PRESENCE: Teleoperators and Virtual Environments 6 (2), 179–197. https://doi.org/10.1162/pres.1997.6.2.179.

Schoonmaker, R.E., Cao, C.GL., 2006. Vibrotactile force feedback system for minimally invasive surgical procedures. In: IEEE International Conference on Systems, Man and Cybernetics, 3, pp. 2464–2469. https://doi.org/10.1109/ICSMC.2006.385233.

Shen, R., Boulanger, P., Noga, M., 2008. MedVis: A Real-Time Immersive Visualization Environment for the Exploration of Medical Volumetric Data. In: 2008 Fifth International Conference BioMedical Visualization: Information Visualization in Medical and Biomedical Informatics. https://doi.org/10.1109/MediVis.2008.10.

Shing, C.Y., Fung, C.P., Chuang, T.Y., Penn, I.W., Doong, J.L., 2003. The study of auditory and haptic signals in a virtual reality-based hand rehabilitation system. J. Robotica. 21 (2), 211–218. https://doi.org/10.1017/S0263574702004708.

Shoemake, K., 1992. ARCBALL: A user interface for specifying three-dimensional orientation using a mouse. In: Proceedings of Graphics Interface '92, pp. 151–156. https://doi.org/10.20380/GI1992.18.

Silén, C., Wirell, S., Kvist, J., Nylander, E., Smedby, O., 2008. Advanced 3D visualization in student-centred medical education. J. Med. Teach. 30 (5), e115–e124. https://doi.org/10.1080/01421590801932228.

Sousa, M., Mendes, D., Paulo, S.F., Matela, N., Jorge, J.A.P., Lopes, D.S., 2017. VRRRRoom: Virtual Reality for Radiologists in the Reading Room. In: 35th Annual ACM Conference on Human Factors in Computing Systems. CHI '17, pp. 4057–4062. https://doi.org/10.1145/3025453.3025566.

Steam, 2020. SteamVR. https://store.steampowered.com/app/250820/SteamVR/ (accessed 15 September 2020).

Steinberg, A.D., Bashook, P.G., Drummund, J., Ashrafi, S., Zefran, M., 2007. Assessment of faculty perception of content validity of PerioSim, a haptic–3D virtual reality dental training simulator. J. Dent. Educ. 71 (12), 1574–1582. https://doi.org/10.1002/j.0022-0337.2007.71.12.tb04434.x.

Surgical Theater, 2020. VR surgical planning and navigation. https://www.surgicaltheater.net/ (accessed 15 September 2020).

Sutton, D., 1993. Textbook of Radiology and Imaging, 5th ed. Churchill Livingstone, London.

Swennen, G.R.J., Schutyser, F., 2007. Three-dimensional virtual approach to diagnosis and treatment planning of maxillo-facial deformity. In: Bell, W.H., Guerrero, C.A. (Eds.), Distraction Osteogenesis of the Facial Skeleton, 1st ed. BC Decker Inc., Hamilton, p. 24.

Tietjen, C., Meyer, B., Schlechtweg, S., Preim, B., Hertel, I., Strauß, G., 2006. Enhancing Slice-based Visualizations of Medical Volume Data. In: Conference EuroVis06: Joint Eurographics - IEEE VGTC Symposium on Visualization, pp. 123–130. https://doi.org/10.2312/VisSym/EuroVis06/123-130.

Troulis, M.J., Everett, P., Seldin, E.B., Kikinis, R., Kaban, L.B., 2002. Development of a three-dimensional treatment planning system based on computed tomographic data. Int. J. Oral. Maxillofac. Surg. 31 (4), 349–357. https://doi.org/10.1054/ijom.2002.0278.

Unity, 2020. Unity software. https://www.unity.com/ (accessed 15 September 2020).

Unity haptic plugin, 2020. Unity 5 Haptic Plugin for Geomagic OpenHaptics 3.3 (HLAPI/HDAPI). https://assetstore.unity.com/packages/essentials/tutorial-projects/unity-5-haptic-plugin-for-geomagic-openhaptics-3-3-hlapi-hdapi-34393 (accessed 15 September 2020).

Van Vlijmen, O.J., Maal, T., Bergé, S.J., Bronkhorst, E.M., Katsaros, C., Kuijpers-Jagtman, A.M., 2010. A comparison between 2D and 3D cephalometry on CBCT scans of human skulls. Int. J. Oral Maxillofac. Surg. 39 (2), 156–160. https://doi.org/10.1016/j.ijom.2009.11.017.

