STEREOSCOPIC BIMANUAL INTERACTION FOR 3D VISUALIZATION

by

Isaac Cho

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
In partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Computing and Information Systems

Charlotte

2013

Approved by:

____
Dr. Zachary Wartell

____
Dr. William Ribarsky

____
Dr. Xiaoyu Wang

____
Dr. Paula Goolkasian

____
Dr. Mark Faust

ABSTRACT

ISAAC CHO. Stereoscopic bimanual interaction for interactive 3D visualization (Under the direction of DR. ZACHARY J. WARTELL)

Virtual Environments (VE) are being widely used in various research fields for several decades such as 3D visualization, education, training and games. VEs have the potential to enhance the visualization and act as a general medium for human-computer interaction (HCI). However, limited research has evaluated virtual reality (VR) display technologies, monocular and binocular depth cues, for human depth perception of volumetric (non-polygonal) datasets. In addition, a lack of standardization of three-dimensional (3D) user interfaces (UI) makes it challenging to interact with many VE systems.

To address these issues, this dissertation focuses on evaluation of effects of stereoscopic and head-coupled displays on depth judgment of volumetric dataset. It also focuses on evaluation of a two-handed view manipulation techniques which support simultaneous 7 degree-of-freedom (DOF) navigation (x,y,z + yaw,pitch,roll + scale) in a multi-scale virtual environment (MSVE). Furthermore, this dissertation evaluates auto-adjustment of stereo view parameters techniques for stereoscopic fusion problems in a MSVE. Next, this dissertation presents a bimanual, hybrid user interface which combines traditional tracking devices with computer-vision based “natural” 3D inputs for multi-dimensional visualization in a semi-immersive desktop VR system. In conclusion, this dissertation provides a guideline for research design for evaluating UI and interaction techniques.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my wife, Woneui, for her encouragement, patience and unwavering love and support. I could not have finished my dissertation without her support.

I would like to thank Dr. Zachary Wartell, my thesis advisor, for his tolerance and support. This research would not be possible without his support and help. I am sincerely grateful for his efforts to keep me on the right track to be a researcher. He was the best advisor I could have asked for all my research interests and problems as well as personal concerns.

I would like to thank my other committee members: Dr. William Ribarsky, Dr. Derek Xiaoyu Wang, Dr. Paula Goolkasian and Dr. Mark Faust for all their help and guideline. In particular, I would like to express my sincere gratitude to Dr. Paula Goolkasian and Dr. Mark, who always gave me invaluable feedback of statistical analysis for my research.

I would like to thank Dr. Chang Geun Song and Dr. Ha Jin Kimn who encouraged and motivated me to start my Ph.D. course. My life would be totally different if I had not met them.

I would like to thank all present and past faculty members of the Computer Science Department and members of the Charlotte Visualization Center at the University of North Carolina at Charlotte. I would also like to thank for volunteers who participated my research. I would also like to thank all my friends who I met in Charlotte,

I would like to thank God and my family, My Mom, Kwang Ok Kim and my Dad, Young Hee Cho, for their sacrificial support and love. I am sincerely proud of being your son. I am also truly grateful to my sisters, Hae Ryung and Hae Jin, and my family in law

for their support and help.

Last but not the least, I would like to thank my friends and colleagues in South Korea, especially Dong Pyo Hong, Young Jun Kim, Sean Kang, Jung Hun Kwon, Eun Joo Kwak, Myung Jin Song, Tae San Park, Se Kwang Oh, Jong Woo Jung, Ji Young Jeong, Mun Hyun Lee and Dong Kyu Kim for always being by my side since we were teenagers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	analysis of variance
AR	augmented reality
CAVE	cave automatic virtual environment
DOF	degree of freedom
FOR	field of regard
FOV	field of view
GUI	graphical user interface
HCI	human computer interaction
HMD	head-mounted display
HTD	head-tracked display
HUI	hybrid user interface
IT	interaction technique
LSD	least significant difference
MSVE	multi-scale virtual environment
MUI	multimodal user interaction
UI	user interface
VE	virtual environment
VR	virtual reality
WIMP	window-icons-menu-pointer

CHAPTER 1: INTRODUCTION

Virtual Environments (VEs) are being widely used in various research fields, such as 3D visualizations [28,29], educations [74,174,168,60], trainings [42,135] and games [179,50,99]. Stereoscopic VEs have a lot of potential to visualize various types of real world datasets and to be a general medium for human-computer interaction (HCI) [44,45,102]. Over the last decade, many researchers have employed complex real world datasets including scientific and medical 3D volumetric datasets as well as conventional polygonal datasets. Furthermore, many researchers have developed a number of 3D user interfaces (UIs) to enrich interaction between the user and the computer generated 3D world.

VEs can be categorized as Head-Mounted Displays (HMD), which mount the display on a heads-worn apparatus, and Head-Tracked Displays (HTD) [139], which uses stationary displays that are attached to a wall, tabletop or desk. Examples of HTD VEs are the Cave Automatic Virtual Environment (CAVE) [37], fish tank VR [154], the virtual workbench [165] and desktop VR [121]. This dissertation deals with HTD systems, mainly semi-immersive desktop VR. A primary component of HTD systems is a stereoscopic and head-coupled display which can enhance human depth perception of computer generated 3D imagery by providing binocular (retina disparity) and monocular (motion parallax) cues [155,160].

1.1 Immersive Virtual Reality technology

Much previous research has demonstrated the utility of stereoscopy and head-coupled motion parallax for enhancing human perception of complex three-dimensional (3D) datasets. For example, studies by Ware et al. examine the effect of stereoscopic display and kinetic depth on understanding 3D networks which are represented by tubes or lines [155,161,160]. User performance at finding paths in a complex 3D network improves when using stereoscopy and structure-from-motion.

A significant number of visual analytic domains, however, also heavily use 3D volumetric data. Examples include medical imaging, weather and environmental simulations, and fluid flow. Volumetric data is characterized by large amounts of transparency, occlusion and ambiguous spatial structure. There has been a fair amount of evaluation of perception of volumetric data under different rendering algorithms and different parameterizations such as modifying transfer functions [77]. For a simple volume dataset one might find a combination of transfer function settings, rendering algorithms and rendering augmentations (such as edge enhancement) such that adding VR type displays would do little to further improve depth and shape perception. However, for more complex volumetric dataset, even with perceptually optimized transfer functions and rendering parameters, VR display capabilities may further improve depth and shape perceptions.

For surface and 3D networks, this type of result is typical. When the 3D geometry is relatively simple, well-chosen rendering parameters can maximize shape perception causing VR display technologies to have less additional benefit. However, with more realistic complex geometry, VR technologies show significant positive effects on shape and depth perception.

One would expect similar results for volumetric data. Further, the addition of VR technologies could be especially important with time-varying volumetric dataset that are viewed in real-time where extensive preprocessing for optimized transfer functions and volume rendering parameters is not possible. An example would be real-time, streaming Doppler weather radar data [164]. With the increasing affordability of semi-immersive VR displays and GPUs capable of advanced volume rendering, there is a need to quantify the effectiveness of stereopsis and structure-from-motion on volumetric data, and also to quantify how these display parameters interact with other volumetric rendering conditions.

Although stereoscopic head-coupled display can enhance depth perception of the user in a computer generated 3D world, sometimes the user sees two separate 2D images rather than a solid 3D image. Many factors influence fusion problems, but typically these translate into a range of distance in front of and behind the screen where a stereo 3D image can be comfortably fused. Fusion problems increase eye strain and motion sickness especially with a head-tracking display. Ware et al. [151] reduce these problems by changing the modeled eye separation value, such as using 3cm instead of 6cm, to reduce stereo 3D image depth. However, with the head-tracked stereo, this false eye separation causes a distortion with a head-tracked shearing component [163]. Changing the view scale factor via cyclopean scale is an alternative way to change image depth without a inducing a dynamic, non-linear distortion.

1.2 3D User Interface and Interaction

The ubiquitous Windows-Icon-Menu-Pointer (WIMP) UI and its 2D mouse UI techniques began with Xerox Parc’s and others seminal work. Similar to 2D interaction techniques (ITs) [22], 3D ITs often require physical devices, such as ChordGloves [92] (also pinch gloves), a bat [157] , or a Cubic Mouse [52], to provide full 6 Degree of

Freedom (DOF) interaction to the user. Although a significant number of researchers have produced 3D UIs, there still is no standard 3D UI while various WIMP standards exist. Hence, many researchers still employ the conventional WIMP interfaces for their 3D applications due to its familiarity even though it does not provide a complete solution for 3D ITs because of its limited DOFs (i.e. 2DOF) [22].

The complexity and size of 3D datasets targeted by 3D visualization applications continues to grow rapidly. For example, a real-time scientific visualization generates complex geometry in massive datasets [28]. This growths data complication grows the complexity of interaction of the application due to heavy mathematical computations for the communication between the user and complicated data. Further complication have arisen because a significant number of researchers have begun to use 3D volumetric datasets to meet the requirements established by a number of application areas. The 3D volumetric dataset incorporates a 3D discrete regular grid of voxels and commonly is stored in a volume buffer, which is a large 3D array of voxels [73]. Examples of applications that employ 3D volumetric datasets include medical imaging, weather, environmental simulation, fluid flow, and geospatial visualization (such as deep-ocean sonar and ocean current applications). Even though 3D applications have increased in their interaction's complexity, the traditional WIMP UI remains the most common UI for 3D volumetric visualizations. 3D volumetric visualizations have an additional complexity because raw volumetric data does not provide edge and surface information for selection. Also data size can be relatively larger than for non-volumetric datasets. As such, 3D ITs for VEs designed for surface (typically polygonal) datasets are not always appropriate for 3D volumetric data [144]. Addressing these challenges requires evaluation of 3D UIs for

3D volumetric datasets and, potentially, the development of new interaction methods.

To address these challenges with traditional 3D UIs, researchers introduced various UIs and ITs. One example is the hybrid user interface (HUI). Feiner et al. [49] first introduced the HUI for their augmented reality (AR) system, which uses a see-through HMD unit with a 2D desktop screen for multi-computer interactions. Conventionally, a HUI consists of heterogeneous displays and input devices that use multimodal physical and natural human input devices. However, previous works generally focused on output and simultaneous multiple heterogeneous displays. These previous studies on HUIs rarely endorsed hybrid interaction. Rather, they typically encourage developers to use different UI techniques individually to control varying heterogeneous displays [10,39].

Similar to the concept of the HUI, Multimodal User Interaction (MUI) provides combined natural human input modes, such as voice, gaze, and body gesture with traditional UIs [107,108,109]. No references currently exist that clearly detail the differences between HUIs and MUIs. According to previous works, however, MUI tends to focus on natural human input modes, particularly ‘speech’ (or voice), the key component of MUI. Several studies use physical input devices (e.g. stylus input) for MUI. Conversely, hybrid user interface research tends to focus on both non-traditional input and non-traditional output devices. The HUI provides high usability, flexibility, and the blended benefits of different UIs and ITs for HCI by combining multimodal input devices.

Using immersive VR technologies can also pose challenges for 3D UIs. For example, some multi-scale virtual environments (MSVEs), which are VEs that contain geometric details whose sizes cover several orders of magnitude (i.e. a spatial multi-scale world [53] that is 3D), require manipulating of view scale as a separate 7th DOF in the view model

[122,123]. This 7th DOF, view scale, is not equivalent to Field-of-View (FOV), travel gain factors, or a host of other possible view model parameters. View scale has significant effects on usability in systems with head-coupled display [112], stereoscopic display [156] or direct 3D manipulation [97]. Many 7DOF object manipulation techniques exist including Mapes and Moshell's work [92]. Many 3D ITs can be used for either object manipulation or travel using the grab-the-world approach [157]. However, none of previous works [127,133,176] provide a simultaneous 7DOF (position + orientation + scale) object manipulation technique.

1.3 Summary of Research Challenges

In summary, this dissertation focuses on several research challenges:

1) Stereoscopic and head-coupled displays have not been fully evaluated for 3D volumetric dataset to determine their effects for human depth perception.
2) There is yet no simultaneous bimanual 7DOF IT presented and evaluated, especially for a MSVE.
3) Dynamic adjustment of stereo view parameter techniques, which are to reduce stereoscopic fusion problems, need to be evaluated to determine their effectiveness in head-tracked stereoscopic VE systems.
4) HUI that supports 2D and 3D physical UIs with natural hands input without an acquisition time penalty for multi-dimensional visualization in semi-immersive VR has not been presented.

1.4 Contributions

The purpose of the research presented in this dissertation is (1) to determine benefits of VR technologies for depth perception in 3D volumetric datasets, (2) to develop and evaluate bimanual simultaneous 7DOF ITs for navigation in a MSVE, (3) to evaluate dynamic adjustment of stereo view parameter techniques to reduce stereoscopic fusion problems and (4) to develop and evaluate a HUI that is a two-handed, hybrid of 3D and 2D user inputs with human natural hand and finger input for multi-dimensional visualization in desktop VR.

1.5 Thesis Organization

Chapter 2 broadly reviews previous research and background information concerning VR, 3D UIs and various application domains in a VE.

Chapter 3 presents user experiments to evaluate benefits of fundamental VR technologies for depth perception in 3D volumetric datasets. Two user studies on depth ordering and depth discrimination were conducted to determine stereo and motion effects for 3D volumetric datasets in semi-immersive desktop VR. In addition, the chapter compares the results with previous human depth perception studies on other types of data.

Chapter 4 introduces a simultaneous 7 DOF object manipulation technique and presents user studies to show its advantages for view manipulation tasks compared to previous one-handed and two-handed object manipulation techniques. The chapter also describes an implementation of this technique using a natural hand input and discusses its limitations.

Chapter 5 evaluates several dynamic adjustments of the stereo view in a MSVE to reduce stereoscopic problems. The chapter shows different effects of auto-adjustment techniques on the user's ability to accomplish a task between desktop VR and a CAVE with a one-handed, scene-in-hand 7 DOF travel technique.

Chapter 6 introduces a bimanual HUI for multi-dimensional visualization in a semi-immersive, desktop VR system. The chapter introduces a cross-dimensional application and concept of a two-handed HUI that simultaneously uses both natural hand inputs and traditional physical 3D input devices. This chapter also presents an experimental evaluation of the HUI by comparing it with two other traditional UIs.

Chapter 7 discusses design guidelines that have resulted from this research and Chapter 8 summarizes this dissertation and discusses future works.

CHAPTER 2: BACKGROUND AND LITERATURE REVIEW

2.1 Depth Perception

Prior experiments with surface and 3D network datasets show that stereoscopic display can aid depth perception when either the visual stimuli lacks other depth cues, as can occur in teleoperator environments and remote sensing or in less sophisticated computer graphic presentations, or when the visual stimuli contains a high depth complexity as measured by many occlusions [38]. As a specific recent example, Ware et al. examine the effect of stereoscopy and the kinetic depth effect on a person's understanding of a 3D network which is represented by 3D tubes or lines [155,161,160]. They demonstrate great benefit for stereoscopic and kinetic depth when a user must find paths in between nodes in a complex 3D network. Grossman et al. introduce some volumetric aspects, although the focus is on comparing types of stereoscopic and head-coupled motion parallax display technologies instead of volumetric rendering algorithms. They perform three experiments to evaluate the depth perception of 3D scenes with 3 different types of displays: a volumetric display, a stereoscopic display, a stereoscopic display with head-tracking [55]. The participants performed 3 tasks: ranking the depth of a sphere, tracing the path in a 3D network, and judging whether an object will collide with another object. Their results show the volumetric display has significantly better user performance than the others, but their results also show that stereoscopic display with head tracking is better than stereoscopy alone.

The most detailed evaluations of how to improve a user's perception and spatial understanding of volumetric data focus on comparing different rendering techniques and/or different transfer functions. A. Steward introduces a shading model named “vicinity shading” which provides shadowing to enhance the perception of surfaces within a volume [137]. Svakhine et al. introduce the 2D outline illustration technique which can be merged with a 3D volumetric medical dataset to improve the depth perception [140]. Bruckner et al. present a volumetric halo drawing technique to emphasize depth information of a volumetric medical dataset [25]. Boucheny et al. present perceptual studies that examine how transparency affects depth perception in Direct Volume Rendering (DVR) [17]. The participant determines the depth ordering of two cylinders with semi-transparency, two luminance conditions of cylinders (left cylinder dark or left cylinder bright) and three different background images matching to the intersection’s brightness. Overall performance of participants was relatively weak, but above chance level.

A few authors have studied the effectiveness of stereoscopic display for volumetric data. Maciejewski et al. introduce the Interactive Volume System which implements stereoscopic and haptic rendering with interactive transfer function to provide visual and haptic feedback to users [89]. They evaluate depth cues and accuracy of the user’s docking of two proteins using their system. The result shows the benefit of stereoscopy for finding the best docking configuration. However, the haptic transfer function did not provide any benefit. Hancock evaluates the depth perception of a volumetric dataset rendered using DVR on a stereoscopic display [61]. In these experiments, observers determine the relative depths of three smaller non-overlapping spheres which are either: presented directly; embedded in a large transparent in volumetric sphere; or embedded in a larger transparent sphere enhanced by modulating its opacity based on surface gradient. Observers viewed the objects using both a stereo and monoscopic display. Observers were most accurate under the transparent sphere condition while stereoscopic display improved accuracy only in the ‘presented directly’ condition. Overall, this indicated a dominance of aerial perspective over stereopsis for their scenes.

Some work examines how stereoscopy and structure-from-motion interact to effect perception of volumetric data. Kersten et al. show the effectiveness of stereopsis and simulated aerial perspective for the depth perception of a volume dataset of cylinders. The data are digitally reconstructed radiographs (DRR) [75]. Their results show that the stereoscopic display and stimulated aerial perspective provide depth cue better than opacity and spatial frequency for recognizing the rotation direction of the cylinder. Mora et al. show the effectiveness of order-independent direct volume rendering (Maximum Intensity Projection (MIP) and X-ray projection rendering techniques) [100]. They showed advantages of stereoscopy and transfer functions for enhancing the depth perception of a volume object rendered by MIP and X-ray projection renderings that lack other spatial depth cues. While compelling, their observations regarding stereoscopy’s effectiveness are anecdotal. Finally, Hubbold et al. present a technique named “tunneling” which allows users to see internal features and details of volumetric dataset rendered via DVR [72]. In their experiments, participants assess patterns of small blobs inside a volumetric brain. The experimental conditions include combinations of head-tracking, kinetic depth and stereoscopy. The results show that stereoscopy improves the users’ reported understanding of the depth structure and the combination of head-tracking with stereo is most preferred.

2.2 3D User Interface and Interaction

Many researchers have introduced 3D UI techniques for VEs. Bowman et al. [22,23,20,21] conducted many of the most recent, broad reviews of 3D UIs and ITs and have reviewed and evaluated a number of 2D and 3D ITs. They also have identified specifications of ITs that will improve the usability of 3D interactions in real-world applications and have proposed guidelines for future ITs [19]. Liu et al. explored modern ITs for 3D desktop personal computers [87]. A number of other articles also include review of physical input devices for 3D UIs [64,81], and ITs for large displays [56].

Several taxonomies of spatial input technologies (hardware) [32] have been created as well as taxonomies of 3D spatial user interaction techniques (software) [91]. A traditional mouse is a 2D held-device with 2 position DOFs. A 2D mouse with the ability to yaw perpendicular to the motion plane [91] is referred to here as a planar-3DOF device. Multi-touch is a body-tracking, 2D input with roughly 20 DOFs (10 fingers × 2 position DOFs). VIDEOPLACE was an early body-tracked 2D interface [90]. Notably the user was completely unencumbered (i.e. requiring no worn apparatus of any kind, not even fiducial markers).

3D input interacts in a 3D space. The bat [157] is an isotonic, 3D held-device with 6DOF pose (position and orientation). A bending-sensing data glove with a 6DOF tracker attached is categorized as 3D body-tracking, not a held-device. The ideal implementation of body-tracking, of course, is a completely unencumbered system. Wang et al. [150] demonstrate unencumbered hand plus finger-tracking. Our operational definition of body-tracking treats encumbered and unencumbered implementations as sub-categories.

Various researchers have demonstrated [177,79,65] that having a 3D held-device grasped in the hand is beneficial due to the tactile feedback (passive haptics) it provides for 3D manipulation. Such feedback does not exist in hand or finger-tracked 3D UIs, but does exist in 2D multi-touch UI's or 3D systems augmented with haptics.

Most ITs for 3D volumetric objects are quite different from ITs for polygonal objects, because of the ambiguity of the volumetric objects' surface. Many researchers have studied such interactions for polygonal datasets. For example, Steed [136] provides a literature review of selection methods in VEs, and Rick et al. [120] introduce GPU implementation of 3D object selection by the mouse cursor position for desktop environments. These studies introduce selection techniques for 3D polygon objects in the 3D virtual world. With 3D volumetric datasets, however, selection techniques like ray-selection and selection by volume are not always appropriate for selecting 3D volumetric objects because of the unique characteristic of the datasets. Therefore, several selection algorithms for 3D volumetric datasets are introduced such as using transfer function [120], selection box [144] and 2D interaction with a stylus tip with optical tracking [114].

2.2.1 Two-Handed Interaction Techniques

Bimanual interaction enriches interaction because humans often use two hands to accomplish real world tasks. A significant amount of research shows the advantages of bimanual interactions [7,8,33,48] based on Guiard's Kinetic Chain theory [58] that classifies different categories of bimanual actions.

Early two-handed spatial user interfaces used a 2DOF (degree-of-freedom) puck and slider [33], dual mice [14], a puck and stylus [83]. In 3D 6DOF UIs a number of two-handed interfaces were developed building off of one-handed techniques. These can be used for either object manipulation or travel using the grab-the-world approach [157]. Many travel techniques exist including those based on both rate control and position control [19].

Mapes and Moshell [92] develop the two-handed pinch glove based interface for an HMD based virtual reality system. Hinckley et al. [66] develop a two-handed interface for non-immersive, desktop 3D medical imaging system using hand-held props. Shaw and Green [129] present a two-handed interface on a non-immersive desktop system for polygon surface design. They add 3 buttons to a pair Polhemus Fastrak receivers [115] creating dual button enhanced bats [157]. Zhai [177] presents and studies the FingerBall, a small ball with a single button activated by squeezing. The size was selected to allow a precision grip and their experiments demonstrate that for a one-hand 6DOF docking task the FingerBall is faster than a pinch glove based technique.

A common form-factor for 6DOF devices is using a joystick handle [19]. This power-grip approach has been carried over to two-handed 6DOF systems such as the SpaceGripsTM which are used by Schultheis et al. [127]. They add the Spindle visual to the IT of Mapes and Moshell based on visual feedback in a two-handed 2D IT of Balakrishnan and Hinckley [7]. Experimentally they find the two-handed IT out-performs a mouse IT and 6DOF Wanda technique for a 6DOF docking task and an object construction task. While the IT allows view scale change, the objects that are docked do not need to be re-scaled. Hence, the view scale adjustment is available for the user to augment translational travel and to find an ideal view scale for performing the docking task, but strictly speaking the docking task does not require a view scale adjustment.

Ulinski et al. [146,145] explore different two-handed techniques for selecting subsets of volume data. The techniques create and manipulate a 3D box which has 9 DOFs total, although it can be reduced to 7 if the box is assumed to be a cube. They also develop and compare a number of two-handed techniques based on Guiard’s [58] kinetic chain theory.

Similar to Mapes and Moshell, their Two-Corners Technique does not support pitch about the axis between the button balls.

Most recently, Song et al. [133] present a gesture-based 7DOF mid-air two-handed object manipulation technique by Microsoft Kinect. However, the IT doesn’t provide continuous 7DOF: it allows simultaneous 6DOF and separate pitch control. Furthermore, the free hand mid-air interaction increases arm fatigue rate significantly and fails to avoid the “Gorilla arm” problem.

2.2.2 Hybrid User Interface

The term hybrid user interface (HUI) refers to a UI with multiple methods for spatial input, frequently supporting both bimanual or unimanual interaction and 2D and 3D interaction. Benko et al. [10] combine a multi-touch 2D surface with hand and finger 3D gestures and 3D interaction in an augmented reality system. They coin the terms HUI and cross-dimensional gestures.

Some earlier devices support a similar notion of cross-dimensional interaction. The VideoMouse [67] and the Logitech 2D/6D Mouse [88] are a single device that support both 6DOF mode and planar-3DOF mode. However, in neither system was this concept extensively developed into a hybrid 2D/3D UI nor was two-handed interaction supported. The utility of confining the motion of 6DOF device to a physical plane, such as a held tablet, to reduce the physically manipulated DOF’s has been demonstrated [22]. However, these prior works do not use a significant displacement between the physical device and its representative 2D or 3D cursor (as in [129]) and neither of these works’ UI’s implement the 6DOF to planar-3DOF mode switching.

Massink et al. introduced HyNet, a HUI system for desktop-based navigation [94]. This work uses a traditional mouse for navigating the 3D world with a conventional desktop system. However, the system only uses 2D GUIs with 2D UIs and does not provide a solution for 3D visualizations and VE systems. The authors also introduce a programming abstraction for the HUI, with traditional desktop-based systems that used conventional mouse and keyboard inputs. The HUI addresses both theoretical abstraction and 3D input modalities.

Alencar et al. [3] present HybridDesk that combines 2D and 3D interactions with a tracked Wiimote and WIMP interface for an oil platform visualization. There are three UIs in HybridDesk used to evaluate their HUI techniques: VR-Nav for navigation and selection, VR-Manip for manipulation, and the traditional WIMP UI. More recently, Magic Desk [15] utilizes multi-touch input, a mouse, and a keyboard within a traditional desktop environment for unimanual and bimanual interactions. The authors explore suitable physical positions of multi-touch input relative to the user during the experiment.

Althoff et al. [4] present a multimodal interface for navigation in arbitrary virtual VRML worlds, which uses a mouse, keyboard, joystick, and multi-touch input. However, their environment was limited to 2D visualizations and 2D interactions. The Slice WIM interface [35], which uses a multi-touch table with a head-tracked, stereoscopic wall screen display for a medical imaging volumetric dataset, allows multi-touch interaction on the table to control 3D data using two widgets.

Multimodal user interfaces (MUI) generally use more than just spatial input; for instance they combine voice and gesture [36,70]. Bolt [16] introduces a system called “put-that-there,” which uses voice and gaze inputs. Within GIS systems, voice and gaze inputs also are popular interaction methods in MUIs [2,119]. The main advantage of natural human input modes is that they do not require any held-device and users need less training.

HUIs and MUIs can be combined with augmented reality as well. ICARE is an example of such a mixed environment [18]. Bianchi et al. [13] develop a hybrid AR system, which used a hybrid external optical tracker for the user's head pose and a subsequent visual landmark-based refinement of the pose estimation that uses AR's overlaying of virtual objects on the user's real environment [5]. Other previous works include medical volumetric datasets design for use by surgeons [114,126].

Many HUI and MUI systems incorporate hand-held, mobile devices. Song et al. [132] introduce an application called what-you-see-is-what-you-feel that uses a mobile device for input and a wall-mounted display for medical imaging volumetric data visualization. Users employ 2D multi-touch input on the handheld device to manipulate the 3D medical volume data on the large wall-mounted display through the wireless network.

Researchers also can use HUIs and MUIs in collaborative systems. Each user can handle a different system employing heterogeneous displays with various techniques to share the visualization or data with other colleagues. Schmalstieg et al. [126] introduce a mixed reality environment that combines AR, ubiquitous computing, and a desktop metaphor for a collaborative system used with medical volume data.

2.2.3 Natural hands input

In the past few years, the multi-touch input has become a more popular natural input mode for direct, intuitive interaction. Much research has revealed the benefits of multi-touch input through different types of heterogeneous displays, such as tabletops [41], desktop-based systems [11] and wall-size displays [110,147]. Benko et al. [11] introduce precise multi-touch selection techniques for a desktop-based vertical multi-touch display using different type of traditional WIMP 2D GUIs like x-menu, slider, stretch, and offset.

Forlines et al. also demonstrate that the benefits of direct touch input outweighed those of mouse input. Direct input reduces selection time for many, but not all, selection techniques (c.f. docking object by unimanual selection) with a vertical orientation display [51]. Unfortunately, these previous studies do not speak to 3D interaction.

Bradley and Roth [24] demonstrate untethered computer vision tracking of a fist-sized ball, but occlusion remains a problem, especially for a two-handed scenario. Current battery and sensor technology still precludes constructing an accurate, small-form factor wireless 6DOF ball, but this area of engineering is very active [166]. Finally, non-isomorphic rotation techniques [22] can ameliorate cord entanglement during rotation operations.

Multi-touch tabletops also support multi-user collaboration. DiamondTouch [41], for example, is designed for multi-user interfaces. Kin et al. [76] explore the benefits of direct-touch and multi-finger input with bimanual interaction for the single task of multi-target selection. Their results show that bimanual multi-direct-touch input provides fewer additional benefits than the mouse input. In addition, using more than two fingers can reduce the targeting accuracy, and does not provide any additional benefit for multi-target selection. Daniel et al. [169] introduce a multi-touch interface with two-sided (bottom and top of the surface) interactive table display.

One of problems with multi-touch input is a lack of owner identification of fingers or hands. Westerman [167] describe methods for tracking and identifying multiple fingers and hands on multi-touch surface systems that generate X-ray-like images of a hand with an opaque sensing surface. Similarly, Marquardt et al. [93] introduce glove-based multi-touch input, which identified parts of the user’s hand (fingertip, knuckles, palms, sides, and back of the hand).

Another approach to user identification is to design special purpose non-planar interactive surfaces for the multi-touch and gesture direct input. Benko et al. [9] design special displays for direct two-hand touch input, with a 3D volumetric dataset: Sphere, which is a spherical multi-touch sensitive system; Pinch-the-Sky Dome, which is an above-the-space depth-aware interaction large curved display; and DepthTouch, which is a horizontal, depth-aware, multi-touch system. Like the work of Benko et al., Grossman et al. [57] introduce a volumetric display with 2D and 3D gestures.

Several research papers introduce multi-touch systems with stereoscopic display. Toucheo [59] interacts with co-located 3D stereoscopic visualization via multi-touch input and has two screen layers. The top layer is for the stereoscopic display, and the bottom layer is for multi-touch input using a transparency panel to display a 3D image. Valkov et al. [147] suggest a 2D input of a vertical wall-size display for 3D stereoscopic objects.

Some approaches provide 3D ITs for 3D objects in multi-touch tabletop systems. Wilson [170] presents a 3D tangible tabletop system that uses depth-sensing video cameras for a 3D input. Zimmerman and Lanier [178] present a glove-based gesture interface for 3D object manipulation. Segen et al. [128] introduce a camera system that recognized 3D gestures and postures. Their results show that the system is fast and robust enough to compute different poses and provided examples of how to use the system with 3D graphical editors, VR applications, and video games.

2.2.4 7DOF Travel and Docking

View scale has significant effects on usability in systems with head-coupled display [112], stereoscopic display [156] or direct 3D manipulation [97]. Therefore, not only is it desirable for MSVE systems to support 7DOF multi-scale 3D UIs—rather than just 6DOF multi-scale 3D UI's—it often is a requirement.

Many 7DOF multi-scale travel techniques exist including those based on both rate control and position control. In general, many IT's can be used for either object manipulation or travel using the grab-the-world approach [158].

2.3 Multi-Scale Virtual Environment

Some 3D UIs for MSVEs do not support view scale as 7th DOF because their underlying view model lacks the sophistication of [122]. Instead, "zooming-in" occurs through 6DOF view adjustment (dollying) with some auto-adjustment applied to travel velocity and possibly to the near/far clipping planes to manage zbuffer precision. However, early VR work [122] demonstrate 7DOF travel techniques as well as the benefit of view scale differences in multi-user VEs [84]. (This had been observed for 2D multi-user environments earlier [79]).

Various previous works use specific navigation techniques for MSVEs. Pierce and Pausch [113] propose a navigation technique for better scalability to large virtual world, with visible landmarks allowing users to travel in the vicinity with a single gesture and with symbolic place representations allowing users to travel to distant locations with a small number of gestures. Kopper et al. [78] present the design and evaluation of two navigation techniques for MSVEs. They find that automatic scaling is more efficient than manual scaling and target-based navigation performs better than steering-based navigation. Wu et al. [172] evaluate way-finding aids interface (view-in-view map, animation guide and human system collaboration) in a MSVE. The result of their experiment shows the view-in-view map offers the best performance overall. Bacim [6] provide understanding and classification of way-finding information (hierarchical and spatial information) needed for travelling in a MSVE. The result shows new techniques help users perform better in both travelling and way-finding aid, although from different perspective. Trindade et al. [143] propose improvements to two existing interfaces in order to assist and facilitate the task of navigating in a 3D VE. For flying they include support for collision handling and automatic navigation speed adjustment with respect to scale. For exo-centric travel, they use a point-of-interest technique with an automatic pivot point based on the construction and maintenance of a depth CubeMap. Their result show significant improvement in the execution of navigation tasks.

Hougast et al. [71] and Wartell et al. [164] develop a virtual workbench application which balances interaction and stereoscopic display for a multi-scale volumetric weather visualization. They find a trade-off between direct manipulation and stereoscopic display, which must be optimized to help users perceive the environment. However, no formal evaluations are described.

Oh and Hua [103] present a user study on three multi-scale visualization interfaces on a 3D workbench display: focus plus context, fixed focus plus context, and overview plus detail, with the purpose of identifying the differences of these interfaces with two tasks (path following and 3D map reading) in large scale information visualization on the 3D workbench.

Glueck et al. [54] argue that the design of 3D ITs in VEs is an ill-defined problem. They develop an abstract model to illustrate the cyclic relationship between interaction and navigation in VEs and argue that navigating and understanding must be evaluated simultaneously. Finally, they highlight strategies to support the design of interactions in MSVEs and propose general categories of research focus.

2.4 Stereoscopic Fusion Limit

Stereo fusion problems have been studied in stereo media [86] and computer graphics [68] and continues to be investigated [130]. Stereo image depth adjustment via by deliberate altering the eye separation pre-dates computer graphics [86]. Underestimated modeled eye separation (e.g. using 3 cm instead of 6 cm) can compress the depth non-linearly of the stereo image to reduce stereo fusion problems. The distortion is more specifically a non-affine homology [163]. Setting the modeled to the true separation and creating only a virtual-to-physical difference is equivalent to applying a uniform scale transform to the image and this technique also predates computer graphics [125]. Ware et al. [151] develop the latter into the cyclopean scale, a dynamic adjustment where the VE is dynamically scaled with the scale's fixed point between the stereo frustum's center of projections. Wartell et al. [162] classify 9 other prior stereo image depth adjustment methods circa 2001.

Later, Holliman et al [69] propose an approach for stereoscopic image creation which allows a defined region of interest in scene depth to have an improved perceived depth representation compared to other regions and which can keep this mapping constant even if total scene depth is changing. They also present a novel three-region algorithm for stereoscopic image capture.

Lambooij et al. [82] review the concept of visual fatigue to clarify the importance of various causes and aspects of visual comfort in relation to stereoscopic display and image generation. They indicate that even within the sufficient range allowing for satisfactory depth perception provided by one degree limit of disparity, visual discomfort may still occur due to the factors: (1) excessive demand of accommodation-convergence linkage, (2) 3D artifacts resulting from insufficient depth information in the retinal images yielding spatial and temporal inconsistencies, and (3) unnatural amounts of blur.

Carvalho et al. [34] propose a technique to dynamically adjust stereo parameters based on a CubeMap structure [95] during the usage of two VR tools: fly and examine, in an MSVE.

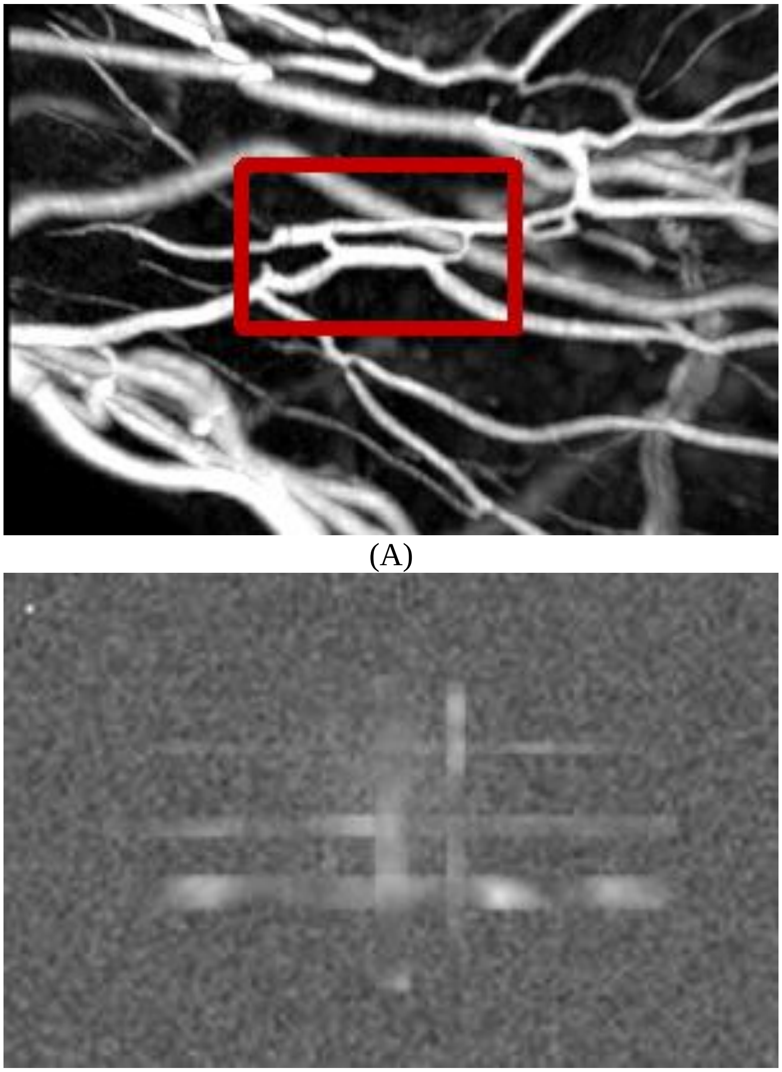
2.5 Interactive 3D Volume Visualization

Some researchers have explored the use of VEs for 3D volumetric visualization with stereoscopic and head-tracking displays like Stereoscopic Field Analyzer for flow visualization [43], ocean flow [31], Virtual WindTunnel for complex fluid flows [29], and molecular visualization [12]. However, these previous works use either traditional WIMP interfaces or unimodal 3D UIs, and do not explore HUIs.

Due to the unique characteristics of 3D volumetric datasets, the design of UITs may need to differ from UITs for 3D polygonal data. Bruckner et al. [26] introduce VolumeShop, which is an interactive system for medical 3D volumetric datasets. However, their system does not provide a full solution for 3D interaction because of the lack of DOF by 2D GUI. Engelmeier et al. [47] introduce a system for the 3D medical volumetric data that allowed medical professionals to navigate through a patient’s 3D scans. This MUI combines natural speech inputs, eye tracking, and glove-based hand gesture recognition. Their work leads to considerable enhancements in the speed and efficiency of diagnoses for doctors who used the new MUI.

Ulinski et al. [146] examine asymmetric and symmetric bimanual selection to draw a selection box to select a part of 3D temperature volume dataset. Their work demonstrates that symmetric bimanual interaction was more efficient than asymmetric methods for drawing the selection box. However, their approach focuses only on 3D volumetric selection, and did not examine other ITs. Although previous research involving 3D UI primarily involved 3D volumetric datasets, some limited the types of interaction (e.g. selection, and GUI) and others were limited to WIMP based interaction with 3D volumetric datasets.

CHAPTER 3: EVALUATION EFFECT OF VIRTUAL REALITY TECHNIQUES FOR 3D VOLUMETRIC DATASETS



(A)

(B)

Figure 1: Similarity comparison between our artificial dataset and actual MRI blood vessel scan. (A) Maximum Intensity Projection (MIP) rendering of blood vessels [100]. (B) Our artificial dataset.

3.1 Introduction

This chapter, we present results of two experiments on the benefits of stereoscopy and head-tracking for a person's correct perception of depth ordering of volumetric objects. The experiment is motivated by dataset such as the MRI scan of blood vessels shown in Figure 1A reproduced from [100]. As is typical of volumetric data, this dataset is has a heavy presence of transparency, occlusion and a highly ambiguous spatial structure. In Figure 1A, it is particularly challenging to determine the depth order of the blood vessel inside the red square. As discussed in [89], the volume rendering technique used here makes it appear that the square-shaped loop vessel is in front of the diagonal one. However, in fact the diagonal one is in front of the square-shaped loop.

We mimic this type of ambiguity by generating controlled experimental dataset such as Figure 1B where the user’s task is to determine the depth ordering of various occluding transparent cylinders. The subjects view the datasets under a variety of display conditions including combinations of stereoscopic display, head-tracking, and small object rotations. We present a set of cylinders of various size, opacities and depth orderings to mimic datasets such Figure 1A but in an experimentally controlled manner. To isolate the effect of semi-immersive display conditions the user is not allowed to alter that transfer function or other volumetric rendering parameters such as lighting or switching between rendering methods. The two experiments include a depth-ordering task, in which participants must understand the full depth ordering between six volumetric cylinders and a depth discrimination task, in which participants must distinguish the relative order of just two cylinders within a limited exposure time (2 sec). In addition, the experiments are also designed to differentiate the differences between experienced and less-experienced users with respect to experience with 3D games and VR related technology. Results from both groups show an overall benefit for stereoscopy with head-tracking in enhancing depth perception of volumetric data. More interestingly, our study also suggests that familiarity with 3D games and VR related technology significantly affects the user's depth perception accuracy.

3.2 Environment

Our study tests the effectiveness of a semi-immersive VR display on depth perception of volumetric data. We examine the effects of display environment on two tasks: a depth ordering task, in which participants determine the general depth ordering among six volumetric cylinders with no time limit; and a depth discrimination task, in which participants must distinguish the depths of a pair of cylinders within a short time limit (2 sec).

Figure 2: Environment system. This picture shows our environment system for experiments. A 3D display with shutter glasses, Polhemus tracker, and a chin-rest.

As shown in Figure 2, both tasks use a desktop VR setup which consists of a stereo display (22" Samsung Sync Master 2233RZ) and a tracked pair of stereo glasses. The tracker is a Polhemus Fastrak. The glasses are by Nvidia. The display rate is 120 Hz time-multiplexed to 60 Hz per eye. The subjects sit roughly 60 cm in front of the screen. The independent experimental variable is the display condition which a combination of stereoscopy, head-tracking and/or a small object rotation for a kinetic depth effect.

We informally experimented with using various freely available volumetric dataset including an engine block, an orange, a tree and medical data. We considered having participants identify a variety of different shape characteristics. Base on pilot test for a variety of scenarios, we narrowed our experimental design to depth discrimination and ordering tasks. Further, we chose to create an artificially generated dataset loosely based on the types of depth ambiguities found in the blood vessel example (Figure 1A).

Figure 3: This figure shows one of our volumetric datasets. It has six-cylinders (three horizontal and three vertical). Perlin Noise modulates the 3D texture. In the actual experiment, the voxels are far more transparent.

These decisions are motivated as follows. First, there is a much larger range of rendering algorithms and rendering parameters for volumetric data than there are for surfaces and 3D networks. The choice of volume rendering technique or transfer function can make a huge difference to an observer’s perceptual understanding of a volumetric dataset’s spatial structure. This complicates creating a controlled experiment on shape or depth perception of volumetric data. Second, after testing various datasets with various volume rendering techniques and rendering parameters and after considering various perceptual questions, we concluded it was necessary to artificially generate a volumetric dataset where we could completely control these many factors and in order to present different volumetric datasets on each trial to avoid learning effects. At the same time, we wanted the created dataset to roughly mimic perceptual ambiguities found in a real-world, volumetric dataset. Therefore, we developed a dataset that contains depth ambiguities between volumetric tubes inspired by 3D medical scans of networks of blood vessels.

As shown in Figure 3, our synthetic volumetric dataset contains six overlapping cylinders of varying diameters and transparency. In this image, we decreased the transparency levels from that used in the experiment for expository purposes. Three cylinders are vertical and three and horizontal. The voxel resolution is 512 × 256 × 256. Perlin noise is used for the internal texture of the cylinders and the texture of a large background polygon. In the experiment an additional background polygon approximates the visual effect of having the cylinders embedded in a more complex volumetric environment.

To minimize a participant's knowledge gained from performing perceptual tasks on the same dataset over many repeat trials, the synthetic dataset present varies from trial to trial. Each cylinder in a trial is assigned a random depth location, a unique noise texture, a direction (vertical or horizontal), and a random cylinder size (thin, medium and large). A non-randomized dataset might allow a participant to consciously or non-consciously pick up and respond to coincidental relationships between these parameters such as "the horizontal cylinders are always further away" or "the thinnest cylinder is always the one farthest away." Avoiding such conflating factors is a key reason we use an artificial dataset rather than a single real-world data set.

3.3 Rendering Technique

We choose a high quality GPU-based ray-casting rendering techniques [80,124]. Compared to other rendering techniques, such as per pixel lighting and MIP rendering [17,85], the GPU based ray casting technique yields more accurate depth cues [124]. The renderer is available as an OpenSceneGraph plugin [105].

We add a black polygon with a square hole in front of the volumetric data to act as a window to hide the ends of the cylinders. This was necessary because being able to see the cylinder ends made the depth ordering task trivial. In real-world datasets such as the blood vessel example (Figure 1A), the complex intertwined paths of the tubes typically tend to obscure the tube endpoints. We also tested scaling up the voxel volume's rendered size (to extend the cylinders' ends off screen) but this failed because aliasing artifacts were too visible. Attempting to increase the volumetric resolution beyond 512 × 256 × 256 to counter aliasing problems exceeded the renderer's memory limitations.

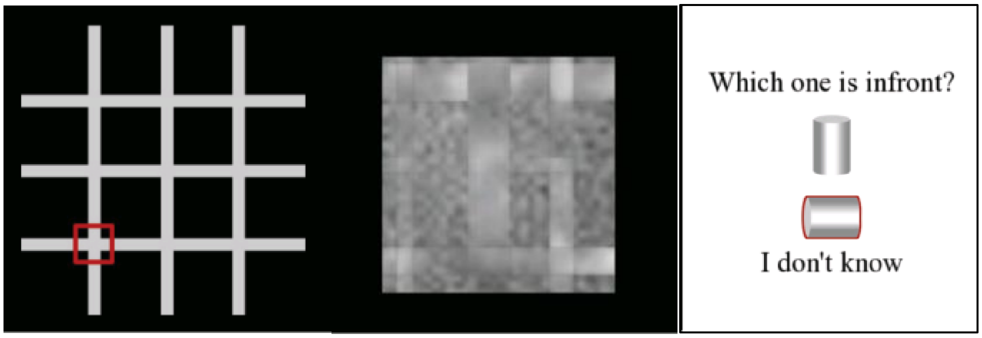
Following a previous study [100], we fix the data parameters, such as the Alpha gradient and transparency, to represent a reasonably clear outline of each semi-transparent cylinder (alpha = 0.9, transparency = 0.2, density = 0.025). Note that, to isolate the effects of stereo and structure-from-motion, we do not allow users to interactively adjust the transfer function in our study even though many previous studies demonstrated its utility in depth perception [77,111].

3.4 Experiment Design

Our two experiments examine the effect of stereoscopy and/or head-tracking on the perception of volumetric data. Experiment 1 has four display conditions and Experiment 2 has six. We use a within-subject design with repeated measure. Each subject is randomly assigned a sequence of display conditions using Latin squares. The measures in our experiment are answering time and error rate. Before each experiment, the participant provides demographic information such as gender, academic major and degree being sought. A questionnaire inquired regarding their familiarity with stereoscopic display, VR technology and gaming. Questions include: how often do you play games on a computer or game console with/without motion capture devices such as Xbox Kinect and Sony Playstation Move; how often do you watch 3D movies in the theatre; how often do you sue 3D displays for movies or games. After the experiment, each participant fills out a post-questionnaire regarding their confidence in their answers to the task's spatial questions and their opinions on various visual aspects of the volumetric dataset such as transparency, the noise background, etc.

We recruited twenty eight participants, twelve for Experiment 1 and sixteen for Experiment 2. Sixteen of them are undergraduate students and twelve are graduate students. Fourteen participants major in Computer Science and fourteen participants are of other majors including psychology, nursing history and fashion design. All participants have (corrected) 20/20 vision. We provide a tutorial to familiarize the participants with the stereo display and head-tracking hardware. We designed two experiments. Experiment 1 examines the effect of stereoscopy and head-tracking on a depth discrimination task. In this task, subjects are exposed to the volumetric dataset for a short amount of time (2 sec) so that they do not have time to cognitively reason about the depth order based on factors such as transparency, window size, etc. (In many psychophysics studies the exposure time is usually in the range of a few hundred of milliseconds but 2 sec is common in stereoscopic VR studies [152]). Experiment 1 requires the subject to first locate an intersection of a pair of cylinders based on a provided instructional cue, and to then report on the depth relation of the cylinder pair. The 2 second exposure time allows for vergence eye movements [173]. Experiment 2 explores the effect of stereoscopy and head-tracking on the task of depth ordering which requires distinguishing the depth order of multiple cylinders, not just a single pair. Experiment 2 examines the displays' effect within the context of an unlimited exposure time.

3.5 Experiment 1: Depth Discrimination



(A) (B) (C)

Figure 4: Three screens displayed in Experiment 1's trial. (A) The instructional cue indicating the target cylinder pair to examine. (B) The volumetric cylinders seen through an aperture. (C) The question the participant answers for the trial.

Experiment 1 evaluates how stereoscopy and structure-from-motion affect performance on a depth discrimination task. The participant determines which of two cylinders, one horizontal and the other vertical, is in front of the other. The volumetric dataset contains six cylinders, but in each trial a pair of cylinders is designated as the target pair for the trial. On each trial, the first screen displays a 2D picture (Figure 4A) where a red box designates which of the nine intersections of the six cylinders is the target pair. The next screen displays the volumetric dataset for 2 seconds (Figure 4B). The final screen displays a menu with three choices (Figure 4C): "the horizontal cylinder is in front", "the vertical cylinder is in front", or "I don't know". Note, we choose not to use a force-choice protocol in this experiment because we want to gather data on how often a user feels they cannot determine the depth ordering. A force-choice protocol would have conflated results for trials where participants were guessing at the depth order with those trials in which they felt they could determine a specific ordering.

Experiment 1 has six display conditions. The conditions are: non-stereo without no motion (NS-NM), stereo with no-motion (S-NM), non-stereo with head-tracking (NS-H), stereo with head-tracking (S-H), non-stereo with kinetic-depth effect (NS-KD) and stereo with kinetic-depth effect (S-KD). The last two conditions were added because in pilot tests, not all users utilized the head-tracking when limited to the 2 second exposure time. In particular, some users did not attempt to use a quick head motion to gain motion parallax cues even when we were careful to specifically remind them this was possible. Hence, the kinetic-depth effect condition automatically rotates the cylinders left and right by 10 degrees. For a small range of motion the visual effect is similar to having the participant quickly move her head left and right. In the non-head-tracking conditions, a participant uses a chin rest. In this condition, the view frustums are calibrated for this fixed head position.

Each participant performed 324 trials in blocks of 54 where each block used one of the 6 display conditions. Display condition block order was counter-balanced using Latin squares.

3.5.1 Result

Quantitative

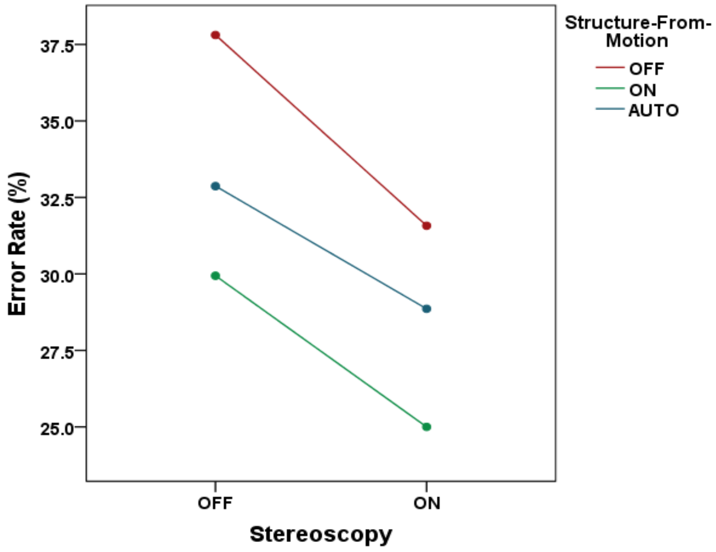


Figure 5: Effect of stereoscopic and motion on error rate of Experiment 1.

Table 1: Average and standard deviation (SD) of error rate for each condition of Experiment 1.

Display Condition	Error rate (%)	
	Mean	SD
S-H	25.00	15.82
S-KD	28.86	17.15
NS-H	29.94	16.12
S-NM	32.41	16.61
NS-KD	32.87	15.68
NS-NM	37.81	17.00

We analyzed the results using a two-way repeated measures (*rm*) ANOVA (2 × 3) followed by Fishers' least significant difference (LSD) for pairwise comparisons with α=.05 level of significance. Error rate measures the percentage of incorrect depth judgments counting "I don't know" answers as incorrect. The result shows no significant interaction between stereo and motion on error rate (Figure 5). Stereoscopy has a simple main effect on error rate (F(1,11)=8.5, p=.014, η_p^2=.316) decreasing the error rate from 33.5% to 28.5%. The motion condition simple main effect was not significant.

Table 1 shows means and standard deviations of error rate for each condition. The one-way *rm* ANOVA shows a main **e**ffect of the general display condition (NS-NM, S-NM, NS-H, S-H, NS-KD, S-KD) on error rate (F(2.512,27.635)=3.549, p=.034, η_p^2=.244). LSD post-hoc tests show the following. The mean error rate of S-KD is significantly lower than NS-NM (p=.012) and S-NM (p=.034). The mean error rate of S-H is significantly lower than S-NM (p=.038), NS-H (p=.038) and NS-NM (p=.009). And the mean error rate of S-NM is significantly lower than NS-NM (p=.038).

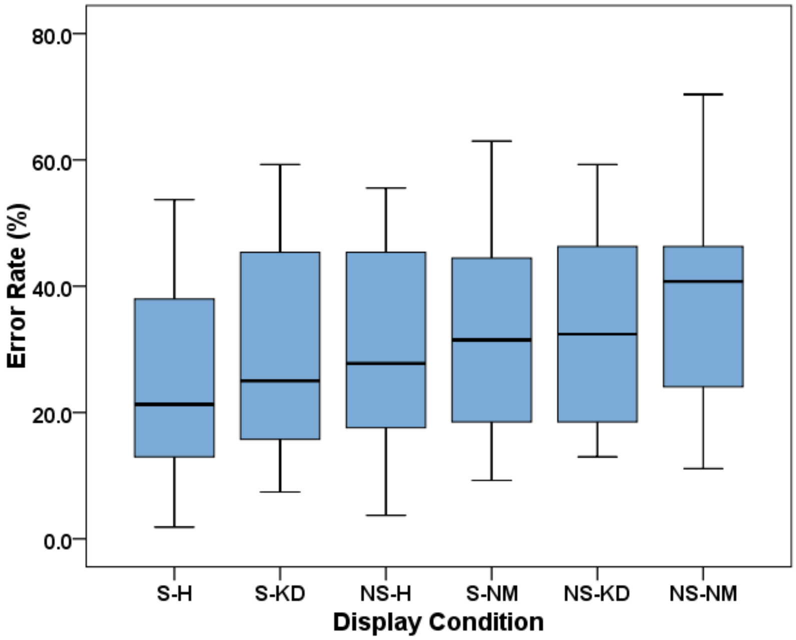


Figure 6: Box plots of error rate by display conditions of depth discrimination experiment.

We observed that some participants appeared more confident with their task performance and more comfortable with using our semi-immersive VR environment. Also, our participants came from two pools, a computer science (CS) pool all of whom were computer science majors, and a psychology (PSYC) pool, which were psychology and liberal arts majors. The one-way between-subjects ANOVA shows a main effect of pool: CS pool average error rate is 27% while PSYC pool average is 36% (p=.001).

Participants from the two pools were randomly assigned a display condition order and they participated over the same time period, therefore, we analyzed the pre-questionnaire. An example question is: “How often do you play 3D computer games?” Answers are on a 7 point scale with 1 being “Never” and 7 being “A Great Deal”. On a number of these questions CS pool scored significantly higher on this scale for game playing experience. In particular some results were: 2D game playing (mean 4.8 vs. 2.3 at p<.001), gaming on a PC (mean 4.8 vs. 3.2 at p=.017), gaming on a console (mean 4.3 vs.

2.7 at p=.001), gaming with motion capture devices (mean 2.5 vs. 1.5 at p=.048), and stereoscopic 3D TV usage (2.2 vs. 1.0 at p=.05).

We apply a mixed three-factor within subjects ANOVA to evaluate the effect of pool × motion × stereo (2 × (3 × 2))) on error rate. There is no significant 3-way interaction (F(5,20)=1.014). Regarding the two-way interactions, stereo × pool is not significant. (The significant main effect of stereo is reported in the prior section). However, motion × pool is significant (F(2, 20)=3.690, p=0.043, η_p^2=.270).

The simple main effects for motion are as follows. For the PSYC pool motion is not significant but for the CS pool the main effect of motion is significant (F(2,10)=4.269, p=0.029. η_p^2=.299). LSD pairwise comparisons for the CS pool show head-tracking conditions are better than no-motion conditions, with average error rates 18.8% vs. 32.3% (p=0.09), and kinetic-depth-effect is better than no motion, with average error rates are 18.8% vs. 28.7% (p=0.038).