Venson, J.E., Albiero Berni, J.C., Edmilson da Silva Maia, C., Marques da Silva, A.M., Cordeiro d'Ornellas, M., Maciel, A., 2017. A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality. Studies in Health Technology and Informatics. In: MEDINFO 2017: Precision Healthcare through Informatics, 245, pp. 244–248. https://doi.org/10.3233/978-1-61499-830-3-244.

VIVE, 2020. HTC VIVE VR. https://www.vive.com/ (accessed 15 September 2020).

Webster, R.W., Sassani, J., Shenk, R., Good, N., 2004. A haptic surgical simulator for the continuous curvilinear capsulorhexis procedure during cataract surgery. Studies in Health Technology and Informatics. In: Medicine Meets Virtual Reality 12, 98, pp. 404–406. https://doi.org/10.3233/978-1-60750-942-4-404.

Wirth, M., Gradl, S., Sembdner, J., Kuhrt, S., Eskofier, B.M., 2018. Evaluation of Interaction Techniques for a Virtual Reality Reading Room in Diagnostic Radiology. In: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. UIST '18, pp. 867–876. https://doi.org/10.1145/3242587.3242636.

Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J., 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11, pp. 143–146. https://doi.org/10.1145/1978942.1978963.

Yu, L., Svetachov, P., Isenberg, P., Everts, M.H., Isenberg, T., 2010. FI3D: Direct-Touch Interaction for the Exploration of 3D Scientific Visualization Spaces. IEEE Transactions on Visualization and Computer Graphics 16 (6), 1613–1622. https://doi.org/10.1109/TVCG.2010.157.

Ángyán, L., Téczely, T., Ángyán, Z., 2007. Factors affecting postural stability of healthy young adults. Acta. Physiol. Hung. 94 (4), 289–299. https://doi.org/10.1556/APhysiol.94.2007.4.1.

Zhenxing Li received his MSc degree in Radio Frequency Electronics from Tampere University of Technology and MSc degree in User Interface Software Development from University of Tampere, in 2009 and 2013 respectively. He is a researcher in Tampere Unit for Computer-Human Interaction (TAUCHI) Research Center at Tampere University and working toward PhD degree in Interactive Technology. His research interests include haptics, VR, gaze, multimodal interaction, mathematical modelling and algorithms.

Maria Kiiveri received her MSc degree in Computer Science from University of Tampere in 2019. She has worked in Tampere Unit for Computer-Human Interaction (TAUCHI) Research Center in Tampere University in a project that develops innovative technical solutions for medical and surgical use. She is currently employed in a company that provides software solutions for clinical laboratories.

Jussi Rantala received his MSc in Computer Science (2007) and PhD in Interactive Technology (2014) from the University of Tampere. He is a Senior Research Fellow at the Tampere Unit of Computer-Human Interaction (TAUCHI) at Tampere University. His research focuses on human-computer interaction, with special interest in haptics, olfaction, multisensory interfaces and XR.

Roope Raisamo is a professor of computer science in Tampere University, Faculty of Information Technology and Communication Sciences. He received his PhD degree in computer science from the University of Tampere in 1999. He is the head of TAUCHI Research Center (Tampere Unit for Computer-Human Interaction), leading Multimodal Interaction Research Group. His 25-year research experience in the field of human-technology interaction is focused on multimodal interaction, XR, haptics, gaze, gestures, interaction techniques, and software architectures.

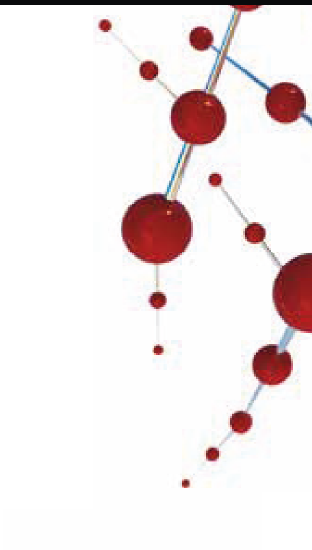
Publications in the **Dissertations in Interactive Technology** series

---

Details of the dissertations are available at
https://research.tuni.fi/tauchi/dissertations/