Plausibly CS pool subject’s greater experience with gaming trains a person to better attend to various depth cues when viewing computer generated 3D images and increases their sense of confidence in using VR type technologies. Alternatively, the CS majors might have simply been more interested in the technology employed and hence were somewhat more motivated. However, given that participants perform 324 trials, we suspect CS pools greater experience played a larger role than interest level.

In summary, for all participants stereo had a generally positive, significant main effect while only for the CS pool does the motion condition have a generally positive significant effect. In general, the CS pool participants performed better overall than PSYC pool participants.

Qualitative

On a 7-point Likert scale, regarding how confidence the participant is for her answer in such conditions (1 *not at all* through 7 *great deal*), participants answered they were more confidence with S-H (M=4.67) than other conditions (S-NM=4.5, NS-H=4.17, NS-NM=4.17, S-KD=4.17, and NS-KD=4.08). However, there was no statistical significant effect (*p*=.499). When asked which condition was the most effective, six out of twelve answered S-H, three answered S-KD, two answered NS-H, and one answered NS-KD. When asked which VR technique was better either motion or stereo, six answered stereo was better, five answered motion, and one answered both stereo and motion were same.

3.6 Experiment 2: Depth Ordering

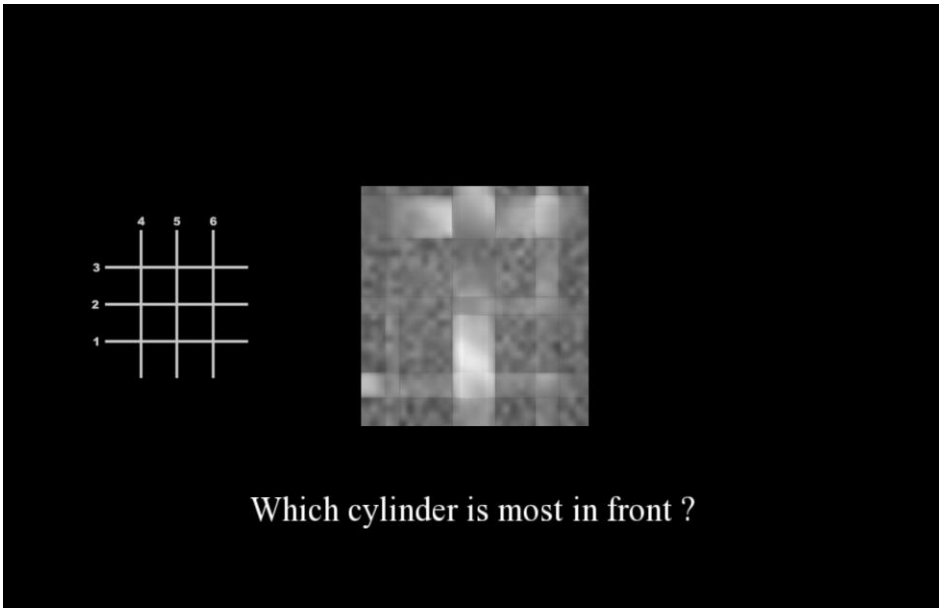


Figure 7: Single screen used in depth ordering experiment.

In Experiment 2, participants perform a depth ordering task on the six volumetric cylinders. Because the trial duration is unlimited and participants have ample time to use head-coupled motion parallax, the kinetic-depth-effect (e.g. auto-rotation) conditions are not included. The four display conditions are: non-stereo without head-tracking (NS-NH), stereo without head-tracking (S-NH), non-stereo with head-tracking (NS-H), and stereo with head-tracking (S-H). Each cylinder is labeled with a number (1-6). The participant must designate which of the six cylinders is at a particular position either: the front, the middle, or the back. The particular position queried is randomly determined per trial. (Two answers are counted as correct for 'middle'). For each trial, the cylinders are rendered with random depth ordering. Figure 7 shows the displayed screen.

The participant designates which cylinder is at the queried position by pressing the corresponding number key on the keyboard. In non-head-tracking conditions participants use a chin rest as in Experiment 1. Each participant undergoes 36 trials per display condition which means 144 trials total. Trials are in blocks by display condition and the block ordering uses Latin squares.

3.6.1 Result

Quantitative

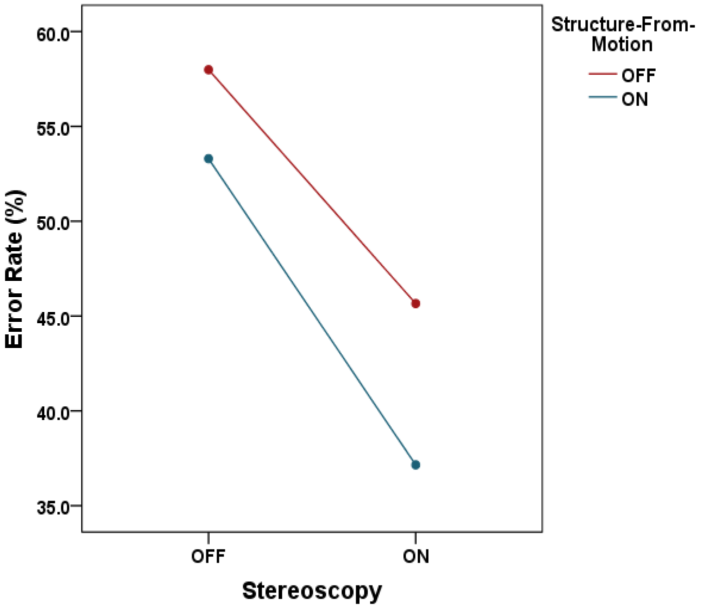


Figure 8: Error rate of stereoscopy by structure-from-motion of Experiment 2.

Table 2: Mean and standard deviation (SD) of error rate and answering time for all display conditions in Experiment 2

Display Condition	Error rate (%)		Answering Time (sec)	
	Mean	SD	Mean	SD
S-H	37.15	19.27	13.89	2.46
S-NH	45.66	14.59	12.61	2.30
NS-H	53.30	11.08	15.56	4.43
NS-NH	57.99	12.13	14.11	3.56

We analyze the effect of the display condition on answering time and error rate. Error rate is computed as the number of incorrect answers divided by total number of questions in each trial (36 questions per condition).

We analyzed the results using a two-way repeated measures (*rm*) ANOVA (2 × 3) followed by Fishers' least significant difference (LSD) for pairwise comparisons with α=.05 level of significance. Table 2 shows mean and standard deviation of error rate and answering time for all display conditions. The results show no interaction effect on error rate between stereoscopic and head-tracking display. Head-tracking has a main effect on error rate (F(1,15)=6.934, p=.019, η_p^2=.316). Stereoscopy has a main effect on error rate (F(1,15)=23.07, p<.001, η_p^2=.606). Figure 8 indicates the lack of interaction. There is no effect of display order condition and unlike in Experiment 1, no interactions with participant pool (CS vs. PSYC) are significant.

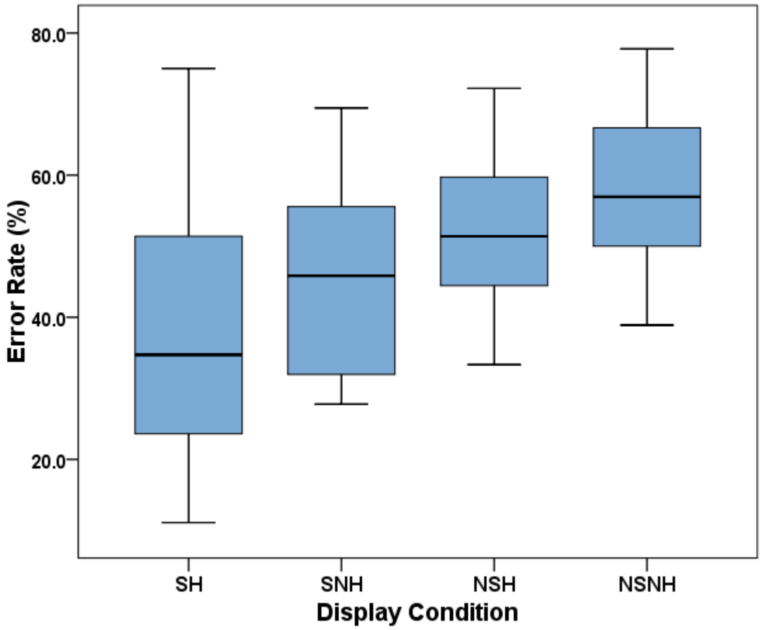


Figure 9: Box plot of error rate by display conditions of Experiment 2.

The one-way *rm* ANOVA shows a main effect of the combined display condition on error rate (F(3,45)=11.047, p<. 001, η_p^2=.424). LSD tests show the mean error rate for condition S-H (stereo with head-tracking) is significantly lower than all other three display conditions (NS-NH (p<.001), S-NH (p=.012) and NS-H (p=.004)) and S-NH is significantly lower than NS-NH (p<.001). Figure 9 illustrates mean error rate across all four conditions. Unexpectedly, head-tracking alone (NS-H) does not lead to significant improvement in accuracy over the no stereo no head-tracking (NS-NH) condition (p=.308).

Average task completion time across all conditions is 14.4s. We expected the addition of stereo to reduce response time. While the stereo means were faster the differences were not significant. Further, tests for three-way (including participant pool), two-way and one-way ANOVA are not significant.

Qualitative

On a 7-point Likert scale, regarding how confidence the participant is for her answer in such conditions (1 *not at all* through 7 *great deal*), participants answered they were more confidence for their answer with S-H (M=5.31) than other conditions (NS-H=4.38, S-NH=4.0 and NS-NH=3.38). The one-way repeated measures ANOVA shows that there was a main effect on task confidence of display condition (F(3,45)=19.825, p=.001, η_p^2=.431). LSD pairwise comparisons show confidence rate of S-H (M=5.3, SD=1.1) was higher than NS-H (M=4.4, SD=1.1, p=.014), S-NH (M=4.0, SD=0.7, p<.001) and NS-NH (M=3.4, SD=1.4, p<.001). When asked which condition was the most effective, fourteen of sixteen answered S-H, and two answered NS-H. When asked which VR technique was better either motion or stereo, nine of participants answered motion, six answered stereo, and one answered both stereo and motion were same.

3.7 Discussion and Conclusion

The task in Experiment 2 is more difficult than in Experiment 1. Over all display conditions error rate range is 37.1% to 58% compared to 25% to 38% in Experiment 1. In Experiment 2 chance guessing would be expected to yield an error rate of 77% while Experiment 1 would be 50%. This indicates even in the worse condition—no stereo, no motion--participants perform better than chance.

In both the prior report and our current analysis, there is no main effect on response time; however, we test and find no three-way nor two-way interactions which could have theoretically masked the main effect. Further, we note here that there is a non-significant trend for shorter mean response time with stereo and head-tracking. Possibly, a more statistically powerful experiment could find a small effect which is being masked by the depth ordering task’s relative difficulty. Another possibility is that with the NS-NH condition yielding an average 58% error rate participants are essentially giving up after 14s thus capping the response time. Perhaps a study that gives user feedback on the correctness of their answer and allows for a limited number of additional attempts to answer the depth ordering question correctly might find a larger variation in the response time required to obtain a correct answer.

3.7.1 Prior Work Comparisons

Ware et al. [155] compared display conditions' effect on a participant's ability to determine whether two nodes in a network are connected. The conditions are 2D rendering, 3D rendering (no motion, no stereo), stereo, passive rotation, stereo plus passive rotation, hand controlled rotation, stereo plus hand controlled rotation, head-coupled motion parallax, stereo plus head-coupled motion parallax. The network had 75 nodes and 100 arcs. The results confirmed stereo plus motion is the most effective and show that which method is used for producing motion is not particularly important.

While the 3D network data and our volume data clearly differ, some useful comparisons can be drawn. Our Experiment 1 and Experiment 2 are consistent with network study's finding that stereo plus motion is most effective. However, in the network experiment motion alone showed a greater advantage than stereo alone. In contrast, in our experiments, the motion alone conditions did not demonstrate significant improvement over the stereo alone conditions. For Experiment 1 this might be explained by our shorter exposure time (2s) compared to the network experiment where user response time varied from 5 to 15 seconds depending on the node count. With a shorter exposure time, there is simply less time to gather structure-from-motion cues. However, in our Experiment 2 the average task time is a similar 14.4 seconds. Here motion cues can perhaps become more useful. However, again the motion only condition did not exhibit significant improvement for depth ordering accuracy. Overall in both our volume data experiments, the stereo conditions had more significant pair-wise comparisons which suggest that for volumetric data, unlike for 3D network data, stereo may be more significant than motion while stereo plus motion still yields the best accuracy.

Ware et al. [161] repeat a modified version of their network experiment comparing stereo, the kinect-depth-effect, stereo with the kinect-depth-effect, and plain 3D (no stereo or motion) for viewing 3D graphs of varying sizes. A major difference is the use of a 3840 × 2400 Wheatstone stereoscope rather than the 1024 × 768 time-multiplexed display in the earlier study. This change caused the improvement due to stereo plus motion to be roughly an order magnitude, rather than merely the threefold improvement found earlier.

Their participants had up to 5 seconds to view each trial after which the screen went blank until the participant responded. Average response time ranged between 1.5 and 3 seconds. This a similar range to our depth discrimination task's limit of 2 seconds, but less than our depth ordering tasks average of 14.4 seconds. Note, our Experiment 1's kinect-depth-effect rotates through 20 degrees in 2 seconds. The 3D network experiment rotates 360° per 36 seconds, implying 30° for a 3 second view. In the high-res network experiment for inexperienced observers stereo was the most useful cue, while for the experienced observers (the experimenters themselves) motion was the most useful (i.e. gave greater incremental improvement in accuracy). The former result of inexperienced subjects is inline with our Experiment 1 and 2 results where post-hoc comparisons show the stereo conditions' better performance to be statistically significant. We did not include ourselves in our experiments, but as we note earlier in Experiment 1, for the CS pool subjects (who had more gaming experience) head-tracking alone did show a significant effect, although it was not stronger than stereo. This suggests repeating our study with highly experience observes to see whether motion proves a stronger cue.

The volume depth discrimination’s average range of error, 25% to 37%, across display condition is similar to the range of error found in the largest tested 3D networks (1000 nodes), but is beyond the range in the 33 node network, roughly 5% to 15% error. Given the similar error rate range and trial duration, this suggests a similar level of task difficulty between the 1000 node task the volumetric depth discrimination task. The volume depth ordering task appears even harder given its 37.1% to 58% error rates.

Our depth ordering task took significantly longer per trial (average 14.4s) than the high-res 3D network task (1.5s to 3s) but it still improves in accuracy with stereo. In the 3D network task, display condition does significantly affect response time, with the stereo group performing 15% faster, but in our depth ordering task response time did not improve. This further suggests the volume depth ordering task is more difficult and possibly the increased difficulty swamp any improvement due to display condition. Another possible explanation is subjects are essentially giving up in the worst case condition (no stereo, no motion). This hypothesis could be further explored as discussed in earlier of this section.

Finally, the large improvement that the two 3D network studies found when going from a 1024 × 768 time-multiplexed stereo display to a 3840 × 2400 Wheatstone display (sans cross-talk) strongly suggests one may find similar greater enhancements in volumetric depth tasks for the stereo plus motion case with higher resolution systems.

3.7.2 Qualitative Results

Some participants in our study noted that the auto-rotation condition that approximately simulated the head-tracking condition was not really the same as head-tracking, because the auto-rotation only rotated about the vertical axis. This roughly corresponds to side-to-side head motion, but does not mimic up-down head motion. In the post-questionnaire, most participants state that the background noise image does not make the depth judgement difficult. However, they frequently comment that the individual cylinders' textures did affect their depth judgement. In particular, when the overlapping portions of two cylinders happened to be a brighter texture region, is makes depth judgement is easier. Typically, participants report that neither cylinder size nor the presence of the black window (which hid the cylinders' ends) affects their depth judgement. Further based on the post-questionnaire, participants' confidence in the accuracy of their depth judgements is the highest for stereo without head-tracking; the third highest for head-tracking alone; and the lowest for non-stereo with head-tracking. Interestingly 84% of participants answered that head-tracking gives better depth perception than stereo in the Experiment 1 but in Experiment 2 only 56% of participants answered this way. Yet, the quantitative results suggest stereo is more important for accuracy.

3.7.3 Conclusion

In this chapter, we examine the effect of stereoscopy and structure-from-motion on depth discrimination and depth ordering tasks for a volumetric dataset. The stereo plus motion condition is the most effective in both experiments. And we found that stereoscopy by itself improves depth perception in a depth ordering task. In the depth discrimination task, head-tracking by itself helps depth judgement for our CS-pool participants who report playing more computer games. However, stereoscopy alone does not aid depth discrimination and head-tracking alone does not benefit participants overall in either experiment.

The work of Ware and Mitchell suggest stereoscopy’s enhancement of depth perception of volumetric data may be even greater for display resolutions approaching that of the human eye. This is open to further experimentation.

A challenging area is evaluating how display conditions and volumetric software rendering parameters interact to effect perception when both are varied. There are a large number of potential independent variables and interactions to evaluate. For all such studies, our results indicate we should develop a more robust pre-experiment questionnaire that would allow separating participants into groups based on degree of experience with gaming, VR type technologies, and expertise in viewing stereoscopic volume data. Future work should include a range of more elaborately generated volumetric datasets to better mimic real-world data sets while at the same time providing randomly varying volumetric structures to avoid spurious learning effects across repeated trials.

CHAPTER 4: BIMANUAL 7DOF OBJECT MANIPULATION

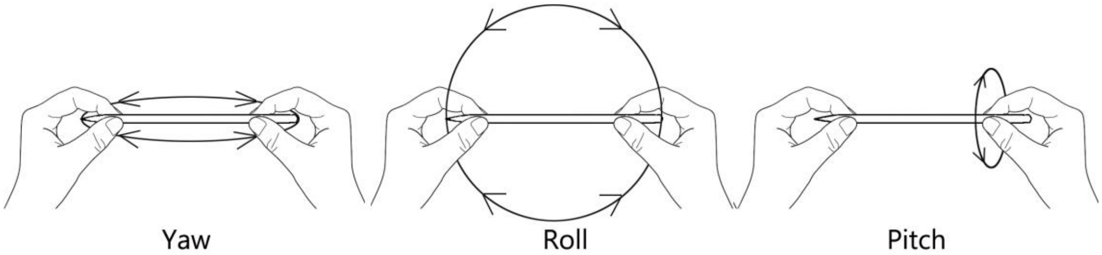
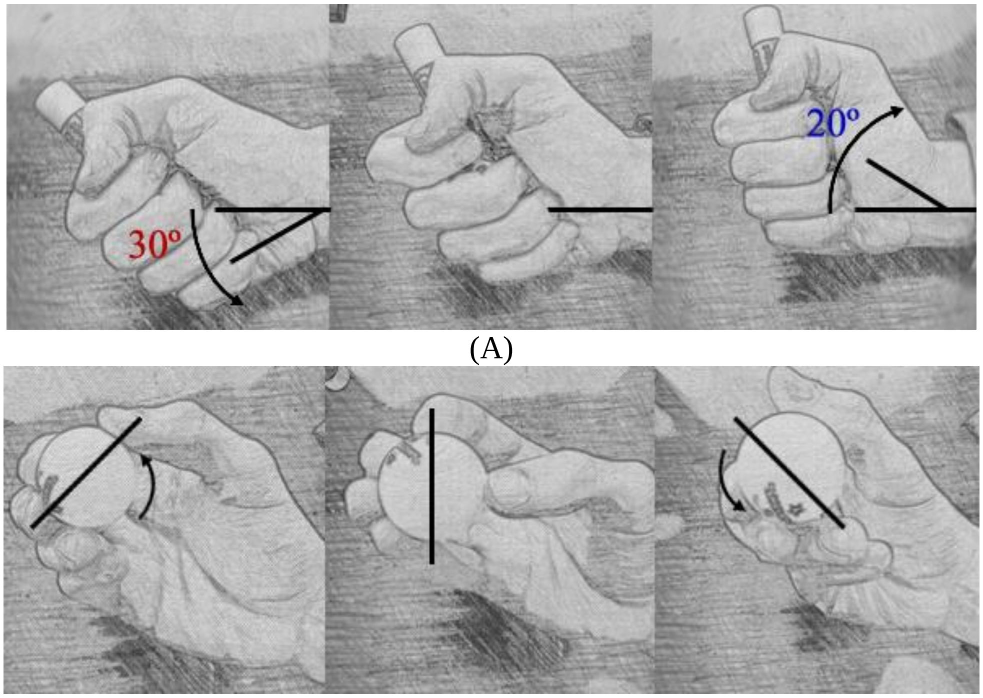


Figure 10: Real world two-handed rotation

Figure 11: Button ball input

4.1 Introduction

Previous two-handed interaction techniques (ITs) are based on the work of Mapes and Moshell [92]. They present an IT using 6DOF tracked pinch gloves for adjusting the scale and pose (position + orientation) of an object. The IT is engaged with a pinch gesture. Two 3D cursors are displayed corresponding to the user’s hands. Translating the hands rigidly translates the target object. Rotating the hands relative to one-another rotates an invisible axis between the two cursors adjusting the target object’s orientation. Expanding or contracting the distance between the hands scales the target up or down. The commercial product SmartScene evolved from this work and includes this IT (among others). More, recently Schultheis et al. [127] add a visual representation of the axis (the "Spindle") drawn between cursors with a small sphere indicating the center point. They found this improved the user's understanding of the IT. This Spindle IT is implemented using power-grasped joystick handles called SpaceGrips^{TM}. In prior work, we developed various two-handed 6DOF user interfaces (UIs) using precision grasped buttons balls (Figure 11) combining the designs of Zhai's FingerBall [177] and Shaw and Green's [129] pair of button enhanced bats [157]. Ulinski et al. use button balls to evaluate techniques for manipulating the 9DOFs of a box used for selecting volume data [146,145]. We also chose this smaller form-factor button ball to maintain a precision-grasp during interaction.



(A)

(B)

Figure 12: (A) A power-grasp, and (B) a precision-grasp.

In this chapter we present the Spindle+Wheel, an IT that extends the Spindle IT by using precision-grasped button balls instead of power-grasped joystick handles. The user can twist or roll a button ball within her fingers independent of any relative hand translation. This allows immediate control of the rotation around the spindle axis or pitch (the rightmost picture in Figure 10). Power-grasped handles do not strongly afford this physical manipulation, nor do pinch gloves. With these two earlier input devices the user can only perform 20˚ radial deviation and 30˚ ulnar deviation (side bending toward the thumb or little finger (Figure 12A)). This is not particularly comfortable and likely for this reason prior two-hand ITs do not support rotation around the spindle axis. In contrast, when holding a small ball, the fingers can freely rotate the ball in either direction and do so continuously with physical clutching (Figure 12B). While our current button balls are chorded which does restrict the rotation, wireless technologies have been demonstrated and would not have this restriction. Although current wireless 6DOF technologies do not fit our desired small ball form factor, we anticipate their future availability.

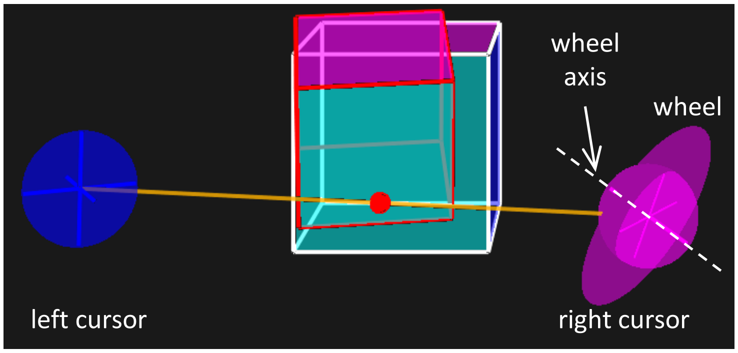


Figure 13: Spindle+Wheel visual.

The Spindle+Wheel IT includes a "*Wheel*" visual representing the additional pitch DOF (Figure 13). We conduct two user studies with twelve participants each for a scaled docking task that is implemented as a 7DOF multi-scale travel technique using the scene-in-hand metaphor. In Experiment 1, we compare Spindle+Wheel and Spindle only IT conditions. In Experiment 2, we compare Spindle+Wheel, Spindle+Wheel with separate scale and one-handed 6DOF+scale [176] IT conditions. Both experiments demonstrate that the Spindle+Wheel IT has better performance than other ITs on completion times and button clicks for the 7DOF docking task.

4.1.1 Spindle+Wheel interaction technique

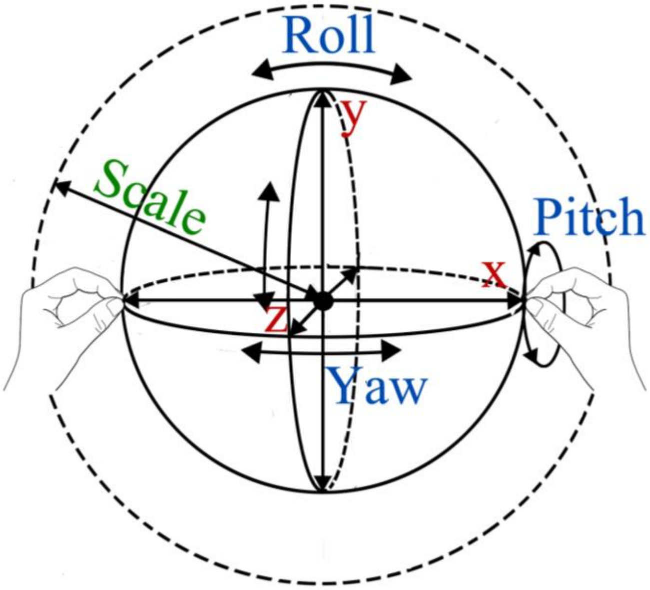


Figure 14: Illustration of 7DOF of the Spindle+Wheel IT.

The Spindle+Wheel IT mode works as follows. In the following description we assume that the user is right handed. For a left-handed user the roles of the left and right button balls would be reversed. Two 3D sphere cursors are shown corresponding to the button balls, but at a comfortable translational offset [129] for our stereoscopic, desktop VR setup [40]. Based on Schultheis et al. [127], a cylinder is drawn between the cursors with a red sphere at the center point (Figure 13). Pressing and holding a button on the left button ball engages the travel technique. As in Mapes and Moshell, translating the hands rigidly translates the view point. Moving one hand individually, such as rotating one hand about the other—while keeping their distance constant—rotates the view in yaw and roll (Figure 10). Moving the hands closer or farther apart scales the view. The precise center of scale (or rotation) depends on the hand motion. For instance, holding the left hand still and moving the right hand inward or outward scales about the left cursor; whereas, moving both hands an equal distance inward or outward scales about the spindle center.

The Spindle+Wheel Twist: Spinning or twisting the dominant button ball with the fingers around the axis of the wheel rotates the view around the spindle axis (i.e. pitches, see rightmost figure in Figure 10). Note that the user need not have the wheel aligned perpendicular to the spindle axis to perform this maneuver (Figure 13). We chose this design based on initial informal evaluation and pilot testing. When the IT is first engaged the wheel axis is reset parallel to the spindle axis and then the wheel axis orientation remains fixed relative to the right button ball orientation. During typical yaw and roll maneuvers (Figure 10) the orientation of the wheel axis will deviate to varying degrees from the spindle axis (Figure 13). Note, in Figure 13 the wheel axis is just shown for illustration and not actually displayed by the IT. In our experience the visual feedback of the wheel makes it easy to pitch the view regardless of how one has oriented the spindle during a maneuver. Figure 14 illustrates how all 7DOF works simultaneously.

During pilot studies, we observed some users accidentally changing the view scale factor while they were manipulating the view position. If a travel task doesn’t require any view scale changes, this may reduce user performance. Hence we added Spindle+Wheel with separate scale IT (simultaneous 6DOF pose + scale) in Experiment 2 in order to evaluate effects of accidental scale changes to user’s ability on task accomplishment.

4.2 Experiment Design

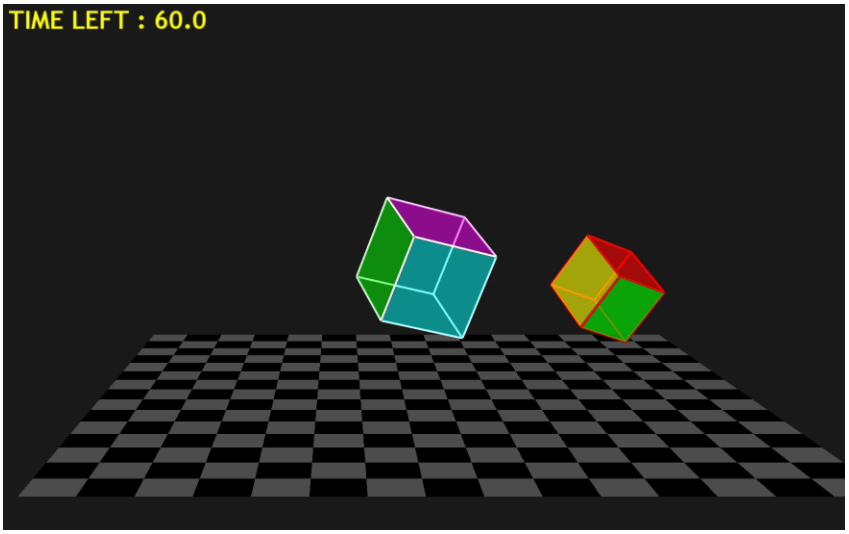


Figure 15: Screen capture of virtual environment displayed on desktop VR in Experiment.

The VE has a checker-board ground-plane (Figure 15). It is 40 cm square with half appearing behind the display surface and half appearing in front. In the center of the screen is a transparent box of fixed size (white-outlined box) and at a random orientation per trial. Each face is a different color. This cube's pose remains stationary relative to the display screen during travel and is called the *objective* cube. At each trial, a second *target* cube appears at random location on the ground-plane (red-outlined box). This *target* cube's location, size and orientation vary randomly across trials. A timer appears in the upper right of the screen. The user must travel to align the target cube with the objective cube. This requires view pose and view scale maneuvering to match the cube sizes. The size of the target cube varied across 3 sizes (25%, 100% and 400% of the objective cube's size.)

When the distance between the target cube's corresponding vertices is within a tolerance (0.84 cm) of the objective cube's vertices, the outline of the target cube turns green and a success sound is played. The user must release the IT engagement button to stop the clock. The user then presses a button to advance to the next trial. The user uses the dominant hand for the IT engagement button and the less-dominant hand for a pitch control. Pilot tests indicated 60 seconds was sufficient for this task.

The system is a stereoscopic, desktop VR setup [40] using Nvidia 3D Vision glasses and a 120Hz LCD 22" monitor. The tracking system is a Polhemus Fastrak. The user sits with his torso roughly 1 meter from the display. Software is written in OpenSceneGraph [105], VRPN [141], OpenAL [104] and OSGVE [138]. In all IT conditions of both experiments, the participant uses the button balls (Figure 11). At the start of a session the user is asked to hold the button balls and rest their elbows on the chair's arms and the experimenter sets a translational offset [129] that places the 3D cursors in the center of the screen.

The 3D cursors and spindle are mapped using absolute position control. However, travel is only engaged when a button is pressed. Based on the ground-plane size and a typical person's arm reach, a user will typically perform several clutching translation maneuvers to reach the far side of the ground-plane. Because we are comparing two position control ITs (rather than rate control), the target box distance range is restricted to the range of the ground plane. In a proper (large) multi-scale virtual environment larger ranges of travel would occur inducing either a larger number of translation clutches or strategic use of view scale and translation maneuvers. Since our goal is to compare two position control ITs within the domain of multi-scale environments, we keep the distance short but use the variation in target box size to force the user to change view scale.

4.2.1 Experiment 1: A Comparison of Spindle and Spindle + Wheel Interaction Techniques

Twelve users performed 90 trials (15 trials × 3 box sizes × 2 IT conditions) each in a within-subject comparison of the Spindle IT vs. the Spindle+Wheel IT for a 7DOF docking task. The two IT conditions are presented in counter-balanced order across the twelve participants. All participants are from Computer Science Department, ten are males and two are females (eight Ph.D. students, one master, and three undergraduate students). All participants have (corrected) 20/20 eye vision and no disability using their arms and fingers. All participants have high daily computer usage (6.3 out of 7) and eleven of them have experience of 3D UI typically using the Microsoft Kinect or Wiimote controller.

4.2.2 Experiment 2: A Comparison of One-handed, Spindle + Wheel with Separate Scale Control and Spindle + Wheel Interaction Techniques

Twelve users performed 90 trials (10 trials × 3 box size × 3 IT condition) each in a within-subject comparison of three conditions (one-handed IT, Spindle+Wheel with separate scale IT and Spindle+Wheel IT). The three IT conditions are presented in counter-balanced order across the twelve participants. In Spindle+Wheel with separate scale IT, one button engaged 7DOF manipulation, while a second one only engages 6DOF manipulation. In the one-handed IT, one button engages 6DOF manipulation using the scene-in-hand metaphor. A second engages rate controlled scale. A third engages position controlled scale. In both cases the center of scale is determined by the button ball location when the scale button is first pressed [122].

Eight participants are from the Psychology Department Pool and four are from the Computer Science Department, four are males and eight are females (one Ph.D. student and eleven undergraduate students). All participants have (corrected) 20/20 or higher eye vision and no disability using their arms and fingers. All participants have high daily computer usage (6.08 out of 7) and four of them have experience with 3D UI. Three of them had experienced motion sickness before but they were able to finish the experiment without motion sickness.

4.2.3 Data Analysis

We recorded task completion time and the number of button presses per trial. The button press counts indicate the number of clutching maneuvers performed per trial.

We carefully checked distributions of task completion time for each participant. We found that the mean is skewed; further the per-trial sample is too small to trim outliers. Therefore, we use the per-trial median of task completion time for further analysis. For number of button clicks, we use the per-trial mean. The reported F tests use α=.05 for significance and indicate the Geisser-Greenhouse correction to protect against possible violation of the sphericity assumption. The post-hoc tests that were conducted were Fisher's least significant differences (LSD) pairwise comparisons with α=0.05 level for significance.

4.3 Experiment 1: A Comparison of Spindle and Spindle+Wheel Interaction Techniques

We used a 2 × 3 repeated measures (*rm*) ANOVA and used IT presentation order as the between-subjects factor. IT condition (value set {*Spindle+Wheel (S+W), Spindle only (SO)*}), and target box size (value set {25%, 100%, 400%}) are the two variables manipulated within participants. The primary hypotheses are:

H1.Spindle+Wheel is expected to have faster completion times than Spindle Only.

H2.*Spindle+Wheel is expected to incur fewer buttons clicks than Spindle Only.*

4.3.1 Result

Quantitative

Completion Time

Table 3: Average completion time and average number of button presses with standard deviations of Experiment 1

	CT (s)	SD	BC	SD
Spindle (25%)	26.65	4.57	6.71	1.77
Spindle (100%)	22.59	4.69	5.88	1.41
Spindle (400%)	20.75	4.23	4.42	1.61
Spindle+Wheel (25%)	19.05	3.13	4.48	1.74
Spindle+Wheel (100%)	18.71	4.90	4.41	2.0
Spindle+Wheel (400%)	16.93	3.91	3.88	1.68

Table 3 shows average completion time and average button clicks of IT condition and box size. The three-way *rm* ANOVA (Order × IT × Size) on completion times shows no significant interaction effect of IT representation order on completion time (*p*=.312) or button clicks (*p*=.418).

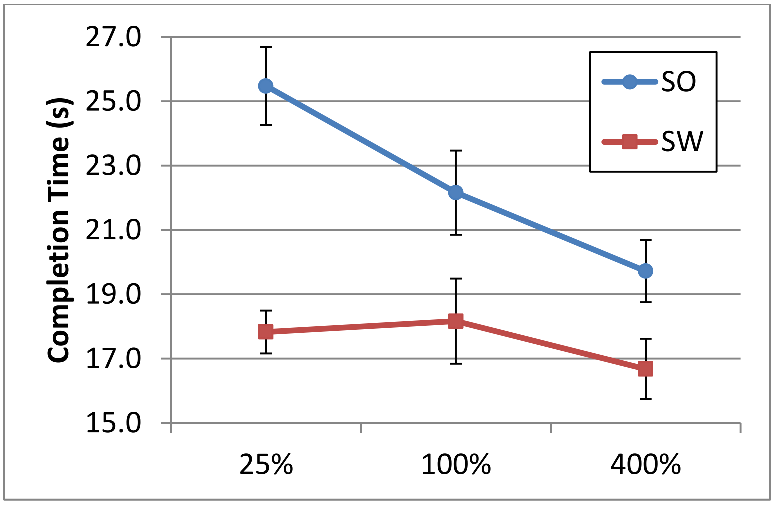


Figure 16: Completion time of IT condition and box size of Experiment 1.

The 3 × 2 *rm* ANOVA (Box Size × IT) shows an interaction effect of box size and IT condition on completion time (F(2,22)=4.579, p=.022, η_p^2=.294). This indicates that box size had different effects on task completion time depending on which IT condition was used. Contrasts reveal that there is an interaction effect when comparing 25% to 400% box sizes for SO compared S+W (F(1,11)=19.431, p=.008, η_p^2=.485) (see Figure 16).

The results show a simple main effect on task completion time of IT condition in 25% box size (F(1,11)=42.835, p<.001, η_p^2=.796), and 400% box size (F(1,11)=6.001, p=.032, η_p^2=353). Completion time of S+W is significantly faster than SO in both 25% and 400% box sizes. However, there is no simple main effect on completion time in 100% box size (F(1,11)=4.742, p=.052, η_p^2=.301). Completion time of S+W was slight better than SO but not statistically significant (100% in Figure 16).

The results show a main effect on task completion time of IT condition (F(1,11)=19.431, p=.001, η_p^2=.639). Overall, task completion time of S+W (M=17.55, SD=3.46) is significantly faster than SO (M=22.45, SD=4.61) (Hypothesis *H1*).

There is also a main effect of box size on completion time (F(2,22)=12.868, p<.001, η_p^2=.539). LSD tests show completion time of 400% box size (M=18.20, SD=3.59) is faster than 25% (M=21.65, SD=5.12, p<.001) and 100% box sizes (M=20.16, SD=4.9, p=.022). However, there is no significant difference between 25% and 100% (p=.070).

Button Clicks

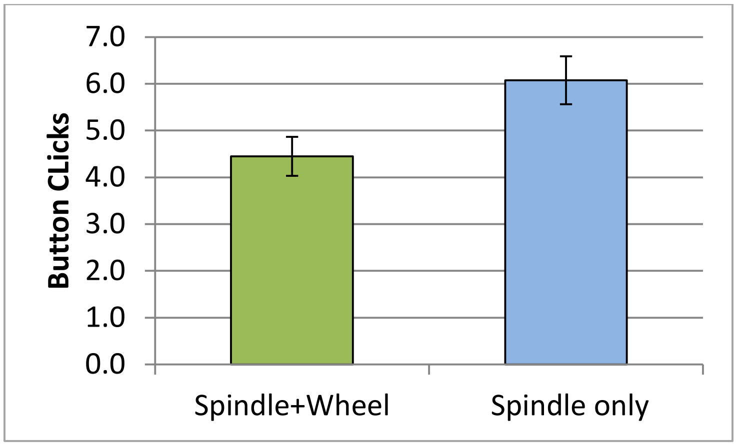


Figure 17: Number of button clicks of IT condition of Experiment 1.

Regarding a number of button clicks, there is no interaction effect of box size and IT condition (F(2,22)=1.915, *p*=.171, η_p^2=148). However, IT condition has a main effect on button clicks (F(1,11) = 43.593, *p*<.001, η_p^2=.799). Button clicks of S+W (M=4.4, SD=0.42) is significantly fewer than that of SO (M=6.1, SD=0.5) (see Figure 17). This is consistent with the faster performance of the S+W condition (Hypothesis *H2*). A strong plausible explanation is that because only the S+W allows for immediate control of pitch, users had to perform additional yaw-roll (Figure 13) manipulations using the SO IT rather than a single pitch maneuver with the S+W IT. This is consistent with participants' comments in post-survey questionnaires.

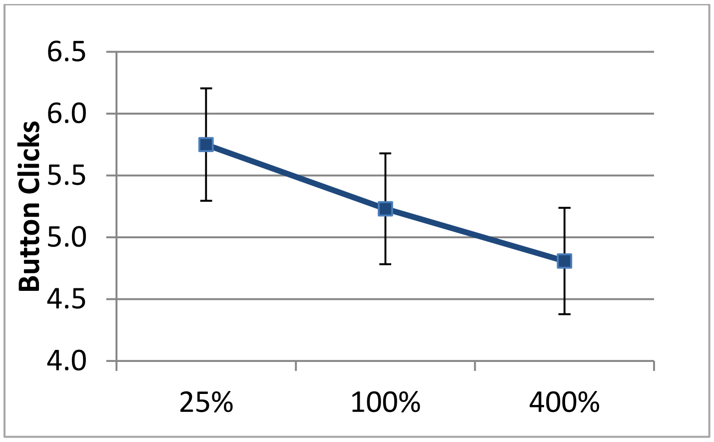


Figure 18: Number of button clicks of box size of Experiment 1.

The results also show a main effect of box size on button clicks (F(2,22)=14.179, p<.001, η_p^2=.563). LSD tests show button click of 100% box size (M=5.23, SD=1.82) is fewer than 25% box size (M=5.75, SD=2.02, p=.008). And button clicks of 400% box size (M=4.81, SD=1.77) is significantly fewer than 25% box size (p<.001). However, there is no significant difference between 100% and 400% box sizes (p=.071) (see Figure 18). A possible explanation is the 100% size box required no scale change. However, we did not see a similar effect between the 400% and 25% size boxes that requires scale changes. This maybe because the 25% box condition generally takes longer time than the other size conditions regardless of ITs.

Qualitative

Participants took a post-survey questionnaire regarding subjective preferences. All (twelve out of twelve) participants preferred the S+W IT over the Spindle IT. When asked whether the S+W is better than the SO for rotation, ten of twelve participants agreed; one rated the ITs equal and one preferred the SO IT. Nine of twelve participants indicated the S+W is more intuitive; two rated the SO IT and one rated both ITs equal.

On a 7-point Likert scale, regarding whether the S+W's pitch control was helpful (1 not at all through 7 very helpful), the average was 5.8. On the same scale rating regarding whether the wheel helped inform them of the rotation axis, the rating was 4.3. On a 7-point Likert scale, user rating of arm fatigue was not significantly different with 3.67 for SO and 3.08 for S+W (1 no fatigue through 7 very painful based on the NASA TLX).

4.4 Experiment 2: A Comparison of One-Handed, Spindle+Wheel with Separate Scale Control and Spindle+Wheel Interaction Technique

For Experiment 2, we use a 3 × 3 repeated measures (*rm*) ANOVA and use IT presentation order as the between-subjects factor. IT condition (value set {*One-handed*

(OH), Spindle+Wheel with separate scale (SWS), Spindle+Wheel (S+W)}), and target box size (value set {25%, 100%, 400%}) are the two variables manipulated within participants. In addition to task completion time and number of button clicks, we analyze two different scale control techniques (rate control and position control) for the OH IT. The primary hypotheses are:

H1. Spindle+Wheel is expected to have faster completion time than others ITs with 25% and 400% box sizes.

H2. Spindle+Wheel is expected to have slower completion times than others ITs with 100% box size.

H3. Spindle+Wheel is expected to incur fewer button clicks than others ITs with 25% and 400% box sizes.

H4. All IT conditions are expected to incur similar button clicks with 100% box size.

H5. In the One-Handed condition, a rate control IT is expected to be preferred to a position control IT.

4.4.1 Result

Quantitative

Completion Time

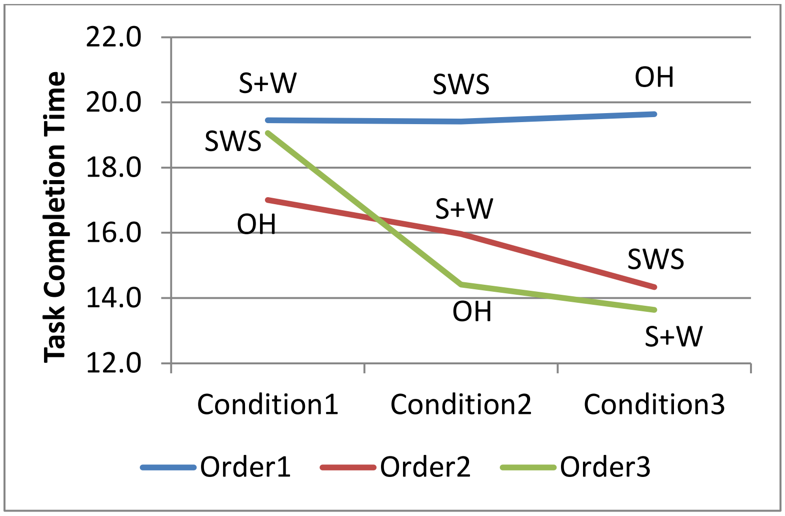


Figure 19: Effect on task completion time of IT condition and IT presentation order of Experiment 2.

The three-way *rm* ANOVA (Order × IT Condition × Box Size) is not significant. The two-way *rm* ANOVA (Order × IT condition) shows a two-way interaction on task completion time (F(4,18)=5.226, *p*=.006, η_p^2=.537). In Order 3, there is a significant effect of IT condition on completion time and SWS is significantly slower than both S+W (*p*=.001) and OH (*p*=.006) (green plot in Figure 19). However, for Order 1 and 2 there is no significant effect of IT condition (red and blue plots in Figure 19).

Table 4: Average and standard deviation of task completion time and number of button presses by IT condition and box size of Experiment 2.

	CT (s)	**SD_{CT}**	**BC**	**SD_{BC}**
OH (25%)	21.47	6.96	9.1	2.6
OH (100%)	10.62	2.91	3.8	1.5
OH (400%)	18.97	3.95	9.6	2.7
SWS (25%)	22.15	6.40	7.7	2.2
SWS (100%)	12.10	4.05	3.4	1.3
SWS (400%)	18.55	4.79	7.0	1.4
S+W (25%)	18.20	5.03	4.7	1.5
S+W (100%)	19.46	5.99	4.3	2.0
S+W (400%)	15.40	3.95	4.5	1.8

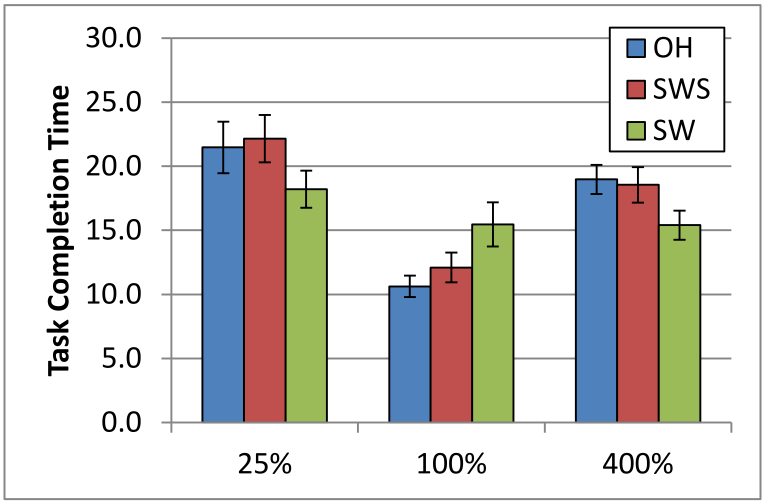


Figure 20: Completion time of IT condition and box size of Experiment 2.

Table 4 shows average completion time and button clicks of IT condition by box sizes. The results show an interaction effect on completion time (F(4,44)=7.924, *p*<.001, η_p^2=.419). Contrasts reveal significant interactions when comparing 25% to 100% both for OH compared S+W (F(1,11)=13.902, *p*=.003, η_p^2=.558) and SWS compared S+W (F(1,11)=17.639, *p*=.001, η_p^2=.573). These also reveal interactions between 100% to 400%

both for OH compared S+W (F(1,11)=18.962, *p*=.001, η_p^2=.633) and SWS compared S+W (F(1,11)=14.777, *p*=.003, η_p^2=.573). These show that completion time of OH and SWS ITs are faster than S+W with 100% but slower with 25% and 400% (see Figure 20). As expected (hypotheses *H1 and H2*), completion time of OH or SWS is significantly faster than S+W with 100%. The most likely explanation is that accidental scale changes with S+W decreases performance in a docking task that requires no scale change, i.e. the 100% box size.

There is a simple main effect on completion time of IT condition with 100% box size (F(2,22)=7.297, *p*=.004, η_p^2=.399). LSD tests show completion time of S+W is significantly slower than SWS (*p*=.036) and OH (*p*=.004) (100% in Figure 20). However, completion time between OH and SWS does not differ (*p*=.221). With 400% box size, there is also a simple effect of IT condition on completion time (F(2,22)=4.687, *p*=.020, η_p^2=.299). LSD comparisons show completion time of S+W is significantly faster than SWS (*p*=.034) and OH (*p*=.018) (400% in Figure 20). Completion time between OH and SWS does not differ (*p*=.736). Unexpectedly, there is no significant simple main effect on completion time of IT condition with 25% box size (*p*=.484) (25% in Figure 20). As we discuss in the previous paragraph, 100% box size does not require any scale changes. OH and SWS have better performance than S+W with 100% box size. Possibly this changes the overall significance of performance of IT conditions overall three box sizes.

Because of this, we break box sizes down into two groups based on whether a scale change is required and analyze these groups separately. The 3 × 2 *rm* ANOVA (IT condition × {25%, 400%}) shows no interaction effect on completion time of IT condition and box size. However, as expected, there is a main effect of IT condition on completion time (F(2,22)=5.497, *p*=.012, η_p^2=.333). LSD tests show completion time of S+W (M=16.8, SD=4.65) is significantly faster than both SWS (M=20.35, SD=5.83, *p*=.019) and OH (M=20.22, SD=5.68, *p*=.009). This clarifies advantages of S+W when a task requires a scale change.

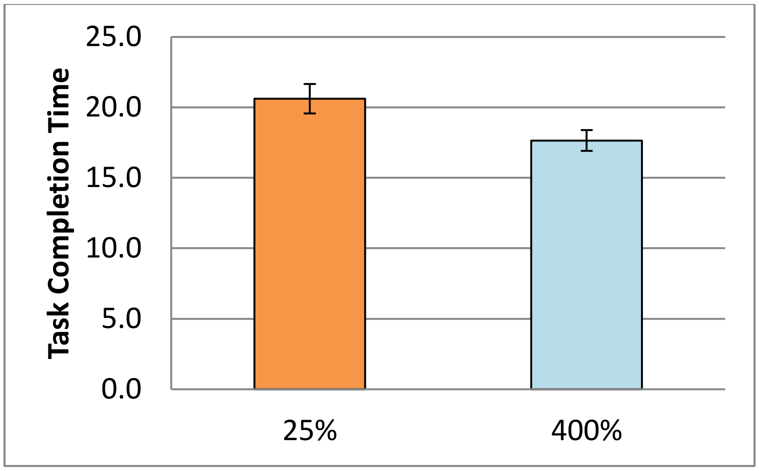


Figure 21: Completion time of 25% and 400% box sizes of Experiment 2.

The results also show a main effect on completion time of box size (F(1,11)=14.598, *p*=.003, η_p^2=.570). Completion time with 400% (M=17.55, SD=0.74) is significantly faster than 25% box size (M=20.61, SD=1.04) (see Figure 21). Interestingly, both box sizes require scale changes. However, it seems scaling down is faster than scaling up. We suggest this is because of differences in arm movements for scaling. As intended, users tend to use their forearms with elbows resting when scaling with S+W. This minimizes shoulder flexion, extension, and adduction and abduction. However, the user may feel more comfortable to bring the hands together from a neutral posture, technically a "medial shoulder rotation", compared to separating the hands, a "lateral shoulder rotation."

Button Clicks

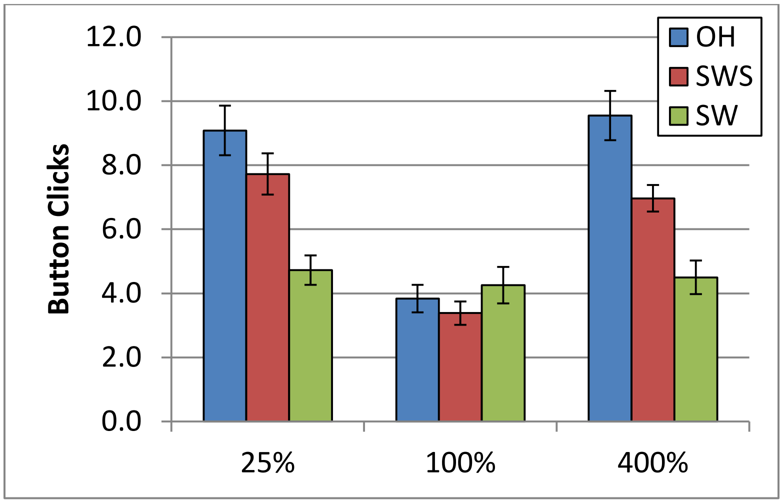


Figure 22: Button clicks of box size and IT condition of Experiment 2.

Regarding button clicks, the 3 × 3 *rm* ANOVA shows an interaction effect of box size (25%, 100%, and 400%) and IT condition (F(4,44)=28.893, $p<.001$, $\eta_p^2=.712$). This indicates that box size has different effects on button clicks depending on IT condition has different effects on button clicks depending on which box size is examined. In Figure 22, for 100% size, IT condition makes little difference, but for the 25% and 400% IT appears to make a larger difference. Contrasts reveal significant interactions when comparing 25% to 100% box sizes both for OH compared S+W (F(1,11)=59.108, $p<.001$, $\eta_p^2=.843$) and SWS compared S+W (F(1,11)=40.967, $p<.001$, $\eta_p^2=.788$). Similarly, for the 100% to 400% box sizes, the S+W to SWS relation differs (F(1,11)=79.209, $p<.001$, $\eta_p^2=.878$) as does the S+W to OH relation (F(1,11)=59.108, $p<.001$, $\eta_p^2=.861$). Further, the relation of OH to SWS also changes between 100% and 400% sizes (F(1,11)=23.047, $p=.001$,$\eta_p^2=.677$). The results are not surprising, since in the scaling conditions (400% and 25%), the OH and SWS condition requires using two buttons (pose and scale) while S+W uses one (pose simultaneous with scale).

In more detail, for the 25% box size, there is a simple effect on button click of IT condition (F(2,22)=28.244, $p<.001$, $\eta_p^2=.720$). Contrasts reveal that button click of S+W is significantly fewer than both OH (F(1,11)=34.532, $p<.001$, $\eta_p^2=.758$) and SWS (F(1,11)=43.612, $p<.001$, $\eta_p^2=.799$) (25% in Figure 22). In addition, these also show that button click of SWS is significantly fewer than OH (F(1,11)=6.149, $p=.031$, $\eta_p^2=.359$). For the 100% box size, there is a simple effect on button click of IT condition (F(2,22)=2.298, $p=.044$, $\eta_p^2=.247$). Contrasts reveal that button click of S+W is significantly fewer than OH (F(1,11)=7.58, $p=.019$, $\eta_p^2=.408$). However, there are no significant differences on button clicks between SWS and OH ($p=.132$) and between SWS and S+W ($p=.283$) (100% in Figure 22). For the 400% box size, there is a simple effect on button click (F(2,22)=30.900, $p<.001$, $\eta_p^2=.737$). Contrasts reveal that button click of SWS is significantly fewer than OH (F(1,11)=20.440, $p=.001$, ηp2=.650) and button clicks of S+W is significantly fewer than both OH (F(1,11)=39.745, $p<.001$, $\eta_p^2=.783$) and SWS (F(1,11)=36.331, $p<.001$, $\eta_p^2=.768$) (400% in Figure 22).

While the interaction, IT condition × box size, is significant, the main effects are also significant. There is a main effect on button clicks of IT condition (F(2,22)=33.384, $p<.001$, $\eta_p^2=.752$). LSD post-hoc tests show that button clicks of S+W (M=5.3, SD=1.8) is significantly fewer than OH (M=7.5, SD=3.5, $p<.001$) and SWS (M=6.0, SD=2.5, $p<.001$). In addition, button click of SWS is significantly fewer than OH ($p=.001$). A good explanation is S+W can change the scale while translating and rotating, and this requires fewer button clicks than the other ITs for the 25% and 400% boxes.

There is also a main effect on button clicks of box sizes (F(2,22)=116.572, $p<.001$, $\eta_p^2=.914$). Number of button clicks of 100% box size (M=3.8, SD=1.6) is significantly fewer than 25% (M=7.2, SD=2.8, $p<.001$) and 400% (M=7.0, SD=2.9, $p<.001$).

However there is no difference on button click between 25% and 400% box sizes (*p*=.609). Intuitively, 25% and 400% box sizes incur more button clicks than 100% box size because they require scale changes.