1. **Timo Partala:** Affective Information in Human-Computer Interaction
2. **Mika Käki:** Enhancing Web Search Result Access with Automatic Categorization
3. **Anne Aula:** Studying User Strategies and Characteristics for Developing Web Search Interfaces
4. **Aulikki Hyrskykari:** Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid
5. **Johanna Höysniemi:** Design and Evaluation of Physically Interactive Games
6. **Jaakko Hakulinen:** Software Tutoring in Speech User Interfaces
7. **Harri Siirtola:** Interactive Visualization of Multidimensional Data
8. **Erno Mäkinen:** Face Analysis Techniques for Human-Computer Interaction
9. **Oleg Špakov:** iComponent – Device-Independent Platform for Analyzing Eye Movement Data and Developing Eye-Based Applications
10. **Yulia Gizatdinova:** Automatic Detection of Face and Facial Features from Images of Neutral and Expressive Faces
11. **Päivi Majaranta:** Text Entry by Eye Gaze
12. **Ying Liu:** Chinese Text Entry with Mobile Phones
13. **Toni Vanhala:** Towards Computer-Assisted Regulation of Emotions
14. **Tomi Heimonen:** Design and Evaluation of User Interfaces for Mobile Web Search
15. **Mirja Ilves:** Human Responses to Machine-Generated Speech with Emotional Content
16. **Outi Tuisku:** Face Interface
17. **Juha Leino:** User Factors in Recommender Systems: Case Studies in e-Commerce, News Recommending, and e-Learning
18. **Joel S. Mtebe:** Acceptance and Use of eLearning Solutions in Higher Education in East Africa
19. **Jussi Rantala:** Spatial Touch in Presenting Information with Mobile Devices
20. **Katri Salminen:** Emotional Responses to Friction-based, Vibrotactile, and Thermal Stimuli
21. **Selina Sharmin:** Eye Movements in Reading of Dynamic On-screen Text in Various Presentation Formats and Contexts
22. **Tuuli Keskinen:** Evaluating the User Experience of Interactive Systems in Challenging Circumstances
23. **Adewunmi Obafemi Ogunbase**: Pedagogical Design and Pedagogical Usability of Web-Based Learning Environments: Comparative Cultural Implications from Africa and Europe
24. **Hannu Korhonen**: Evaluating Playability of Mobile Games with the Expert Review Method
25. **Jani Lylykangas**: Regulating Human Behavior with Vibrotactile Stimulation
26. **Ahmed Farooq**: Developing Technologies to provide haptic feedback for surface based interaction in mobile devices
27. **Ville Mäkelä**: Design, Deployment, and Evaluation of Gesture-Controlled Displays in Ubiquitous Environments
28. **Pekka Kallioniemi**: Collaborative Wayfinding in Virtual Environments

29. **Sumita Sharma**: Collaborative Educational Applications for Underserved Children: Experiences from India
30. **Tomi Nukarinen**: Assisting Navigation and Object Selection with Vibrotactile Cues
31. **Deepak Akkil**: Gaze Awareness in Computer-Mediated Collaborative Physical Tasks
32. **Toni Pakkanen**: Supporting Eyes-Free User Interaction with Vibrotactile Haptification of User Interfaces
33. **Antti Sand:** On Adding Haptic Feedback to Interaction with Unconventional Display Devices
34. **Jukka Selin:** Tietomallin pelillistäminen ja toiminnallisen suunnittelun menetelmä rakennusten suunnittelun apuna
35. **Zhenxing Li:** Efficient and Accurate Hand-based Kinesthetic Interaction for Virtual Reality



Current virtual reality (VR) technology is mainly based on visual and auditive modalities. Implementing kinesthetic interaction in VR could largely improve the immersion of VR and extend the range of its applications. Grounded force-feedback devices can be used to implement kinesthetic interaction. These devices provide a reliable kinesthetic interface with robust and realistic force feedback. However, there is a major challenge while using force-feedback devices for VR interaction, that is, the small workspace. The current solution to this challenge is to employ a large control-display (CD) gain.

The aim of this dissertation is to enable efficient and accurate hand-based kinesthetic interaction for VR. The research was divided into three steps: problem understanding, development and application. First, the research examined the effects of CD gain on kinesthetic interaction and thus provided an empirical basis for designing new kinesthetic interfaces. Second, to address the issue of limited workspace, three new kinesthetic interfaces were developed by using the user's gaze as an input modality. The research also identified the critical factors for designing high-quality kinesthetic interfaces in terms of kinesthetic perception accuracy. Third, the research explored medical applications of a kinesthetic VR interface and demonstrated its potential to be the next-generation user interface in the field of medical diagnosis and planning.

Dissertations in Interactive Technology
Number 35

ISBN 978-952-03-2051-5
ISSN 1795-9489