Rate Control vs. Position Control

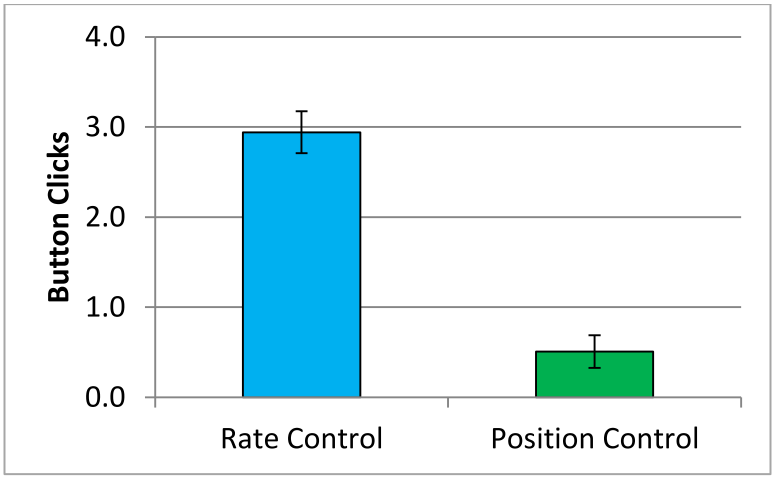


Figure 23: Number of button clicks of rate control and position control of Experiment 2.

Recall, the OH IT has two scale buttons: one for rate control and one for (relative) position control. A 2 × 2 *rm* ANOVA (scale technique *{rate control, position control}* × box sizes *{25%, 400%}*) shows a main effect of scale techniques on number of clicks (F(1,11)=39.701, *p*<001, η_p^2=.783) (see Figure 23). Number of clicks for the rate control (M=2.94, SD=0.81) is significantly more than the position control (M=0.51, SD=0.63). However, there is no interaction effect of scale technique by box size, and no main effect on number of clicks of box size. This indicates users tend to use the rate control more than the position control. This data alone cannot determine whether this occurs because users prefer the rate control, or because the rate control requires more frequent adjustment (hence button clicks) due to overshooting the desired scale; however, a user survey (see next section) indicates the former.

Qualitative

Participant took a post-survey questionnaire regarding subjective preferences. Six of twelve participants preferred the S+W; four rated the OH, one rated the SWS IT, and one rated both OH and SWS. When asked whether the S+W is better than the OH for rotation, six or twelve participants agreed. For translation, six of twelve participants answered that S+W is better than OH. Seven of twelve participants indicated S+W is better than the OH for scaling; three rated the SWS, and two rated the OH. Six of twelve participants answered the OH is the most intuitive; three answered S+W, and three rated SWS. On a 7-point Likert scale, regarding whether the S+W pitch control was helpful (1 *not at all* through 7 *very helpful*), the average was 5.33. On the same scale rating regarding whether the wheel helped inform them of the rotation axis, the rating was 3.67. On the same scale rating regarding whether separate scale was more helpful than integrated scale, the average was 4.25.

For a scale control technique for the OH IT, eight of twelve participants answered they preferred rate control; three rated position control, and one rated no preference.

4.5 Discussion and Conclusion

4.5.1 General Discussion

Experiment 1 demonstrates that the Spindle+Wheel IT performs faster overall then the Spindle IT (M=17.55s vs. 22.45s). There is a significant interaction with the target box size, however, where the improvement is only significant for the 25% and 400% box size conditions. In the 100% box size condition, users perform slower than in the other sizes regardless of the IT used. The number of clutches (button clicks) is significantly less for Spindle+Wheel compared to Spindle (4.43 vs. 6.08). In subjective questionnaires all twelve users prefer the Spindle+Wheel method overall.

Experiment 2 shows that the Spindle+Wheel performs faster for scale change tasks (25% and 400%) than one-handed IT and Spindle+Wheel IT with separate scale (M=16.8s vs. 20.2s vs. 20.4s) but performs slower for the 100% condition (M=10.6s vs. 12.1s vs. 19.5s). The number of clutches is significantly fewer for Spindle+Wheel compared to one-handed and Spindle+Wheel with separate scale (5.3 vs. 7.5 vs. 6.0). In subjective questionnaires six of twelve participants prefer the Spindle+Wheel IT and four prefer the one-handed IT.

Overall, the differences in results for the 25% and 400% versus 100% sizes are consistent with prior work on lesser 6DOF manipulations that find that when a task requires fewer than 6DOF adjustments adding either physical constraints or virtual geometric constraints to an otherwise 6DOF UI can improve performance by constraining the DOFs that the user does not desire to change.

Our results are more nuanced than Schultheis et al. [127] in two respects. Firstly, our results distinguish the scaling trials from non-scaling trials and secondly, the one-handed IT used in Experiment 2 is a closer one-handed counterpart to our experiment’s two-handed ITs. The one-handed IT used in Schultheis et al. is a Wanda device and their UI uses both the embedded tracker’s 6DOFs as well as the Wanda’s track ball. In contrast our one-handed condition uses the exact same device as our two-handed conditions. Overall, our experiments complement their results while at the same time we demonstrate advantages for the “+Wheel” addition.

Nowadays, many researchers focus on free hand 3D gesture rather than tracked held-devices. However, to our knowledge, there has been no free hand bimanual IT that supports simultaneous 7DOFs. The work of Song et al. [133] provides 7DOF control in a certain sense, but similar to the original Spindle technique, their work appears to only simultaneously support 6DOF (x,y,z + yaw,roll + scale) without simultaneous pitch control. To our knowledge, none of the unencumbered hand/finger tracking systems is free from occlusion problems. A free-hand implementation of our S+W IT requires robust and precise finger and hand tracking, and it has proven difficult to implement robustly.

We implemented our S+W IT with Leap motion controller [101] which provides more precise finger tracking data. We use two hands position for scale, translation, and orientation (yaw + roll) controls. For the pitch control, we use the index finger and thumb positions. A computed line between the index finger and thumb acts as the diameter of the wheel of the S+W IT. In pilot tests of the free-hand S+W IT, we found occlusion problems and a comparatively high occurrence of loss of tracking. This makes a comparative study to the button ball version difficult.

4.5.2 Conclusion

This chapter presents a novel 7DOF interaction technique, Spindle+Wheel. Our experiments show a statistically significant advantage of our Spindle+Wheel manipulation technique for 7DOF manipulation on both completion time and number of button clicks when scale changes are required. However, if the task requires only 6DOF (i.e. no scale changes) then Spindle+Wheel interaction technique has worse performance than other interaction techniques with separate scale control.

CHAPTER 5: DYNAMIC ADJUSTMENT OF STEREO VIEW PARAMETERS FOR A MULTI-SCALE VIRTUAL ENVIRONMENT

5.1 Introduction

Stereoscopic head-coupled display can enhance depth perception of the user in a computer generated 3D world [69]. However, sometimes the user sees two separate 2D images rather than a solid stereo 3D image or may experience eye strain and headaches. Stereo fusion problems increase simulator sickness especially with a head-coupled display. Many factors influence stereo fusion problems, but typically these translate into a range of distance in front of and behind the screen where a stereo 3D image can be comfortably fused.

Fusion problems are particularly problematic in a multi-scale virtual environment (MSVE) which is a virtual environment (VE) that contains geometric details whose sizes cover several orders of magnitude. The interaction (stereo × MSVE) occurs because viewing the small details in the MSVE often requires scaling up the world to the point where the rest of the VE geometry extends far behind and in front of the display screen. In non-MSVE environments whose geometry has a simpler geometric distribution [162] stereo adjustment techniques are relatively easier.

A traditional VE usually requires 6 degree-of-freedom (DOF) view control for 3D interaction techniques (IT) such as selection, manipulation and travel. In MSVEs, however, when a 3D user interface (UI) supports direct 3D manipulation, stereo or head-coupled display, the 3D UI benefits from an additional a view scale factor in the view model [122,123]. Proper choice of view scale--and often its dynamic adjustment--is important for reachability during direct manipulation, for maximizing effective stereopsis and for optimizing head-coupled structure-from-motion cues. Adding the 7th DOF, however, can complicate user navigation and also increases the chance for novice users to produce imagery with stereoscopic fusion problems (for example by abrupt manual enlargement of the scale factor).

This chapter evaluates the effect of three different stereo auto-adjustment conditions on a dual stage, multi-scale travel task using a one-handed scene-in-hand [153] travel technique. One adjustment condition is an auto-scale adjustment. Ware et al. [165] introduces this as cyclopean scale with the scale's center between the eyes. A question is whether auto-scale adjustment to control fusion problems will interfere with a user's MSVE travel task when it requires her to reach a particular view scale--not just a particular view pose. Two possibilities are:

P1) Auto-adjusting view scale may help because novice users may find purely manual control of 7DOF travel difficult and automation might reduce the difficulty.

P2) Auto-scale adjusting might hurt by tending to set the view scale to a scale other than the one the user desires.

This question appears to have not been empirically evaluated.

In addition to comparing auto-scale versus no auto-adjustment, we include a third condition an auto-translation based on a modification of Wartell et al. [165]. This condition is included because while it does perform some auto-adjustment (which might reduce novice user's difficulty compared to purely manual 7DOF travel), it does not alter the view scale, possibly avoiding interfering with the user control of scale.

Finally, our experiment uses an *extensive* MSVE, one whose database requires out-of-core paging (see Figure 30 and Figure 31). Such environments have zoomable geometry throughout the VE rather than just at a few select locations. Prior authors often use the latter to demonstrate multi-scale travel techniques to avoid having to implement or leverage an out-of-core renderer. Some of the qualitative observations of our experiment appear to only arise in extensive MSVEs where a larger variety of geometric viewing situations arise during the many trials of a formal evaluation.

Our results show benefits of the adjustments for task completion time and for reducing fusion problems, but only in certain combinations of display and task stage conditions. Further, the auto-adjustments were only beneficial when working at a certain range of target view scale, during the first stage zoom-in. Our use of an extensive MSVE also reveals view configuration examples that require display system specific modifications and demonstrate a further need for more sophisticated adjustment rules for extensive MSVEs.

5.2 Auto-Adjustment Technique

Table 5: True and modeled eye separations (E.S.) in physical and virtual coordinates.

	Physical	Virtual
True E.S.	6.0 cm	60 km
Modeled E.S.	3.0 cm	30 km

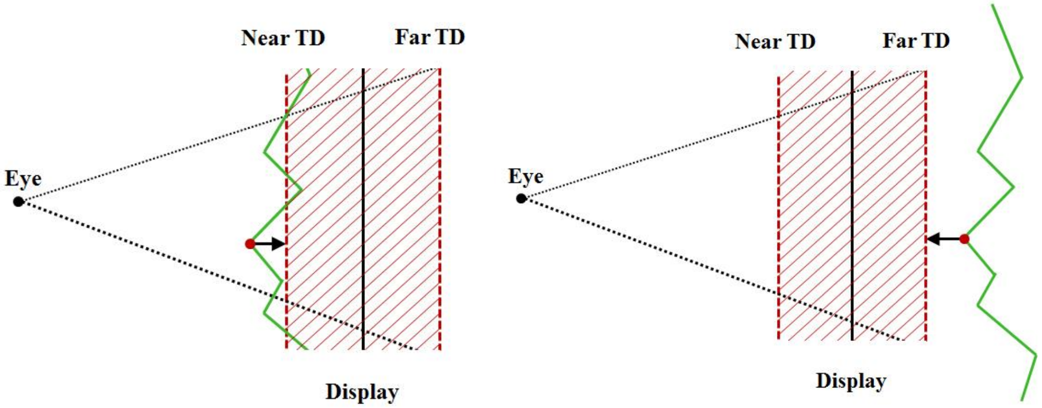
This section establishes this chapter’s terminology. Table 5 illustrates four eye separation measurements that can be distinguished using two independent classifications. Eye separation can be measured in either physical coordinates or virtual coordinates. The latter accounts for the 3D view (isotropic) scale factor. Further, we distinguish the user’s true eye separation versus the modeled eye separation which is used in the view frustum geometry. The *physical true eye separation* is commonly called the inter-ocular distance in psychophysics. The table gives example values for a human subject with an inter-ocular separation of 6 cm and a *physical modeled eye separation* of 3 cm. The view scale in this example is $1/10^6$ which makes a displayed virtual Earth roughly 1 meter in diameter. This yields virtual true separation of 60 km and virtual modeled separation of 30 km.

Head-coupled displays display 3D graphics where the generated perspective graphics image is dynamically adjusted based on head (or possibly more directly eye pupil) position. Head-mounted displays (HMDs) mount the displays on a headset or helmet. In contrast in Head-Tracked Displays (HTDs), the display is stationary mounted on a desk (desktop VR), a table (the virtual workbench) or one or more walls (the CAVE).

The Table 5 example is a case of *false eye separation*. This means the modeled eye separation is deliberately set to a value other than the true value for purposes of distorting the depth of the presented stereo 3D image. Human interocular distance varies subtly with vergence movements, but false eye separation is a technique that assigns modeled eye separation a value whose difference from the true value (modeled eye separation – true eye separation) is significantly larger than that occurring due to vergence movements. False eye separation distorts the 3D stereo image. The *modeled 3D image* is the displayed virtual 3D scene accounting for the view scale. An Earth globe (roughly 10^6 meters in diameter), might appear as a modeled 3D image of 1 m in diameter given a view scale of $1/10^6$.

The *perceived 3D image* is the stereo image the user perceives. There are numerous ways to operationally define the perceived image [149,63]. Here *perceived image* means the expected result of performing a registration experiment [46] between the synthetic image and a physical pointer under the further assumption stereopsis works like a theoretic range finder. Under this definition the perceived image is calculable for stereo cinema [134], HMDs [171], and HTDs [163].

5.2.1 Dynamic-Stereo Adjustment Technique



(A) (B)

Figure 24: Illustration of dynamic adjustment steps.

Our auto-adjustment differs from prior methods. Generally prior methods [162] have a static or dynamic model of the near fusible distance (*nf*) and farthest fusible distance (*ff*) relative to the display screen and then compute a nearest point (*np*) and/or farthest scene point (*fp*). *np* and *fp* are typically the nearest and farthest visible pixels. Dependent on the number of free parameters in a given adjustment technique, the adjustment could map *np* to *nf* and/or *fp* to *ff*. Holliman [69] extends this idea to allow mapping multiple depth ranges in model space to separate ranges in display space.

In this chapter, we only adjust one parameter. Roughly speaking we choose to map *np* to *nf* or *fp* to *ff*, but not both because we rely on adjustment transforms that have only one free parameter. If both the near and far points violate the target depth range, we adjust the near point only. *np* is prioritized because while positive screen parallax reaches a limit, negative parallax can increase without bound. Figure 24 illustrates our approach.

Near Target Distance (TD), is a static approximation of the nearest fusible distance. Far TD is a static approximation of a very conservative limit on the far fusible distance. An independent near clipping plane distance puts a hard limit on the near rendered geometry. It is kept very close to the eye because our 3D UI displays 3D cursors which are allowed to appear at any screen parallax. A standard z-buffer method determines the nearest point, *np*, in the scene. If $np < NearTD$ we adjust to bring *np* to *NearTD*. Otherwise, if $np > FarTD$ we adjust to bring *np* to *FarTD*. (Implicitly, if *np* is in the range [*NearTD*, *FarTD*], no adjustment occurs). This leads to a “buffer zone” such that if the nearest point of the target geometry is in the zone, no auto-adjustment occurs.

Our choice to <u>not</u> incorporate false eye separation or an additional technique limits our fusion control to only adjusting for either near or far point violation but not both. This choice is motivated by the larger context of our investigation.

First, we want to perform an empirical evaluation of comparing a no auto-adjustment condition to one or more auto-adjustment conditions within an MSVE travel task. The question is whether the auto-adjustments help or hurt based on the issues identified in the introduction (P1 and P2). By performing these adjustments alone we avoid potential confounds of any effects of false eye separation, etc.

Second, our travel technique uses a pair of 3D cursors. Best practice applies a fixed translation offset between the tracked input device and the 3D cursor [129]. We desire the effect of the stereo auto-adjustment technique to not affect the 3D cursor in perceived space. Our work is part of a more general investigation in using one and two handed 3D UIs for 3D visualization applications with very, rich 3D interactions including a multitude of ITs for cursor based direct manipulation and selection. Distorting the perceived 3D image of the 3D cursors via stereo adjustment complicates the specification of the offset between the buttonball and the cursor as well as analysis of the effects of using different offset values or offset algorithms. By only employing auto-scale or auto-translation, we implemented the scale and translation adjustments so as to affect the scene without affecting the 3D cursor's size and location (in perceived space). If we add false eye separation avoiding affecting the perceived image of the 3D cursors becomes difficult. (To our knowledge, while theoretically possible it had not been demonstrated).

Auto-Adjustment Condition

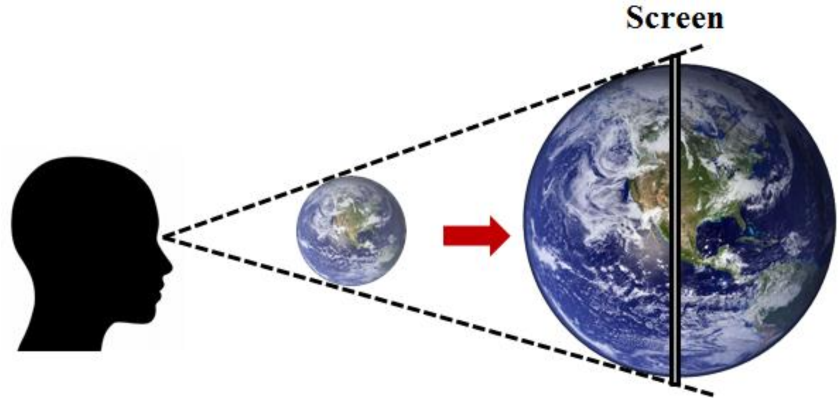


Figure 25: Auto-adjustment of the view scale factor

As explained, our adjustment methods adjust to one of two possible target distances with a buffer zone in which no adjustment will be performed. Our *auto-scale* (AS) condition uses a cyclopean scale to map *np* to *NearTD* or *FarTD* using the algorithm of Figure 24. The AS technique preserves the retinal angle and projected image size of the scene (Figure 25).

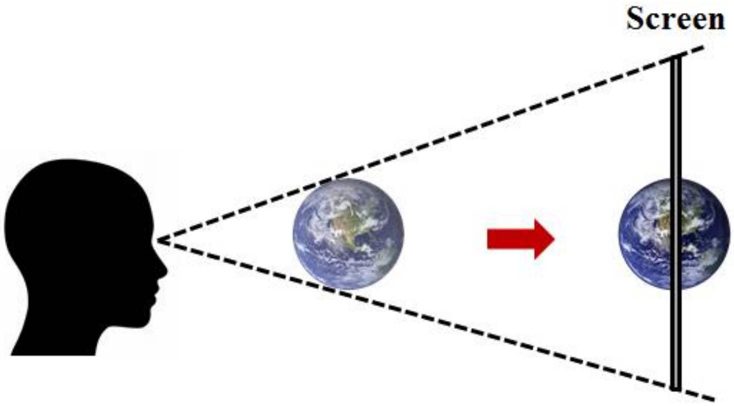
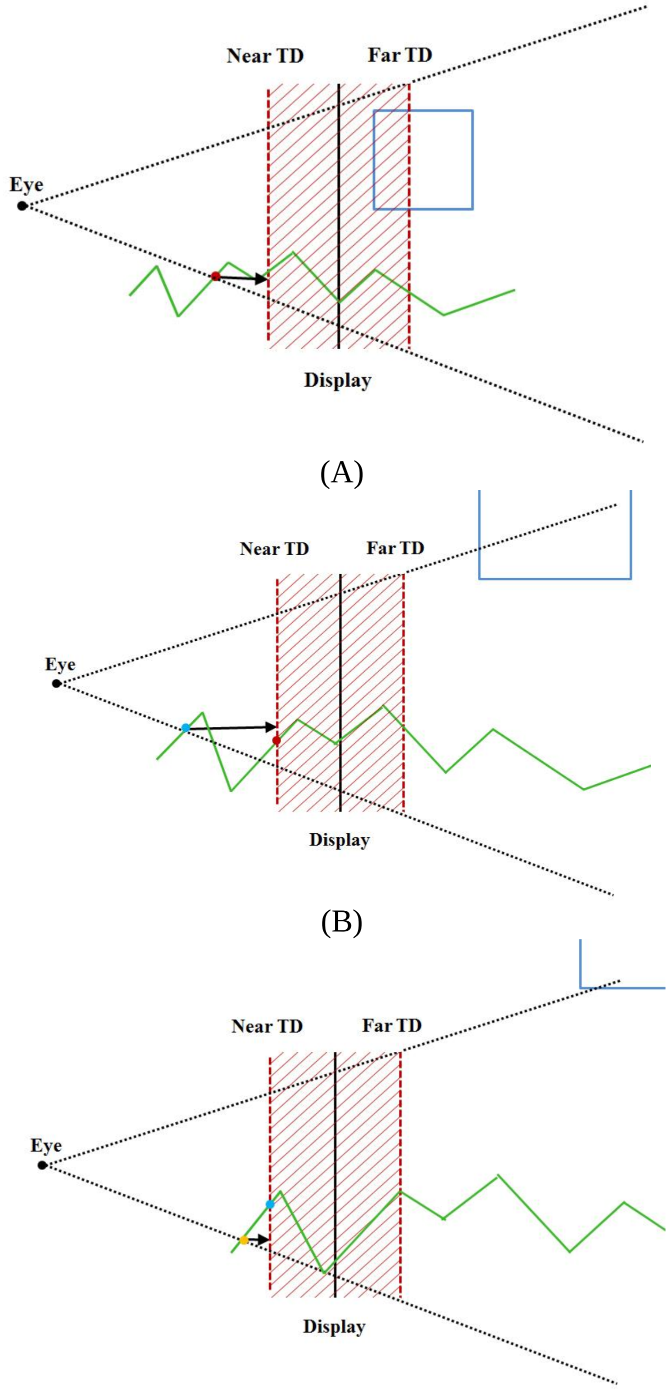


Figure 26: Auto-adjustment of the view location.

Our second adjustment condition adjusts view location. We refer to this technique as *auto-translation* (AT). AT translates the view perpendicular to the screen (Figure 26), similar to [165] but using the algorithm of Figure 1. This condition is included because while it does perform some auto-adjustment (which might reduce novice user's difficulty compared to purely manual 7DOF travel), it does not alter the view scale, possibly avoiding interfering with the user control of scale. This technique preserves 3D view scale, but changes retinal visual angle. The translation is performed smoothly over 0.5 s. Users typically immediately notice this transition compared to the view scale transition. When using a CAVE display, the direction of the translation needs to be changed based on the screen on which the user is fixated, or an estimate thereof.

5.2.2 Auto-Adjustment Stereo View Technique Problem



(A)

(B)

(C)

Figure 27: Illustration of the auto-stereo adjustment techniques' problem.

This section discusses cases where auto-stereo adjustment techniques are problematic. The first problem is an undesired continuous scale adjustment. In this scenario, the usability issues are (1) the adjustment moves a geometry target that the user is trying to advance to away from the user and (2) the adjustment does not switch off in a reasonable amount of time. In Figure 27A, the red dot is the nearest point in the view frustum. The green polyline is some terrain and the blue box is an object the user desires to inspect. The system changes the scale factor to adjust the near point to the Near TD. After the first adjustment, the system detects the new nearest point of the scene (the blue dot in Figure 27B) and the system does a second adjustment. After the second adjustment, the depth buffer detects the new nearest point (the orange dot in Figure 27C) and does a third adjustment. Therefore, the system will keep doing the adjustment until no pixel is in front of the Near TD and meanwhile the blue box, an object of user interest, keeps getting push further away. By itself this can be highly irritating to the user. Further, in an MSVE which truly contains large amounts of geometry in both Gigabytes and spatial-extent, such as a global terrain, if the algorithm continuously finds a new nearest point, the auto-adjustment will keep auto-adjusting. Imagine having an infinite surface (green) or a being inside an infinite cloud of volumetric data. Under such scenarios, each auto-adjustment finds a new non-fusible nearest point.

Use VEs that are actually very extensive because they do not use rendering engines that support out-of-core 3D databases. In contrast, much of our prior MSVE experience, including this chapter’s experiment, uses a global-terrain, out-of-core database. It was precisely in designing and informally testing our adjustment conditions in such an environment (Figure 30 and Figure 31), that this above scenario arose when the view looks over the horizon.

After experimenting with various approaches, we added the following rule to minimize this problem:

Rule 1: If the center of the virtual Earth is out of the view frustum, then the auto-stereo adjustment technique is deactivated.

Additionally, to handle other special cases there are two other rules:

*Rule 2: If the user's eye position is inside of the virtual Earth, then the auto-*stereo *adjustment technique is deactivated.*

Rule 3: If the user's eye position is between the Near TD and Far TD, then the auto-stereo adjustment technique is deactivated.

5.3 User Interface

In both the desktop VR and CAVE applications the user holds a pair of button balls (Figure 11 in CHAPTER 4) tracked by a Polhemus Fastrak. For brevity, in our further descriptions we assume the user is right handed. However, the UI itself accounts for an individual's user's handedness assigning button functionalities based on the user's dominant and less-dominant hands. We use a one-handed travel technique. Holding one button engages a scene-in-hand technique [153] and holding a second button engages rate controlled scaling where the center of scale is the cursor's position when the button is first pressed [122]. This provides full 7DOF travel. In case the user gets lost, a third button resets the view.

Virtual Environment

The desktop VR system uses a 22" Samsung 2233RZ display running at 120Hz as 1680 × 1050 resolution with nVidia 3D Vision glasses. A Polhemus Fastrak provides head and input tracking. The user is seated and holds one or two button balls, one per hand. A 3D cursor is displayed for each button ball at a fixed offset, set by the user at start up. This allows the user to rest her elbows on the desk, her lap, or chair arm [129]. Compared to other display environments where the user stands, this can reduce arm fatigue. Anecdotally the seated posture tends to reduce the users preferred range of arm motion. Based on informal pilot tests, we set the stereo TD to a fixed distance (±8" from the screen).

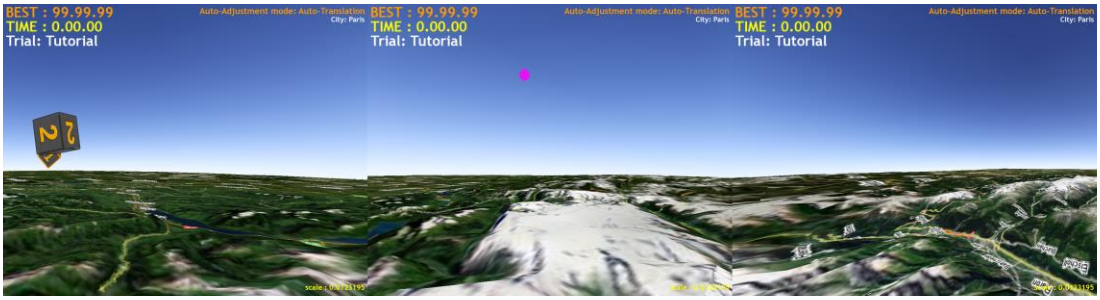


Figure 28: Screen capture of the MSVE application in the CAVE.

The CAVE consists of three large displays (approx. 8 feet × 6.4 feet physical size and 1280 × 1024 screen resolution each) and a Polhemus tracker with the wider range emitter. It provides wider Field of Regard (FOR), approximately 270°, than the desktop VR system. Figure 28 show a screen capture of the MSVE application in the CAVE. The user is stand with no place to rest her elbows or hands in the CAVE.

Since the CAVE system has three displays for navigation, Shaw and Green’s offset, perpendicular to the screen, must be modified. We implement a non-linear offset technique that supports a cursor offset in any directions (360°) based on the Go-Go technique [117,116] that provides non-linear offset for the 3D cursor by the user’s torso position in a HMD system

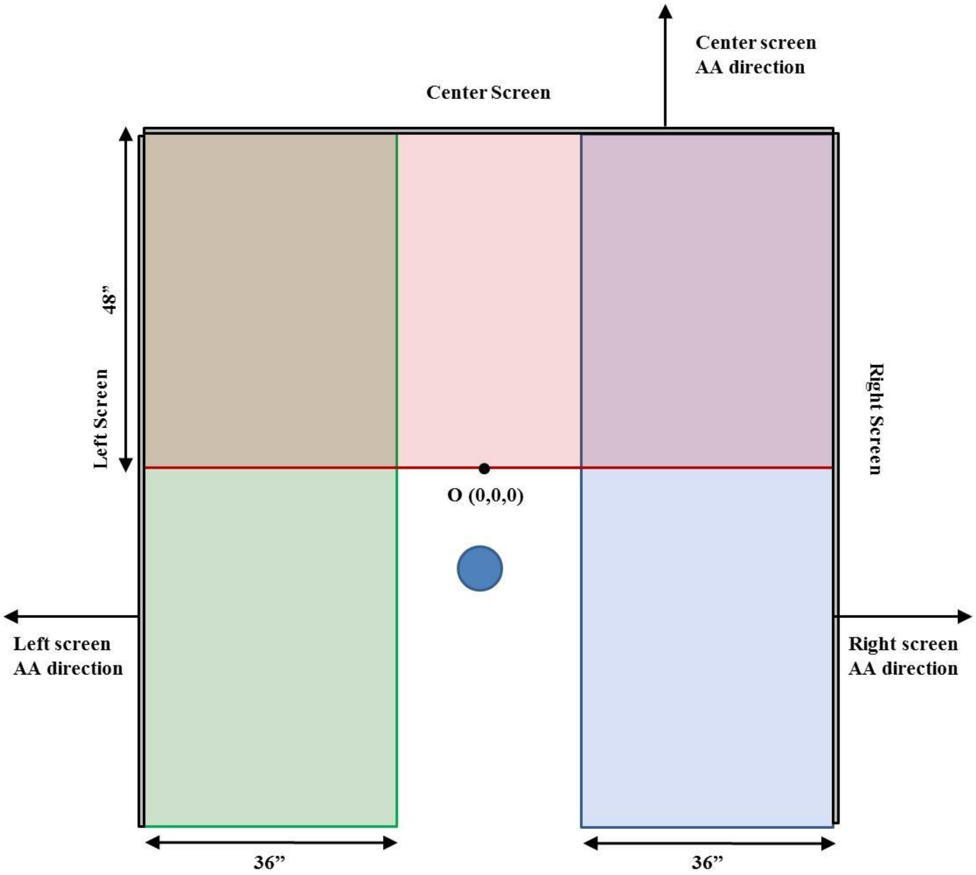


Figure 29: Illustration of target distances for three displays in the CAVE.

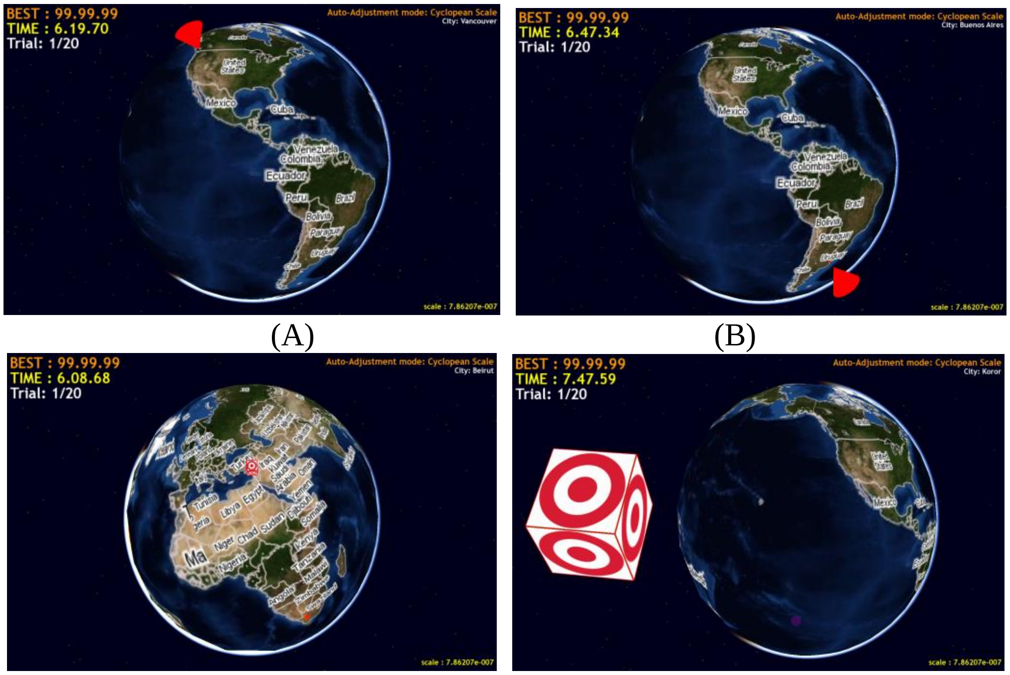
The larger screen size causes the user to stand farther from the screen. This changes the fusible depth range in a non-linear fashion. For the CAVE, the TD is 48" for the front screen and 36" for left and right screens. During the pilot testing, we found that the user tend to stand on approximately 6 Ft from the center screen and 4 Ft from right or left display (blue circle in Figure 29). When the user changes his view to the left or right, she tends not to move her body. With 48" TDs, the AA technique is deactivated because of the *Rule 2*. Hence, we set shorter TDs (38") for left and right displays for the case that the user doesn't move. Figure 29 illustrates different target distances for three displays of our CAVE system.

For the CAVE system, three displays have separate depth cameras for z-buffer sampling. Only one depth camera is activated based on the position of the center of the virtual Earth. For example, if the center of the virtual Earth is in the right screens view frustum, only the right screen's depth camera is used for auto-stereo adjustment.

5.3.1 Application

Our MSVE application is built using OpenSceneGraph [105], osgEarth [106] and osgVE [138]. Our experiment is designed for a global, virtual Earth based on the task of visiting a place of interest, such as a famous city, country or landmark, and then inspecting details of the region. Therefore, we defined the first task as finding a target box which is randomly located on the virtual Earth (see Figure 30). The box appears at one of four different sizes. This condition tests for any interaction of the auto-adjustment condition with the range of view scale change required to reach the target box. To motivate participants, we use pre-defined locations of the target box at capitals or famous cities in the world. In addition, we divided the world into spatial domains by its distance from the start position (America, Africa, Asia, Australia and Europe). This maintains similar travel distance across participants and ensures each spatial domain occurs at least once per box size. A timer appears in the upper left of the screen. The user can see her best time below the timer. The current trial number is shown below the best time. The upper right of the screen displays the auto-adjustment's engagement status as either "*on*" or "*off*". A name of the city, which is a target box location, is displayed below the auto-adjustment engagement status. The view scale factor is displayed on the bottom right of the screen.

5.3.2 Experiment Design



(A) (B)

(C) (D)

Figure 30: The target box has four different box sizes: (A) box size 1, (B) box size 2, (C) box size 3 and (D) box size 4.

(A) (B)

Figure 31: Tasks for experiments. (A) Task 1: target box finding and (B) Task 2: checking inside of surrounding boxes.

Each experimental trial involves two tasks. In Task 1, the user travels to a target box which is randomly located on the Earth. The box comes with four different sizes (see Figure 30). If a target box is too small to be seen by the user at the start position, then a red arrow, whose world coordinate size is dynamic to maintain a roughly constant screen space size, indicates the target box location (Figure 30A and Figure 30B). The user must travel (pose and scale) to position the box within a screen centered wireframe box (Figure 31A). After the user finishes Task 1, the target box disappears and four numbered boxes appear that indicate four cardinal directions; they are the same size as the target box (Figure 31B). Each box has a small hole on one face and a tiny colored sphere inside. The sphere color (red, blue or white) matches the colors of the button ball's buttons (see Figure 11 in CHAPTER 4). The user must carefully maneuver to see the sphere color through the hole. The user indicates the sphere color by pressing the corresponding button on the left button ball. The user examines the boxes in order of their number labels. A success sound plays when the user presses the correct colored button. After the user presses the correct button for all four boxes, a new trial begins. The user can reset the view position to the initial position by pressing a button of the right button ball during a trial if lost. For Task 1, the initial position is where the user can see the entire virtual Earth (Figure 30). For Task 2, the initial position is the last position where the user finished Task 1.

We used a simpler docking-task like navigation task to train participants on how to use buttonball input for the travel technique for 10 minutes. After the training, the instructor teaches the user about stereoscopic fusion problems by showing a case of extreme negative parallax. The instructor also explains how auto-stereo adjustment (AA) techniques try to minimize fusion problems.

We use a within-subject design (AA × BoxSize) repeated measures ANOVA (analysis of variance) for each Task and display condition (desktop VR or a CAVE) to analyze output of our experiments. Participants need to accomplish two navigation tasks with three AA conditions: *Auto-Scale* (AS), *Auto-Translation* (AT) and *No Auto-stereo*

Adjustment (NA). Each participant performs 20 trials with each AA condition. In each trial, a target box appears in a random city with random size for Task 1. Orientation of numbered boxes is also randomized per trial for Task 2. We recorded task completion time and number of resets for both Task 1 and Task 2. AA condition order was fully counter-balanced between subjects using Latin squares.

Our primary hypotheses are:

H1.AS and AT are expected to have faster completion time than NA for the both Task 1 and Task 2. This is because they partially reduce the DOFs the user must manually adjust.

H2.AT is expected to have faster completion time than AS for the both Task 1 and Task 2. This is because AS auto-scaling may interfere the user desired manual scale.

H3.AS and AT are expected to produce less stereo fusion problems than NA for the both Task 1 and 2.

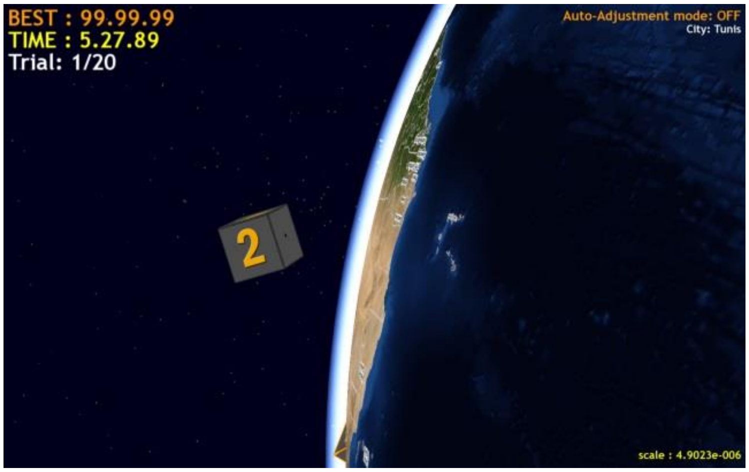


Figure 32: Example of the auto-adjustment techniques' problem with box size 3.

Interestingly, we found that the AA techniques produce another problem during pilot studies. In the box size 3, it's hard to see the inside of the box if the hole is face to the virtual Earth (see Figure 32). This is because AA techniques change view scale or distance when the user tries to rotate the scene to see the downward facing hole. We didn’t define a rule for this problem. Rather, we observed how this problem effect on the user’s ability to accomplish tasks during experiments.

We use the per-trial mean of task completion time and number of resets. The reported F tests use α=.05 for significance and indicate the Geisser-Greenhouse correction to protect against possible violation of the sphericity assumption. The post-hoc tests that were conducted were Fisher’s least significant differences (LSD) pairwise comparisons with α=.05 level for significance. Error bars in graphs represent 95% confidence interval.

5.4 Experiment

For Experiment, we recruited 24 participants (twelve participants for each display condition) from the Computer Science department and the Psychology department participant pool for the experiment. All participants have (corrected) 20/20 or higher eye vision. In the desktop VR group, eight participants are CS major and four are non-CS major (Ten undergraduates, one master, and one Ph.D. student). Eight are males, and four are females. Participants have highly daily computer usage (6.67 out of 7). Nine participants have experience with 3D UIs such as Microsoft Kinect. In the CAVE group, six participants are CS major and six are non-CS major (eight are males and four are females). Five are undergraduates, two are masters, and five are Ph.D. students. Participants have highly daily computer usage (6.42 out of 7). Three participants have an experience with 3D UIs.

5.4.1 Result

Quantitative

Table 6: Average and standard deviation (SD) of task completion time (CT) by AA condition and box size of the desktop VR group with the OH IT.

		NA		AS		AT	
	Box Size	CT	SD	CT	SD	CT	SD
Task 1	1	30.25	16.93	25.33	12.11	20.66	5.16
	2	23.04	7.79	17.02	5.28	16.57	4.30
	3	15.86	6.33	11.07	1.83	11.00	3.00
	4	6.91	1.96	6.12	1.06	6.63	1.79
	Overall	19.01	12.96	14.89	9.71	13.71	6.52
Task 2	1	34.44	12.04	32.08	8.75	34.53	8.18
	2	35.08	12.61	30.68	5.50	31.90	6.83
	3	40.86	12.95	37.11	11.07	37.50	7.56
	4	33.09	12.05	37.19	12.63	33.00	7.26
	Overall	35.87	12.39	34.27	9.99	34.23	7.54

Table 7: Average and standard deviation (SD) of task completion time (CT) by AA condition and box size of the CAVE group with the OH IT.

		NA		AS		AT	
	Box Size	CT	SD	CT	SD	CT	SD
Task 1	1	37.68	12.86	32.08	9.38	34.85	12.75
	2	33.73	11.83	24.70	4.56	24.46	9.22
	3	21.88	5.59	19.71	6.23	19.13	8.20
	4	8.09	1.91	7.69	2.31	8.23	3.27
	Overall	25.35	14.68	21.05	10.78	21.67	13.05
Task 2	1	46.48	12.75	64.70	14.40	55.72	17.41
	2	45.53	13.04	60.18	16.36	53.51	17.59
	3	42.64	11.85	57.11	18.03	44.91	11.50
	4	42.47	18.22	43.04	8.02	40.00	12.48
	Overall	44.28	13.84	56.26	16.41	48.53	15.88

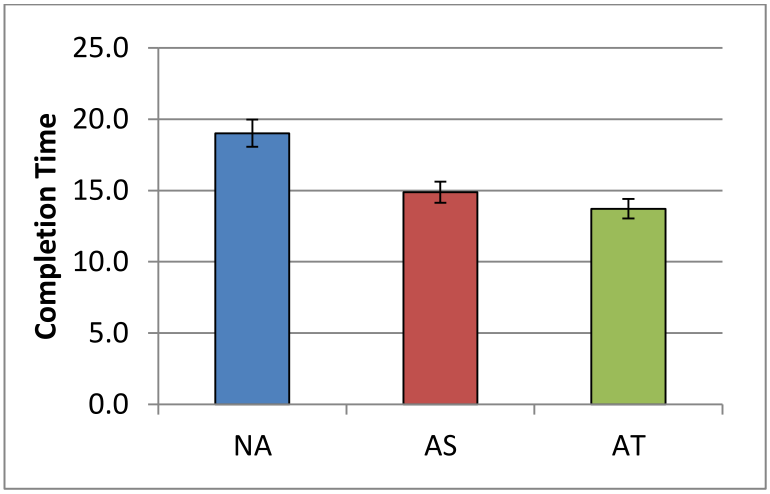


Figure 33: Task completion time of AA conditions of Task 1 of the desktop VR group with the OH IT.

Table 6 and Table 7 illustrate averages and standard deviations of task completion time of AA conditions by box size in the desktop VR and the CAVE. Results of ANOVA for Task 1 of the desktop VR show a main effect on completion time of AA condition (F(2,22)=5.871, p=.009, η_p^2=.348). LSD tests show completion time of NA, 19s, is slower than AS (p=.015) and AT (p=.036), 14.9 and 13.7s (see Figure 33). However, completion times between AS and AT do not differ (p=.361). In the CAVE, there is no significant main effect of AA condition for Task 1 (p=.082).

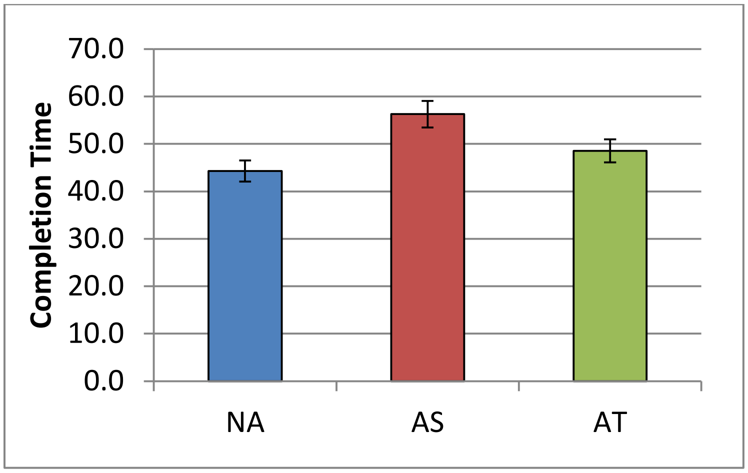


Figure 34: Task completion time of AA conditions of Task 2 of the CAVE group with the OH IT.

There is a main effect on completion time of AA condition for Task 2 of the CAVE (F(2,22)=7.624, p=.003, η_p^2=.409). LSD pairwise comparisons show completion time of AS condition, 56.3s, is significantly slower than NA, 44.3s, (p=.005) and AT, 48.5s, (p=.036) (see Figure 34). However, completion times between NA and AT do not differ (p=.137). In the desktop VR, there is no main effect on completion time of AA condition for Task 2 (p=.132).

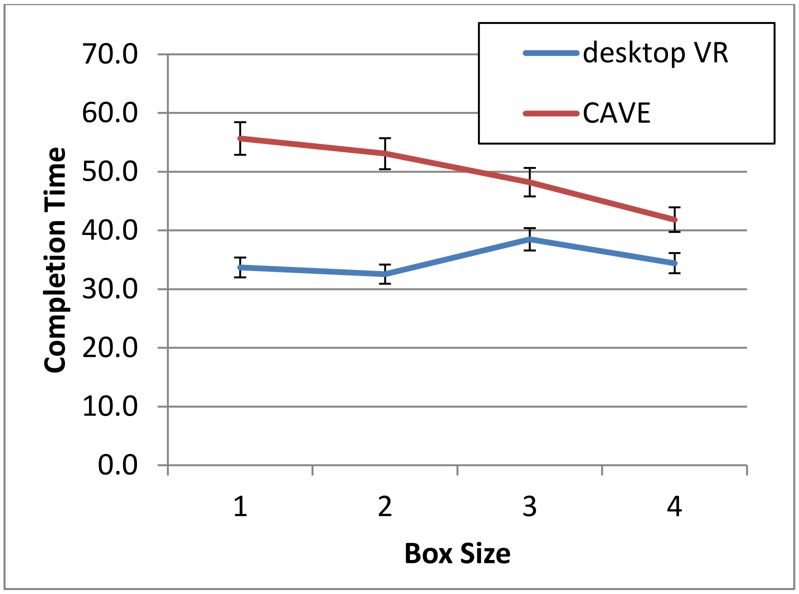


Figure 35: Completion time of box size of Task 2 (blue) of the desktop VR group, and (red) of the CAVE group with the OH IT.

The a main effect of box size on completion time for Task 2 of the desktop VR is significant (F(3,33)=6.294, *p*=.002, η_p^2=.364). LSD comparisons show that box size 3 (M=38.5, SD=10.6) has slower completion time than box size 4 (M=34.4, SD=10.8, *p*=.044), 2 (M=32.6, SD=8.8, *p*=.003) and 1 (M=33.7, SD=9.6, *p*=.016). The blue line in Figure 35 illustrates different completion time of four box sizes for the Task 1 in the desktop VR.

There is a main effect of box size on completion time for Task 2 in the CAVE (F(3,33)=8.918, p<.001, η_p^2=.448). LSD comparisons show that box size 4 (M=41.8, SD=13.2) has faster completion time than box size 3 (M=48.2, SD=15.1, *p*=.050), 2 (M=53.1, SD=16.5, *p*<.001) and 1 (M=55.6, SD=16.4, *p*=.002). Box size 3 has faster completion time than box size 1 (*p*=.024) (see the red line in Figure 35).

There are no interaction effects between box size and AA conditions in either VE system. The number of resets did not differ significantly across any conditions.

Qualitative

Table 8: Average and standard deviation (SD) of arm fatigue and stereo fusion problem rates with the OH IT by AA condition of the desktop VR group.

		NA	SD	AT	SD	AS	SD
Arm Fatigue		3.75	1.48	3.25	1.42	3.67	1.50
Fusion Problems	Task 1	2.50	1.24	2.58	1.56	2.67	1.54
	Task 2	3.17	1.40	2.75	1.29	2.25	0.87

Table 9: Average and standard deviation (SD) of arm fatigue and stereo fusion problem rates with the OH IT by AA condition of the CAVE group.

		NA	SD	AT	SD	AS	SD
Arm Fatigue		4.00	1.41	3.67	1.56	3.17	1.75
Fusion Problems	Task 1	2.67	1.44	1.83	0.83	2.08	1.38
	Task 2	3.42	1.88	2.50	1.45	2.67	1.16

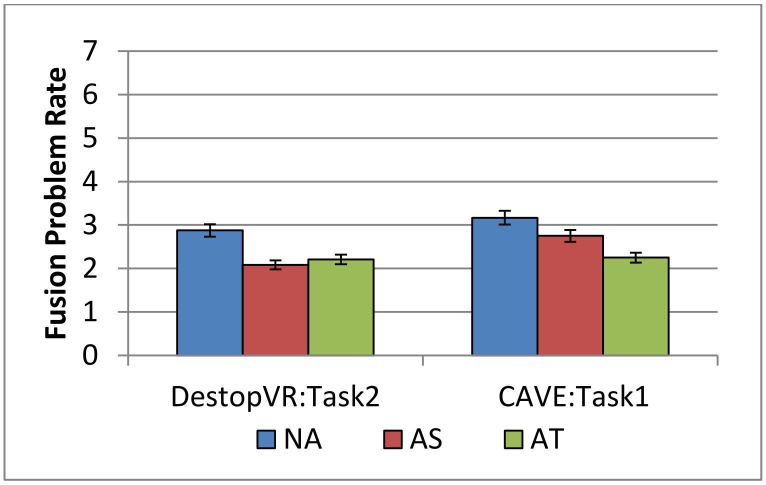


Figure 36: Stereo fusion problem rate of Task 1 of the Desktop VR group (left) and Task 2 of the CAVE group (right) with the OH IT.

We asked how much arm fatigue the user felt after the user finishes the experiment for each AA condition (on a 7-point Likert scale, 1=*not at all* to 7=*very frequently*). Table 8 and Table 9 show average arm fatigue rate by AA condition for desktop VR and CAVE groups. The one-way repeated measures ANOVA shows no main effect of AA condition on fatigue rate of the desktop VR group (p=.125) and the CAVE group (p=.096). This suggests that AT and AS do not induce more arm fatigue than NS for overall trials.

Participants took a post survey after the experiment. On a 7-point Likert scale regarding how frequently the user experienced stereo fusion problems (1=*not at all* and 7=*very frequently*). Table 8 and Table 9 shows average fusion problem rate of Task 1 and Task 2 by AA condition of desktop VR and CAVE groups. The result of the desktop VR group shows no statistical difference of AA condition on a stereo fusion problem rate for Task 1 (p=.693). However, there is a main effect of AA condition on the stereo fusion problem rate for Task 2 (F(2,22)=4.529, p=.023, η_p^2=.292). LSD comparisons show users felt stereo fusion problems less with AT than NA (p=.020). However, there is no statistical difference on stereo fusion problem rate between NA and AS (p=.175) or between AS and AT (p=.111) (see the left graph in Figure 36).

The result of the CAVE group shows a significant main effect on stereo fusion problems rate of AA condition for Task 1 (F(2,22)=4.158, p=.029, η_p^2=.274). LSD post-hoc tests show that the user feels more fusion problems with NA than AT (p=.046) and AS (p=.025) conditions. There is no difference AT and AS (p=.429) (see the right graph in Figure 36). In addition, there is no main effect on stereo fusion problems rate of AA condition for Task 2 (p=.221). This indicates AT reduces the stereo fusion problems. We expected both AT and AS would reduce the stereo fusion problems for both Task 1 and Task 2 (*H4*). However, the result shows that only AT reduces fusion problems for Task 1 not for Task 2.

Regarding whether the AS technique is helpful to accomplish tasks and reduce fusion problems (1=*not at all* through 7=*very helpful*), the average rate of the desktop VR group is 4.82 and the CAVE group is 3.33. Regarding whether the AT technique is helpful, the rating of the desktop VR group is 4.6 and the CAVE group is 4.25. In the desktop VR group, six participants answered they prefer the AS condition, five answered the AT condition, and one had no preference. In the CAVE group, three participants preferred the AS condition, six preferred the AT condition, three had no preference and one disliked both.

5.5 Discussion and Conclusion

5.5.1 General Discussion

Our results show that auto-adjustment of stereo view parameters techniques helps to reduce stereoscopic fusion problems in both the desktop VR system and CAVE. In the subjective questionnaire, the user reports significantly less stereo fusion problems with the auto-scale and auto-translation conditions compared to the no adjustment condition for target finding tasks when using desktop VR system. In the CAVE, however, the user reports significantly less fusion problems with only the auto-translation condition compared to the no adjustment condition for the inspection task. That is the auto-adjustments' fusion problem reduction seems to be muted in the CAVE compared to desktop VR.

One possibility of differing results is FOR differences. Wider FOR is known to increase simulator sickness. In this sense, the three-screen CAVE could be expected to increase general reports of discomfort compared to viewing a single monitor and possibly this general simulator sickness would be reported as a stronger experience of stereo fusion problems. Also, the two display systems use different technologies for stereo image separation (Nvidia 3D Vision active shutter-glasses vs. dual Barco projectors with circularly polarized glasses). This generates different lumens and stereo cross-talk for the two display systems. While both display systems reside in the same room the ambient lighting conditions were not exactly the same due to differences in overhead lighting.

Prior work with stereo fusion control indicates that a combination of AT or AS with an additional stereo auto-adjustment to allow simultaneous control of both the near and far point would further reduce reported fusion problems. However, the goal of this experiment is to examine the interaction of auto-translation and auto-scale with travel tasks and to merely verify AT and AS alone are reducing (or at least not increasing) fusion problems.

Both auto-translation and auto-scale reduce completion time in the desktop VR for the target finding task but show no significant effects for the inspection task. The results support hypothesis H1 for Task 1 but not Task 2. A plausible explanation is that auto-adjustment techniques also automate one of the 7DOFs, leaving the user with a travel task similar to the lesser difficulty of a 6DOF task. We observed novice users having difficulty manually controlling 7DOF travel in the NA condition. These results suggest: (1) in Task 1 auto-adjustment helps, allowing users to complete Task 1, faster, but (2) in Task 2, there is less need for further manual scale change so users experience Task 2 more like a 6DOF task and hence auto-adjustment reduction of DOF complexity becomes superfluous. AT did not perform significantly different than AS. This fails to support H2, that the auto-scaling of AS would interfere more with the user reaching a desired scale than AT would.

For the CAVE, however, neither auto-translation nor auto-scale helps Task 1. This fails to support H1 or H2. We observe that across all conditions CAVE completion times are longer. This might wipe out any time improvement from auto adjustment.

One potential cause for longer CAVE completion times is that the stereo TD is set based on stereo fusion considerations, not on reachability. While in the CAVE the Far TD is farther from the screens than the desktop VR's Far TD is, in desktop VR the Far TD was still closer to the user's nominal seated shoulder position, than the CAVE Far TD was to the user's nominal standing shoulder position. Without any cursor offset this would mean that the Earth, auto-adjusted to the Far TD, would be harder to reach with the 3D cursor in the CAVE. In turn users might engage the scene-in-hand IT or cursor-centered scale IT with the cursor further away from the Earth's surface. Being able to place the cursor close to the surface or even inside the Earth tends to make rotation and scale manipulations more productive. In desktop VR, the combination of the fixed translation offset and the Far TD location generally meant one can easily place the cursor close to or inside the Earth when auto-adjusted to the Far TD. With the CAVE however, the non-linear offset mapping gain factor did not allow the cursor to reach the Far TD unless the user walked several feet towards the screen. Anecdotal observation indicates CAVE users often did not walk much and tended to stay in a central location. Unfortunately, we had not anticipated this.

Another CAVE complication was that the auto-stereo adjustment techniques are activated relative to a particular screen's TD. Our heuristics for dynamically choosing which screen to use for the TD may well be insufficient. They were designed to guess what screen the user was fixated on. Possibly, they chose the 'wrong' screen causing the auto-adjustment to adjust in an unhelpful direction. This might be contributing to the longer CAVE completion times. A good solution is to employ gaze tracking to pick the screen to use for the adjustment TDs.

We observed an anomalous the effect of box size 3 in Task 2. Through further testing, this appears to be due to the relative size of the box to the Earth's size and the box's height above the Earth. The trouble occurs if the hole is facing toward the virtual Earth. In order to look in the hole using the scene-in-hand IT one rotates the view in a manner than tends to place the surface of the Earth between the user's eye and the box, thus occluding the box. Avoiding this occlusion requires further view manipulations. Additionally with an auto-adjustment technique, the initial view rotation tends to make the opposite side of the Earth the near point and the auto-adjustment may push the Earth and box farther away. The peculiarity of this situation demonstrates that auto-adjustments are quite difficult to 'get 100% right'. This scenario would not be uncovered without the many trials of formal evaluation done with a variety of travel and inspection tasks on an extensive MSVE. A simple solution to the anomaly is to allow the user to disable auto-adjustment if desired. More generally, it indicates more sophistication is needed for auto-adjustment to 'always do the right thing.'

Finally, our experience of testing the auto-adjustment techniques by ourselves is that the auto-adjustment techniques did not improve our completion time even in the cases where it improved completion time for the study participants (mainly Task 1). Our anecdotal observation of participants' behavior under the NA condition found they often using suboptimal strategies for manipulating 7DOFs during the task. In contrast the AS and AT conditions tended to help them by automating adjustment of one of the DOFs. This coupled with our own experience of not experiencing a completion time reduction under AT or AS may suggest that expert users of 7DOF travel techniques learn to adopt 7DOF travel strategies that obviate the help provided by auto-adjustments. Hence, it is possible that the AA and AT auto-adjustment methods may be most useful as ‘training wheels’ for novice users of MSVEs that require 7DOF travel. (However, even for experts the stereo fusion control aspect of auto-adjustment may remain useful).

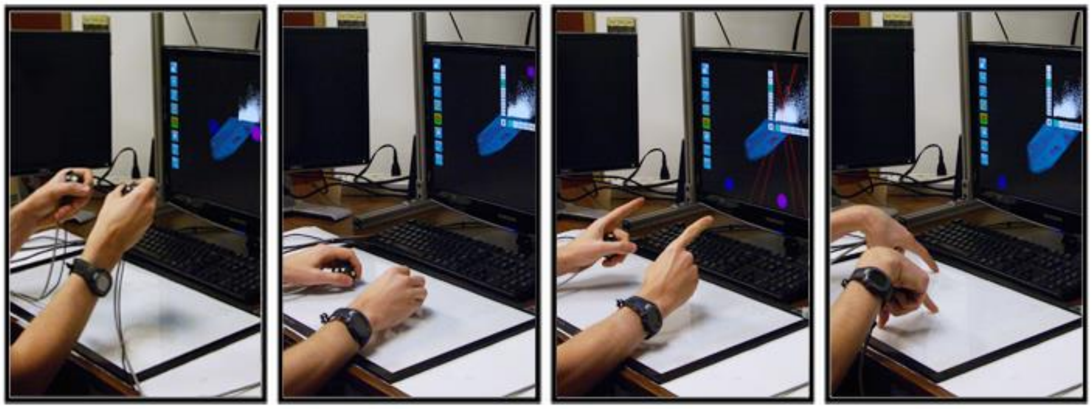
5.5.2 Conclusion

This chapter evaluates two stereo fusion control techniques in a MSVE. The user study demonstrates advantages and disadvantages of auto-stereo adjustment techniques in the desktop VR system and CAVE. Our results show that auto-stereo adjustment techniques reduce stereo fusion problems in both VE systems for certain tasks. In the desktop VR, users report reduced stereo fusion problems during Task 1. In the CAVE, users report reduced fusion problems during Task 2 (inspection). The auto-translation or auto-scale can only control fusion violations for either the near or far point, but not both. A deployed solution would combine auto-translation or auto-scale with false eye separation (or related non-linear technique) to allow fusion control for both the near and far point.

More significantly regarding whether fusion driven auto-scale helps or hinders 7DOF travel in MSVE, in the desktop VR system, both auto-adjustment techniques (auto-translation and auto-scale) had equally faster completion times than no adjustment for the target finding task, but not for the inspection task. This indicates there are two benefits to use the described auto-adjustments in desktop VR: fusion control and easier DOF management. Anecdotally, these may be of less benefit to users with years of experience using 7DOF travel in MSVE on stereo systems.

In the CAVE, both auto-adjustment techniques failed to help with Task 1 and auto-scale was detrimental to performance in Task 2. We suspect this is due to the fact that our methods for addressing the greater complexity of auto-adjustment for multi-screen displays are inadequate. However, this produced two lessons learned. (1) For multi-screen displays, it appears gaze tracking maybe necessary for determining which screen should be used for the auto-adjustments. (2) Results also suggest that the assumptions used to calibrate the gain factor for a CAVE non-linear cursor offset for cursor interactions that occur within 10' of the CAVE floor center need to account for whether a user prefers to stand in the middle of the CAVE floor or walk within the CAVE to reach ‘just-out-of-reach’ objects. Our pilot studies indicated the latter--which drove our design, but our larger study’s results suggest the former--making our design choice sub-optimal and possibly explaining the longer CAVE completion times.

CHAPTER 6: HYBRID FINGER BALL FOR MULTI-DIMENSION APPLICATION IN DESKTOP VR



(A) (B) (C) (D)

Figure 37: The HyFinBall UI supports (A) 6DOF isotonic input, (B) planar-3DOF input, (C) 3D hand and finger tracking and gesture and (D) multi-touch.

6.1 Introduction

In this chapter, we present a minimally immersive, desktop VR [40] interface for a visual analytic application that provides two-handed bat (3D mouse) input, two-handed 2D mouse input, multi-touch and 3D gesture. The primary devices are two 6DOF button balls. We used these previously [146], borrowing from the bat, the FingerBall [177], and the button-enhanced bat [129]. This chapter presents the HyFinBall ("hybrid-finger-ball") user interface (UI) described below:

HyFinBall: The HyFinBall interface starts with a pair 6DOF tracked balls with multiple buttons. Each ball is 45 mm in diameter roughly the size of a ping-pong ball. The software UI has the following properties. When a button ball is held in the air (Figure 37A), a 3D cursor is displayed and 6DOF (xyz + yaw,pitch,roll) interactions are active. When a button ball is placed on the desktop, the UI automatically switches from treating button ball as 6DOF isotonic device to treating it as a planar-3DOF input device (xy-position + yaw) and the 3D cursor is replaced by a 2D cursor in the plane of the screen (Figure 37B). Each button ball independently switches between a 6DOF and planar-3DOF mode. During this switch, the UI techniques available for the button ball switch from 3D ITs to 2D ITs. There is a translational offset between the physical location of the HyFinBall and its displayed 2D and 3D cursors. 6DOF mode uses an elbow-resting posture [129] while planar-3DOF mode uses a palm-resting posture. Strong consideration is given to stereoscopic display issues in the desktop VR environment when displaying the cursors. In particular, certain planar-3DOF ITs use projected 3D cursors.

HyFinBall + Finger-Tracking: The HyFinBall is small enough to hold in a precision grasp [177] and small enough to be held with only the pinky, ring finger and palm in an average adult hand. This leaves the thumb, forefinger and (possibly) middle finger free. The free fingers can either:

*interact on a horizontal 2D, multi-touch desktop display (*Figure 37D)

OR

*perform three finger 3D interaction and gestures when in 6DOF mode (*Figure 37C).

By design, these 2D and 3D finger-tracking modes can be engaged without incurring an acquisition time penalty, i.e. the user does not drop and pick-up the button ball to engage and disengage these finger interaction modes.

The concept of using a single device that switches automatically between 6DOF mode and planar 3DOF mode, while not new (such the VideoMouse [67], and Logitech 2D/6D Mouse [88]) has not, to our knowledge, been integrated into any rich application that requires both 3D interaction and 2D interaction across coordinated views. The design space implied by the HyFinBall interface has not been explored with respect to desktop VR environments (in particular its stereoscopic 3D component) and this type of interface been not been studied for one-handed UIs, let alone two-handed UIs. To our knowledge, there has been no demonstration of a hybrid user interface (HUI) where the user uses a small form factor 6DOF held-device with a precision grip that can be continuously held while allowing the free fingers can engage in 2D multi-touch and/or 3D gesture interaction.

We present the HyFinBall and HyFinBall+Finger-Tracking concept and prototype (hardware and software). We present our anecdotal observations and describe the design space of the resulting hybrid interaction techniques. This is done in the context of a rich, visual analytics interface containing coordinated views with 2D and 3D visualizations and with strong consideration of stereoscopic display issues in desktop VR. Finally, we present a user study that is focusing on the core HyFinBall concept comparing it to a mouse, the planar-3DOF-only mode and 6DOF-only mode across a variety of 2D and 3D combination tasks.

6.2 User Interface

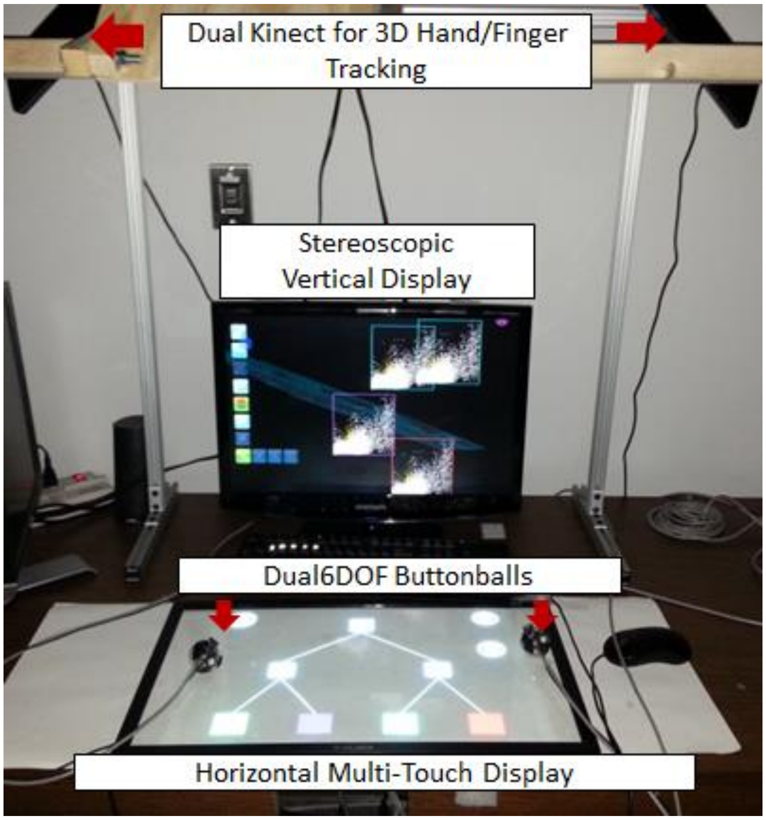
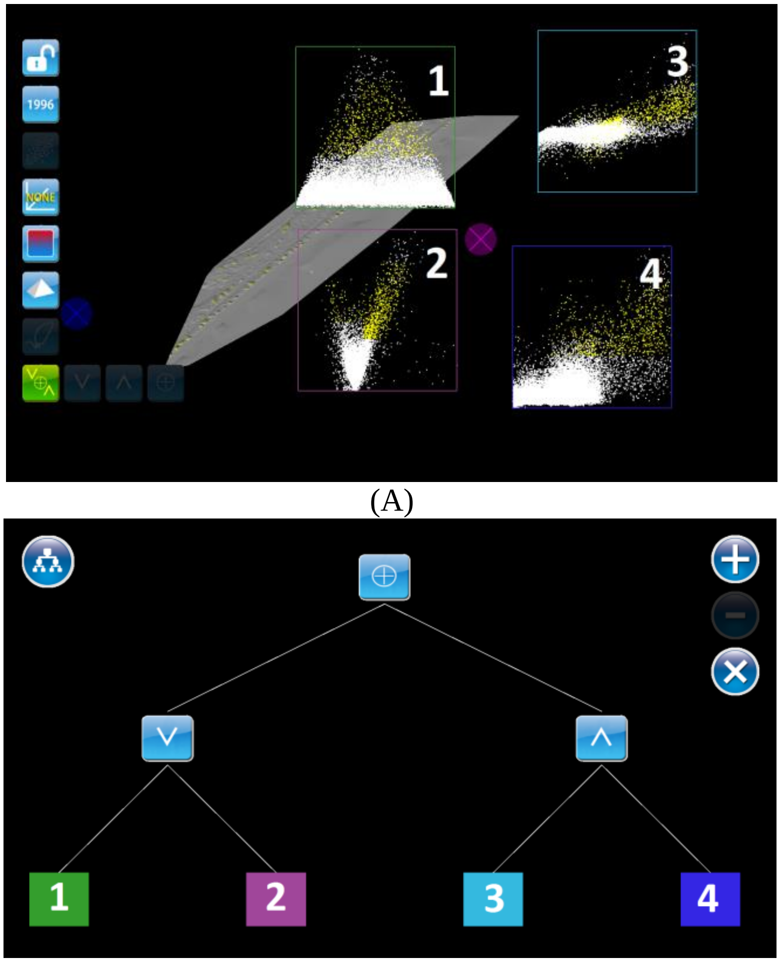


Figure 38: HyFinBall UI. It consists of head-tracked stereoscopic vertical display, projected multi-touch table using PQLab multi-touch frame, dual button balls, and dual Kinects for 3D hand and finger-tracking.

We present the HyFinBall UI in the context of a rich multi-dimensional application called DIEM-VR. DIEM-VR is a tool for analyzing terrain meshes from 10 years of LIDAR scans of the North Carolina Coast from the NOAA Digital Coast database. We extend a linked feature space interface from our prior work [62] that integrates multi-dimensional, feature space 2D views with 3D terrain views. In the HyFinBall system, the user sits at dual screen, desktop VR system. It uses Nvidia 3D vision glasses and a Polhemus Fastrak for head-tracking and for tracking the HyFinBall devices. Two Windows Kinects view the desk space running 3Gear's finger-tracking software [1]. A PQ Labs multi-touch screen [118] is placed on the horizontal desktop with an overhead projector (Figure 38). A pure 2D display and direct 2D manipulation is performed on the multi-touch horizontal display while 3D content as well as limited 2D content appears on the vertical display. For the DIEM-VR application, the images in the vertical and horizontal screens in Figure 38 are reproduced as a screen captures Figure 39A and B. The vertical screen displays a 3D terrain patch as well as feature space 2D scatter plots. The horizontal display shows an interactive Boolean expression tree that controls selection of terrain mesh points using a Boolean combination of the highlighted selections in the 2D scatter plots.



(A)

(B)

Figure 39: (A) Scatter-plots with selected regions and interactive, and (B) Boolean expression tree.

As discussed in the introduction, the HyFinBall UI takes particular advantage of the small form factor of the button ball to enable a number of 3D and 2D interactions' paradigms without having to drop and reacquire the input device. In the DIEM-VR application 3D navigation and volumetric selection of the 3D terrain occurs using one or both of the button balls held in the air, but with the elbows resting. This 6DOF button ball interaction mode is shown in Figure 37A and Figure 40A. Next, interaction with 2D objects on the vertical screen, such as the scatter plots in DIEM-VR, occurs with one or both of the button balls placed on the desk surface (Figure 37B and Figure 40B) in which case planar-3DOF interaction mode is enabled. Third, when the user tucks the button ball in his palm (Figure 37D and Figure 40D and E), the free fingers such as the thumb and pointer finger interact with the 2D graphics on the horizontal display using multi-touch. In DIEM-VR, this multi-touch mode controls the Boolean expression tree mentioned earlier and elaborated upon in Section 6.3.4. Finally, although only experimentally implemented in our DIEM-VR application (due to tracking limitations), when the user tucks a button ball in his palm and makes 3D pointing gestures (Figure 37C and Figure 40C), 3D hand and finger tracking enables a ray-based 3D selection within DIEM-VR.

(A) (B) (C) (D) (E)

Figure 40: (A) Hand off table, 6DOF mode, (B) hand on table, planar-3DOF mode. (C) Dual fingers 3D gesture, (D) fingers on table, multi-touch (side view) and (E) fingers on table, multi-touch (top x-ray view showing held and hidden button ball).

6.3 HyFinBall and DIEM-VR Details

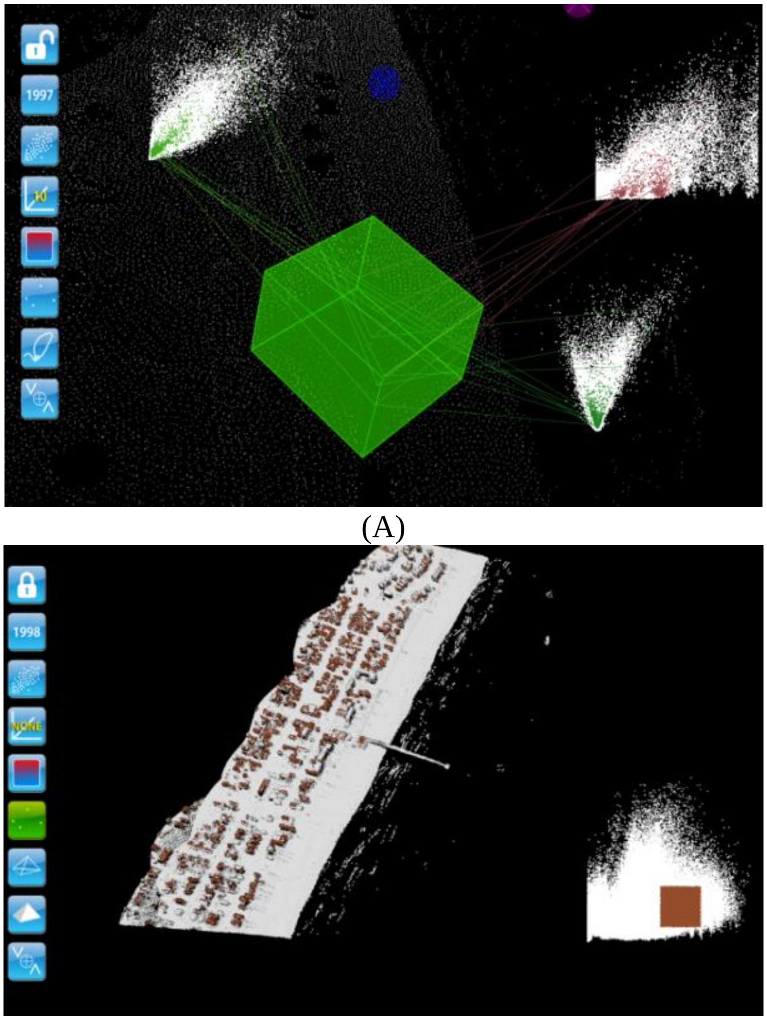
This section discusses the details of the 6DOF, planar-3DOF, and finger-tracking interactions within DIEM-VR and how DIEM-VR demonstrates the 6DOF/3DOF auto-mode switch.

Shaw and Green [129] advocate adding a user adjusted translational offset between the 6DOF button device and the 3D cursor in their two-handed system. This allows the user to keep her elbows resting in her lap, or on the desk or chair arm to combat the common arm fatigue problems in VR interfaces. This offset is part of the 6DOF mode in our system. However, in our prior experimental work with two-handed 6DOF input [146] and in our formative evaluation of the presented HyFinBall interface, we found that while keeping elbows resting on a surface reduces fatigue compared to the naïve 'arms outstretched' approach of early VR systems, this interface is still more fatiguing than using a mouse. With a mouse, the hand and palm—not just the elbow—rests on a surface. Rich data visualizations involve coordinated views of both 2D and 3D components such as in DIEM-VR. Therefore we developed the HyFinBall UI with auto-mode switching between 6DOF and planar-3DOF mode to allow the user to perform one (or two-handed) 3D interactions as well as 2D interactions with the vertical screen while keeping her palm(s) resting on the desk. As we shall explain, in DIEM-VR the 2D scatter plots are intimately tied to the 3D terrain therefore we present these 2D elements on the vertical screen with the 3D terrain while the purely 2D Boolean expression tree remains on the horizontal multi-touch surface. This is a general concept of the HyFinBall UI: pure 2D interactions occur on the horizontal display while 3D interactive graphics and, any 2D interactive graphics intimately tied to the 3D graphics appear on the vertical display.

On the vertical screen, DIEM-VR displays a single patch of terrain which can be optionally color-coded by height or displayed as a wireframe mesh or point-cloud. A series of 2D menu buttons appears on the left of the primary screen. These implement a horizontal, pull “right” menu. All 2D menu items are displayed at zero screen parallax. The user can add and delete multiple scatter-plots whose plot points each correspond to a terrain point. Each plot point's x-y location is determined by a geometric characteristic of the associated terrain point such as the terrain point's average local slope, local degree of roughness, etc. In other words, each original terrain point has several additional geometric characteristics associated with it and by creating scatter-plots along these dimensions, the user can view the terrain in a different feature space such as plotting local roughness versus elevation. The scatter-plots are constrained to the zero-parallax plane.

6.3.1 6DOF 3D Cursors

In its 6DOF mode, the left HyFinBall implements a scene-in-hand metaphor [157] for camera pose manipulation plus separate 3D cursor centered view scaling [30]. In 6DOF mode the left HyFinBall’s virtual representation is a transparent, blue sphere with a user adjustable translational offset [129]. When the left HyFinBall is placed on the desk, planar-3DOF mode is enabled. Now, the HyFinBall’s cursor is replaced by a transparent, 2D blue disc that always remains at zero screen parallax. This cursor interacts like a standard 2D mouse cursor for selecting the menu bar on the left. From our anecdotal observation and several pilot study participants, in the stereo display the switch from the 3D sphere cursor to the 2D disc cursor is immediately apparent.



(A)

(B)

Figure 41: Points selection (A) 3D selection box of terrain as point cloud. And (B) selection of LIDAR points in scatter-plot high-lights house roofs.

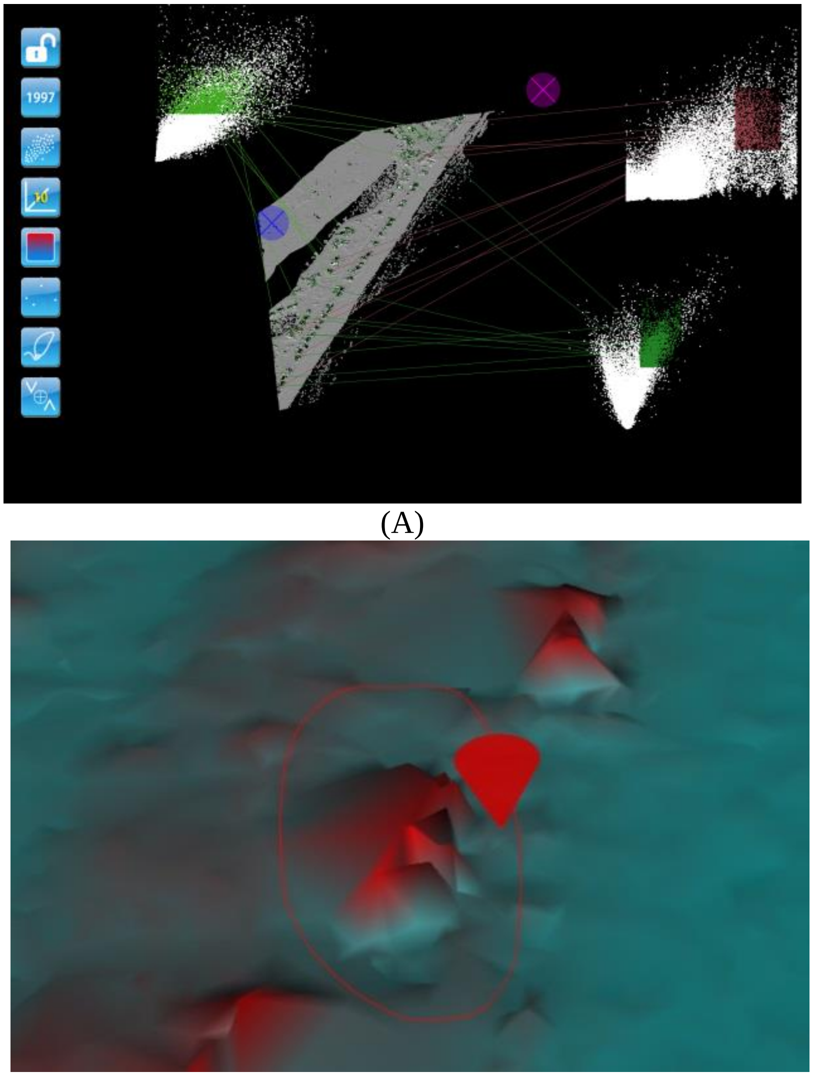
The right HyFinBall's 6DOF mode's 3D cursor is a transparent, purple sphere with a user adjustable translational offset. This 3D cursor implements and initiates 3D selection box creation (Figure 41A). The selection box is used to select points on the 3D terrain. LIDAR scans have multiple returns and are hence multi-planar (not strict height-fields). There are situations where one may want to select points not only within a certain foot-print but also within a limited height range. For example, the user might want to select tree top returns and not the lower layer returns from the underlying ground. While selection in these situations is not as complicated as selection within true volumetric data

[145], we provide a general 3D selection box interface. Further, this general capability for volume selection will be necessary when integrating true volumetric data into the terrain systems as we did in [145]. The 3D selection box can be created, moved, rotated and resized using a technique that is a combination of the two-handed technique of Ulinski et al. [146] and a 3D widget [19].

6.3.2 Planar-3DOF 2D Cursors

When the right HyFinBall is placed on the desk, the 3D cursor is replaced by a transparent, 2D purple disc that remains at zero screen parallax. In this mode, the purple disc acts like a 2D mouse cursor for interacting with any created scatter-plots. When a 2D cursor hovers over a scatter-plot boundary, icons along the x or y axes appear allowing selection of the statistic that will be plotted on the given axis. Various statistics such as average gradient, maximum gradient, and local standard deviation can be selected. The user can move the plot or switch the plot data axis using a button, Button A. The user can select a rectangular region of scatter plot points with Button B. With the 2D cursor, the user can brush points in the scatter-plot. Brushing occurs by creating a rectangular selection region. The selected points are highlighted on the terrain surface using a color pre-assigned to the scatter-plot. In Figure 39B, the scatter-plot in the lower-left plots elevation versus local gradient. The brown selection region is selecting for relatively low elevations with minimal gradient. This causes mostly house roofs to be highlighted in the terrain view.

The user can optionally enable the display of lines connecting the scatter-plot points and the terrain points. This gives a strong visual impression of how the brushed scatter-plot points are spatially distributed on the terrain. (For performance, only a randomly chosen subset of the connecting lines is drawn). Understanding the spatial structure of this “line net” is greatly enhanced by the stereoscopic display. It has some conceptual similarities with traditional 2D parallel coordinates. Figure 42A shows three scatter-plots with line nets connecting their brushed regions to the terrain points. These line nets intimately visually tie the 2D scatter-plots to the 3D terrain and hence keeping these 2D graphics on the same vertical display as the 3D terrain is important.



(A)

(B)

Figure 42: (A) Point-cloud rendering of terrain patch and interactive, coordinated scatter-plot representations of LIDAR points. (B) 2D Lasso Selection.

We assume that during above described 2D interactions with the 2D menu or scatter plots that the user’s eyes fixate on geometry with zero parallax and that the user is not attempting to fixate on geometry with non-zero parallax. (The latter is the condition under which the naïve display of desktop 2D cursor in a stereo 3D system creates problems). Our anecdotal experience indicates this is the case, but future experimentation using an eye tracker is needed. We render the 2D cursors as slightly transparent discs so the user can see through them to any farther 3D geometry. In Figure 42A the left button ball disc is transparent blue and the right is purple.

When displaying these 2D cursors, we automatically reduce the eye separation. If one HyFinBall is in planar-3DOF mode and is performing 2D interaction, then the modeled eye separation is cut in half. If both HyFinBall's are in planar-3DOF mode and performing 2D interactions, eye separation is set to zero. The eye separation changes are animated over a 2s time period recommended by Ware et al [156]. This reduction is again predicated on the assumption that if the user enters planar-3DOF mode they are interacting with the 2D zero-parallax objects and hence fixating at the zero-parallax depth.

We also experimented with enabling a simulation of depth-of-field image blur of the 3D geometry during planar-3DOF 2D interactions. The design space includes the presence/absence of the enabling of depth-of-field simulation and the tradeoff between the fidelity of the depth-of-field rendering and its reduction on frame-rate.

Overall design space issues include presence/absence of the eye separation adjustment, the degree of adjustment, the rate of adjustment, the conditions of adjustment and interaction with depth-of-field implementation. In general, our anecdotal results indicate eye separation reduction is useful when the user is performing planar-3DOF 2D interactions.

6.3.3 Planar-3DOF Projected Cursors

In its planar-3DOF mode, the right HyFinBall can also be used for 2D lasso selection of the terrain points. In this mode, the purple disc is replaced by a different 3D cursor whose 3D position is the intersection of a ray cast from the cyclopean eye through the 2D cursor's computed position on the frustum projection window. In prior work, we used a similar technique where we replaced the display of the desktop 2D mouse cursor with projected 3D cursor. This enabled a mouse controlled travel technique option in our exo-centric, travel technique on stereoscopic virtual workbench [71]. The projected 3D cursor can appear at any screen parallax depending on the location of the intersected terrain under the GUI cursor position. This approach is sometimes referred as geometry-sliding [175].

We chose for the planar-3DOF mode to perform the lasso operation rather than using a 6DOF mode image-plane technique based on the hypothesis the latter would induce greater arm fatigue. During 2D lasso selection we assume the user is fixating on the terrain surface location under the 3D cursor so the eye separation is set to its default setting (see Figure 42B). Our anecdotal experience indicates this assumed fixation point is correct. An experimental evaluation with an eye tracker could confirm this. If the user needs to select a restricted height range, a 3D selection box can be created as described in Section 6.3.1.

Finally, there is an individual terrain triangle selection mode. In this mode the terrain triangle underneath the projected 2D cursor is selected and all other terrain triangles within a range of similar height values are also selected. As the 2D cursor is dragged this selection is continuously highlighted. (Other criteria for selecting 'similar' terrain polygons are, of course, possible).

All these terrain region selections and scatter-plot selections use brushing-andlinking across these coordinated views that are updated in real-time.

6.3.4 Multi-Touch and Finger-tracking

As discussed earlier, when the user tucks the button ball in her palm (Figure 37D and Figure 40D and E), the free fingers such as the thumb and pointer finger can interact with the horizontal multi-touch surface or trigger 3D gestures.

DIEM-VR uses the multi-touch display for the Boolean expression tree once the user creates multiple scatter plots and brushes different regions in each scatter plot. The Boolean expression combines the different selections in various ways to make a final selection where only the terrain points that satisfy the Boolean expression are highlighted in the terrain view. The horizontal multi-touch display shows the tree structure of the Boolean expression (visible in Figure 38 and reproduced in Figure 39B). For example, in Figure 42B, the Boolean expression shows a logical expression of (1 OR 2) XOR (3 AND 4). Numeric labels map elements of the expression to the scatter plot. The user can save the current expression by the (+) menu icon on the right top and an icon is added on the left top. Users can delete, select or modify prior saved expressions. All changes are immediately reflected in the terrain vertex highlighting, the line net display and the scatter plot highlighting.

We specifically chose to touch enable the horizontal display rather than the vertical one, to maintain a palms-resting posture during the multi-touch interaction rather than requiring an outstretched-arm posture that is known in VR to generate user complaints of shoulder fatigue. Further, within the DIEM-VR application the Boolean expression UI is a separate, purely 2D interface unlike the scatter-plots whose line-nets are visually tied to the 3D terrain. Therefore, the Boolean expression UI is highly suited to 2D interaction afforded by the horizontal multi-touch surface.

Again, we can demonstrate hand+finger-tracking while still holding the button ball using 3Gears Kinect based tracking and we integrated a ray-based 3D selection in DIEM-VR using 3D pointing gestures. However, we found the current tracking range and error rate of the hand+finger-tracking to be prohibitively restrictive when trying to pilot test a user study that integrates them with the rest of the HyFinBall UI. For instance, the Polhemus' electromagnetic (EM) tracking of the HyFinBall's never drops out the way it can with the Kinect based tracking and the "error rate" of detecting mechanical button presses is essentially zero. This discrepancy led pilot test participants to want to use the button balls instead of 3D finger-tracking for any practical 3D user tasks such as object selection or manipulation. Nonetheless, because the concept of enabling 3D hand+finger tracking while still holding the button balls is at least demonstrable, we present it as part of the overall HyFinBall UI.

6.4 Design Motivations and Empirical Questions

This section discusses several of the key design considerations and motivations for the HyFinBall interface and as well a number of interesting questions that will require an empirical study. The study focuses only on the 6DOF and planar-3DOF combination. The three devices device conditions are:

I. *Auto-switching HyFinBall UI*
II. *dual planar-3DOF mode only UI*
III. *dual 6DOF mode only UI*

This comparison is done across a variety of 2D and 3D tasks in different sequential combinations. This study does not use the horizontal multi-touch display and the tasks involve 3D terrain manipulation and selection and 2D menu and scatter plot manipulation. The goal is to determine to what degree each of the four device conditions is better suited to pure 2D tasks, pure 3D tasks and to combination 2D and 3D tasks. (To details, see Section 6.5).

6.4.1 Fatigue – Elbows-Resting vs. Palms-Resting

As stated earlier, using the planar-3DOF mode for 2D interactions on the vertical display is motivated by the desire to avoid arm fatigue issues that would arise if the user had to instead use image-plane techniques with the 6DOF mode. Image-plane techniques would require hovering the 3D cursor over the image of the 2D menus or scatter plots to manipulate them. Our experiment tests this hypothesis by comparing user subjective reports of fatigue when doing purely 2D tasks using condition *III*, 6DOF image-plane techniques, and condition *II*, planar-3DOF mode. When the user task is a mix of 2D and 3D tasks, we also expect condition *III* to be more fatiguing than condition *I*, the auto-switching HyFinBall mode, because the auto-switching mode allows the 2D operations to be performed with resting palms. Of course, there is a trade-off with condition *I*, since the user must switch between a palm-resting posture and an elbow-only resting posture in order to switch between 2D and 3D operations.

The overall effectiveness of the 6DOF/planar-3DOF auto-switching will undoubtedly ultimately depend on the balance between the 2D and 3D interaction operations used in a given application and the temporal sequencing and durations of planar-3DOF interactions and 6DOF interactions. Our in-progress experiment is a first step in exploring this. Our anecdotal observations, indicate that users perform better and very much prefer condition *I* or *III* over *II* when the task includes 3D navigation and 3D manipulation of a 3D selection box.

6.4.2 Auto-Switching 2D and 3D Cursors

Section 6.3 described how the HyFinBall UI uses 3D cursors and several types of 2D cursors within a stereoscopic environment. There has been a fair amount of prior work in desktop 2D GUI's regarding having the 2D image of the cursors change to indicate different application states or interaction modes. There has been interesting work in cursors for 3D selection such as Ware and Lowther's One-Eyed cursor [159]. Teather and Stuerzlinger compared four cursor selection techniques in [142], and more recently Bruder et al. [27] explored different offset techniques on a virtual workbench. The HyFinBall raises additional questions because the cursor automatically switches between a 6DOF 3D cursor, a 2D zero-parallax cursor, and a projected 3D cursor (as in HyFinBall 2D lasso mode).

6.4.3 Multi-Touch and Finger-Tracking

Our current implementation of the HyFinBall UI demonstrates the possibility of leveraging the button ball form factor to allow multi-touch and hand+finger tracking interaction without dropping the device. The multi-touch UI is robust enough to consider formal user studies, but the tracking range limitations and 3D gesture error rates of the Kinect-based tracking still need improvement.

At the moment we can only speculate about design issues and questions that could be investigated with more robust 3D hand+finger tracking. If the 3D finger tracking and gesture recognition were as robust as the simpler EM tracking and buttons, it would be interesting to explore what interactions 3D are best performed with hand+finger tracking and what are best performed with the 6DOF button balls. Moehring and Froehlich performed a study using very robust and accurate hand and finger tracking (with Vicon [148] marked gloves) and compared this with a 6DOF held-device (a Flystick) for a series of 3D manipulation tasks [98]. Users preferred the naturalness of finger tracking. However, users of the Flystick performed significantly faster than "bare" finger tracking.

Adding pinch-sensitive finger tracking improved task performance times to be within 10-20% of the Flystick condition. This suggests that manipulating a physical object (such as a button ball) may prove advantageous for some 3D object manipulation tasks over 3D hand+finger tracking within our demonstrated HyFinBall interface. From a practical standpoint it is a bit challenging to test this because the systems that provide robust hand+finger tracking require wearing gloves or thimbles with fiducial markers which may make simultaneously handling a button ball cumbersome.

6.5 Experiment Design

Experiment uses a subset of ITs within DIEM-VR. The three IT conditions are HyFinBall (with auto-switching), 6DOF-only mode, and 6DOF+mouse (dual 6DOF button balls with a mouse). In Experiment 1, DIEM-VR displays five menu icons, and one patch of terrain, and a scatter plot.

Experiment's 3D tasks include 7DOF navigation (pose + scale) and creation and manipulation of 3D selection boxes [146]. In a trial, the user is prompted with a red 3D selection box on the terrain, the user must navigate and create 3D selection box creation to make a matching selection box. Considering all three IT conditions use same 6DOF button ball UI for the 3D task, we set large tolerance (±10% of the cursor box size) to make the 3D task easier to be accomplished. For Experiment 1, both multi-touch and gesture interface are disabled.

2D tasks of Experiment 1 are as follows. Task 1 is selecting a menu icon. A red outlined rectangle appears on one of menu icons to prompt the user to select the icon. Task 2 is selecting axis icons of a scatter plot. An outline appears on one of a scatter plot's data axis icons. The user selects this icon to change the variable plotted on that axis. Task 3 is relocating a scatterplot. A red outlined scatterplot appears in a random position, and then the user moves the target scatterplot to a position indicated by an objective rectangle. Task 4 is brushing points of a target scatterplot. A red rectangle appears inside the target scatterplot bounding a set of plot points, the user needs to draw selection rectangle matching the red target. Each 2D task trial presents four tasks, one of each of the above types, in sequence. The order of 2D task presentation is randomized per trial.

We also perform a within-subject pilot study with five different UI conditions (Mouse only, planar constraint only, 6DOF only, 6DOF+Mouse and HyFinBall). For each of the two participants (non-authors), the experiment took slightly over 2 hours. The qualitative and quantitative results indicated that 3D task completion times for the 2D UIs (mouse (M=65.89) and planar constraint button ball (M=100.88)) are significantly slower than 3D UI (i.e. button ball, M=34.94). This is similar to previous 2D UI studies for 3D interaction [127,131] that show relatively poor performance of 2D UIs for complex 3D interactions. The 2D UI's used standard mouse based 3D navigation techniques (Arcball [131]). The more difficult task of creating and adjusting a 3D selection box was significantly more difficult to perform both with the mouse UI and the dual planar-3DOF UI than with the two-handed 6DOF interface. This was despite several iterations of re-design of the mouse and planar-3DOF 3D selection box interfaces. With a two plus hour experiment, users complained of visual and physical fatigue. They complained heavily regarding having to perform the 3D selection box task with 2D UI conditions as compared to the much more favored 3D UI conditions for the 3D tasks. Due the large quantitative reduction in performance for our 3D tasks when using 2D UI conditions and due to the empirical difficulties in participants performing all five conditions, we therefore removed the 2D UI only conditions from the full Experiment 1.

6.6 Experiment

Twelve participants performed 30 trials each (10 trials (10 × 3D task and 10 × 2D task) × 3 UI conditions) in a within-subject comparison through 6DOF+mouse, 6DOF only and HyFinBall UIs for 3D and 2D tasks. Eleven participants are from the Computer Science department and one from Public Health Sciences; seven are males and five are females. All participants have (corrected) 20/20 or higher eye sight, no disability using their hands and fingers, and passed a stereopsis test. All participants have high daily computer usage (6.25 out of 7). Eleven participants have experience of 6DOF button ball UI from a previous 1 hour user study. The three UI conditions were presented in counter-balanced order across all participants.

We evaluated four different types of tasks: pure 3D task, pure 2D task, and cross-dimensional tasks—a 2D task followed by 3D task (3D-to-2D) and 3D task followed by 2D task (2D-to-3D). For data of cross-dimensional tasks, we averaged all sequential tasks that contain 2D to 3D or 3D to 2D tasks.

We use repeated measures (*rm*) ANOVA. The F tests that are reported use α=.05 for significance and indicate the Geisser-Greenhouse correction to protect against possible violation of the assumption of the sphericity. The post-hoc tests that are conducted are least significant differences (LSD) tests with α=.05 level for significance.

The primary hypotheses are:

H1. Overall, HyFinBall is expected to have faster task completion time than other UIs.

H2. HyFinBall and 6DOF+Mouse are expected to have faster completion time than 6DOF button ball only for 2D tasks.

H3. HyFinBall is expected to have faster completion time for a 3D task than all *other UIs.*

H4. HyFinBall is expected to have faster completion time for cross-dimensional tasks than all other UIs

6.6.1 Result

Quantitative

Overall

Table 10: Average and standard deviation of task completion time of different UIs

	6DOF+Mouse		6DOF only		HyFinBall	
	CT	SD	CT	SD	CT	SD
3D	39.3	10.1	39.8	15.7	33.9	7.7
2D	22.8	1.9	29.3	6.1	21.5	3.7
3D to 2D	63.8	13.2	68.6	19.7	55.8	9.5
2D to 3D	63.9	12.3	69.4	17.4	54.9	8.5

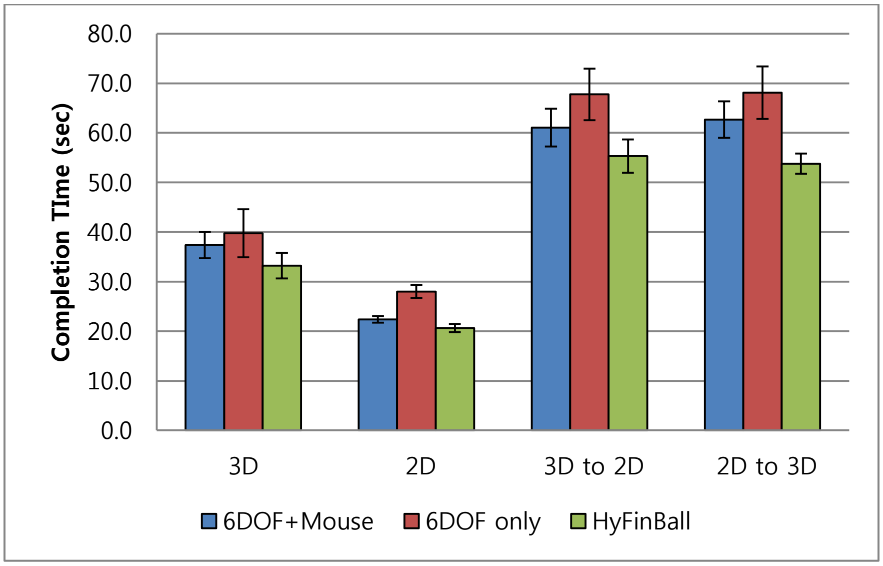


Figure 43: Task completion time of UI condition by Tasks.

The 3 × 3 × 2 *rm* ANOVA (Order × UI × TaskType) for pure 2D and 3D tasks shows no interaction effect on completion time of UI condition and UI presentation order

(p=.315).

The 3 × 2 *rm* ANOVA (UI × TaskType) for pure 2D and 3D tasks shows no interaction effect on completion time of UI condition and task type (p=.165). This indicates no performance change between 3D and 2D tasks across the three UIs. There is a main effect on completion time of UI condition (F(2,22)=7.857, p=.008, η_p^2=.417). LSD post-hoc tests show completion time of HyFinBall (M=27.72, SD=8.66) is faster than 6DOF+Mouse (M=31.08, SD=11.03, p=.021), and 6DOF only (M=34.57, SD=12.82, p=.010). In addition, completion time of 6DOF+Mouse is faster than 6DOF-only (p=.049). As expected (hypothesis *H1*), HyFinBall has the better performance.

The one-way *rm* ANOVA for the 3D task shows no main effect on completion time of UI condition (p=.088). A possible explanation is that the overhead time of switching modes (in HyFinBall) and switching physical devices (in 6DOF+mouse) did not reduce performance for the 3D tasks. That is the size of these overheads is relatively small compared to the 3D task completion time (*3D* in Figure 43).

The one-way *rm* ANOVA for the 2D task shows a main effect on completion time of UI condition (F(1.359,14.944)=11.839, p=.002, η_p^2=.518). Completion time of 6DOF-only is slower than HyFinBall (p=.003) and 6DOF+Mouse (p=.005) (see 2*D* in Figure 43). This indicates 2D task performance of the planar-3DOF sub-mode in the HyFinBall condition and the mouse sub-mode in the 6DOF+mouse condition, are better than the 6DOF-only condition (hypothesis *H2*). This also indicates 2D task performance of planar-3DOF is no worse than the mouse input. We observed in the 6DOF-only condition, most participants had difficulty holding the button ball in a fixed position in the middle of the air when they were performing 2D tasks with image-plane ITs despite the fact they could rest their elbows.

The 3 × 2 *rm* ANOVA for cross-dimensional tasks (3D-to -2D and 2D-to-3D) shows a main effect on completion time of UI condition (F(2,22)=7.606, p=.003, η_p^2=.409). Completion time of HyFinBall (M=55.31, SD=8.82) is faster than both 6DOF+Mouse (M=63.96, SD=12.5, p=.010) and 6DOF-only (M=69.02, SD=18.17, p=.007). 6DOF-only and 6DOF+Mouse do not differ (p=.155).

The one-way *rm* ANOVA for the 3D-to-2D task shows a main effect on completion time of UI condition (F(1.376,15.131)=6.571, p=.015, η_p^2=.374). Completion time of HyFinBall is faster than both 6DOF+Mouse (p=.010) and 6DOF-only (p=.018). However, there is no difference on completion time between 6DOF-only and 6DOF+Mouse (p=.169).

The one-way *rm* ANOVA for the 2D-to-3D task show a main effect on completion time (F(2,22)=7.359, p=.004, η_p^2=.401). Completion time of HyFinBall is faster than both 6DOF+Mouse (p=.013) and 6DOF-only (p=.008). Same as the result of the 3D-to-2D task, there is no difference between 6DOF+Mouse and 6DOF-only (p=.172). The HyFinBall UI has the best task performance for cross-dimensional tasks. This supports hypothesis *H4* that the overhead of switching HyFinBall sub-modes between 3D and 2D takes less time than changing physical input devices in the 6DOF+Mouse UI. Even though 2D task performance of the mouse input is better than 6DOF-only, the device acquisition time penalty may reduce overall performance of 6DOF+Mouse for cross-dimension tasks.

We recorded the number of 3D selection boxes create, the number of scatter plot moves, the number of button presses for navigation, and the number of button presses for modifying the 3D selection box to evaluate effectiveness of different UIs for the 3D task. However, there are no statistical main effects of UI condition on those variables.

2D tasks

Table 11: Average and standard deviation (SD) on completion time (CT) of different UIs for 2D tasks

	6DOF+Mouse		6DOF only		HyFinBall	
	CT	SD	CT	SD	CT	SD
SM	3.5	0.5	2.6	0.4	2.9	0.5
SA	5.4	1.1	6.0	1.6	5.5	0.6
RS	6.6	0.9	7.0	1.1	5.3	1.1
BP	7.3	0.8	13.2	5.6	7.9	2.3

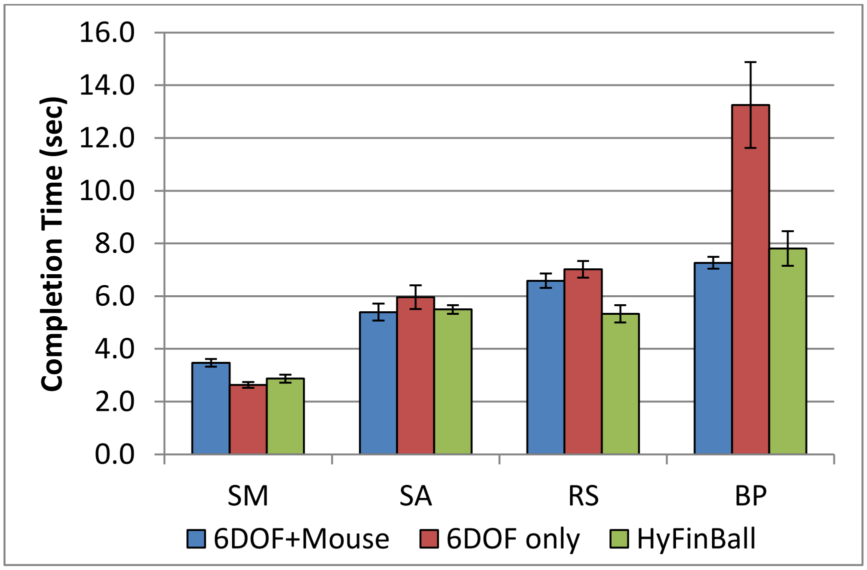


Figure 44: Task completion time of UI conditions by 2D tasks

When considering 2D tasks separately, the 3 × 4 *rm* ANOVA (UI × Task) shows an interaction effect of UI condition and 2D tasks (F(1.533,16.867)=9.773, p=.003, η_p^2=.470). This indicates UI condition has different effects on completion time depending on the particular 2D task. Generally, simple tasks (selecting a menu (SM) and selecting axis icons (SA)) have small differences across UIs. However, precise tasks (relocating the target scatterplot (RS) and brushing points (BP)) have larger differences across UIs (see Figure 44 and Table 11).

There is a simple main effect on completion time of UI condition for the SM task (F(2,22)=9.618, p=.001, η_p^2=.466). Completion time of the mouse is slower than the 6DOF (p=.001) and planar-3DOF button balls (p=.011) (see SM in Figure 44). However, there is no difference between 6DOF-only and planar-3DOF button balls (p=.289). This indicates the benefit of using dual planar-3DOF input rather than a single mouse for the 2D tasks. This is inline with research comparing dual mice UI's to single mouse UI's.

There is no simple main effect on task completion time of UI condition for the SA task (p=.352). The position of a scatter plot is random in the SA task. Normally the user moves cursors of dual mice from left and right of the display. The distance from the cursors of dual mice to a target scatterplot should be similar with single mouse.

There is a simple main effect on completion time in the RS task (F(2,22)=10.917, p=.001, η_p^2=.498). Completion time of planar-3DOF is faster than both mouse (p=.003) and 6DOF button ball (p=.003) (see RS in Figure 44). However, completion time of the mouse and 6DOF button ball do not differ (p=.223). This clarifies the benefit of dual planar-3DOF button balls over a single mouse and the general advantage of a 2D UI for a precise 2D task (3DOF button ball or mouse vs. the 6DOF image-plane ITs).

The results also show a simple main effect on completion time for the BP task, (F(1.257,13.829)=10.437, p=.004, η_p^2=.487). Completion time of 6DOF-only is slower than the mouse (p=.009) and the planar-3DOF (p=.005). However, completion time of mouse and planar-3DOF does not differ (p=.453) (see BP in Figure 44). Since the BP task requires relatively precise manipulation, mouse and planar-3DOF have better performance than 6DOF-only.

There is a main effect on 2D task completion time of 2D task types (F(1.352,14.873)=69.241, p<.001, η_p^2=.863). Completion time of SM is faster than other 2D tasks (p<.001), SA is faster than RS (p=.007) and BP (p<.001), and RS is faster than BP (p<.001). This is because BP requires more accurate control and specifying 2 points.

2D Task precision

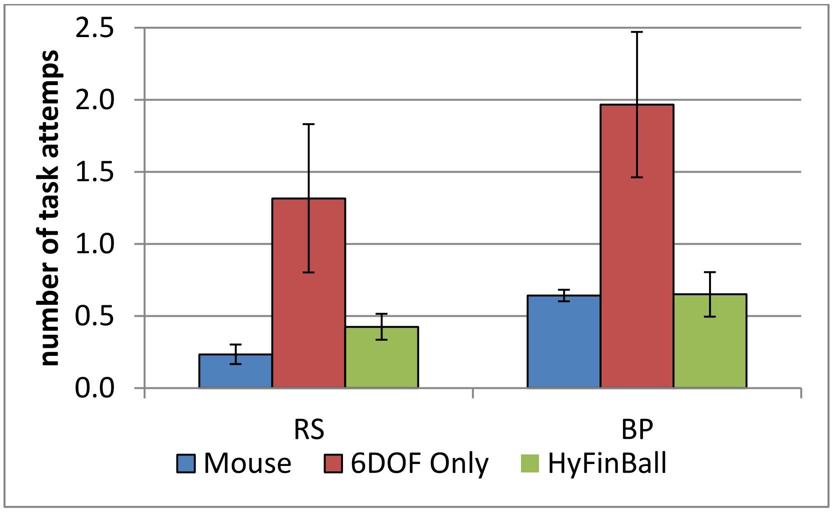


Figure 45: Number of task attempts of UI conditions for RS and BP tasks.

The 3 × 4 *rm* ANOVA to determine usability of UI condition and 2D task on number of task attempts shows an interaction effect on number of task attempt of UI condition and 2D task (F(1.415,15.567)=6.862, p=.012, η_p^2=.384). Similar to the interaction effect of completion time between UI condition and 2D tasks, this explains the user tried more times to accomplish tasks differently depending on task difficulties.

The results show a simple effect on number of task attempts in the RS task (F(1.308,14.391), p<.001, η_p^2=.717). LSD tests show number of task attempts of the 6DOF button ball (M=1.32, SD=0.51) is significantly more than planar constraint button ball (M=0.43, SD=0.31, p=.001) and the mouse (M=0.23, SD=0.23, p<.001) (see RS in

Figure 45). However, there is no difference on number of task attempts between mouse and planar constraint button ball (*p*=.107).

The results also show a simple effect on number of task attempts in the BP task (F(1.155,12.702)=6.56, *p*=.021, η_p^2=.374). Number of task attempts of 6DOF button ball (M=1.93, SD=1.75) is significantly more than planar constraint button ball (M=0.65, SD=0.53, *p*=.027) and the mouse (M=0.64, SD=0.14, *p*=.020). There is no statistical difference on number of task attempts between mouse and planar constraint button ball (*p*=.960) (see BP in Figure 45).

There are no simple effects on task attempts in SM and SA tasks. Evidently 6DOF button ball incurred more task attempts then other inputs for precise 2D tasks.

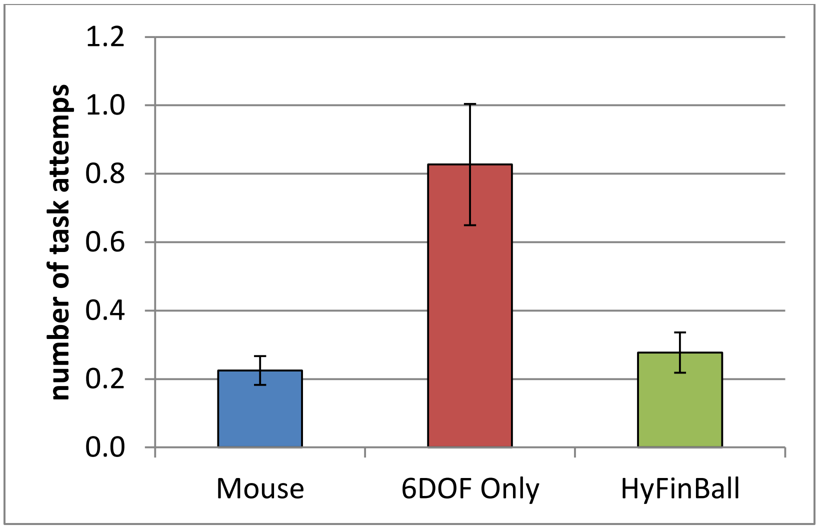


Figure 46: Number of task attempts of 2D UI condition.

There is a main effect on number of task attempts of UI condition (F(1.172,12.894)=13.783, *p*<.001, η_p^2=.556). LSD comparisons show number of task attempts of 6DOF button ball (M=0.83, SD=1.23) is significantly more than the mouse (M=0.23, SD=0.29, *p*=.001) and 3D planar constraint button ball (M=0.28, SD=0.40, *p*=.006) (see Figure 46). This main effect is because 6DOF button ball caused more task attempts on SM and SA tasks.

Percentage of pressed wrong button

Table 12: Average and standard deviation (SD) on 2D task completion time (CT) of different UIs

	6DOF+Mouse		6DOF only		HyFinBall	
	CT (%)	SD (%)	CT (%)	SD (%)	CT (%)	SD (%)
SM	0.0	0.0	15.8	16.2	19.2	22.7
SA	7.5	7.5	30.8	24.7	13.3	8.7
RS	17.5	17.1	13.3	13.0	11.7	14.7
BP	2.5	4.5	30.0	22.6	7.5	8.7

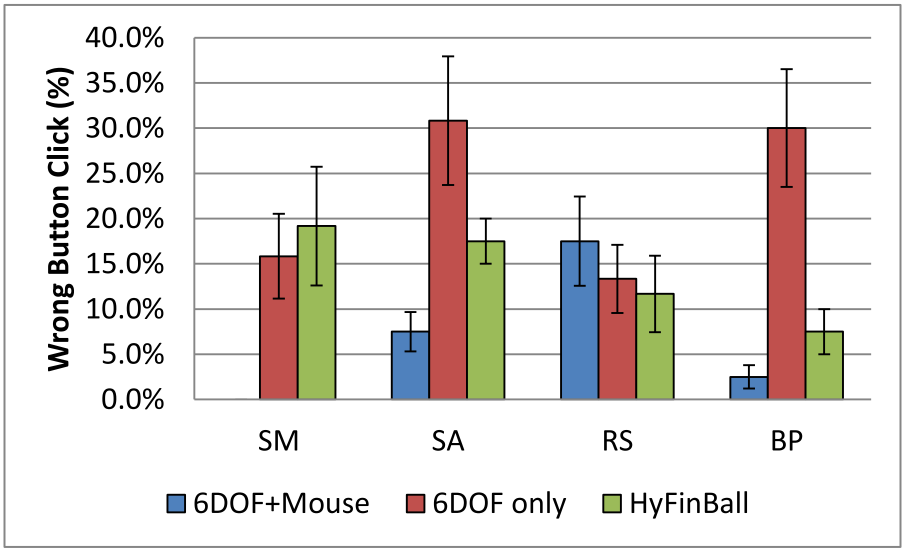


Figure 47: Percentage of wrong button presses of UI conditions by 2D tasks.

We recorded how many trials the user pressed wrong button for 2D tasks. Table 12 show average and standard deviation of percentage of pressed wrong button for 2D tasks by UI condition. The one-way *rm* ANOVA for the SM task shows a main effect on percentage of trials of pressing wrong button on UI condition (F(2,22)=7.008, *p*=.004, η_p^2=.389). Contrasts reveal the mouse has significantly less percentage than both 6DOF (F(1,11)=8.52, *p*=.014, η_p^2=.436) and planar constraint button balls (F(1,11)=11.444,

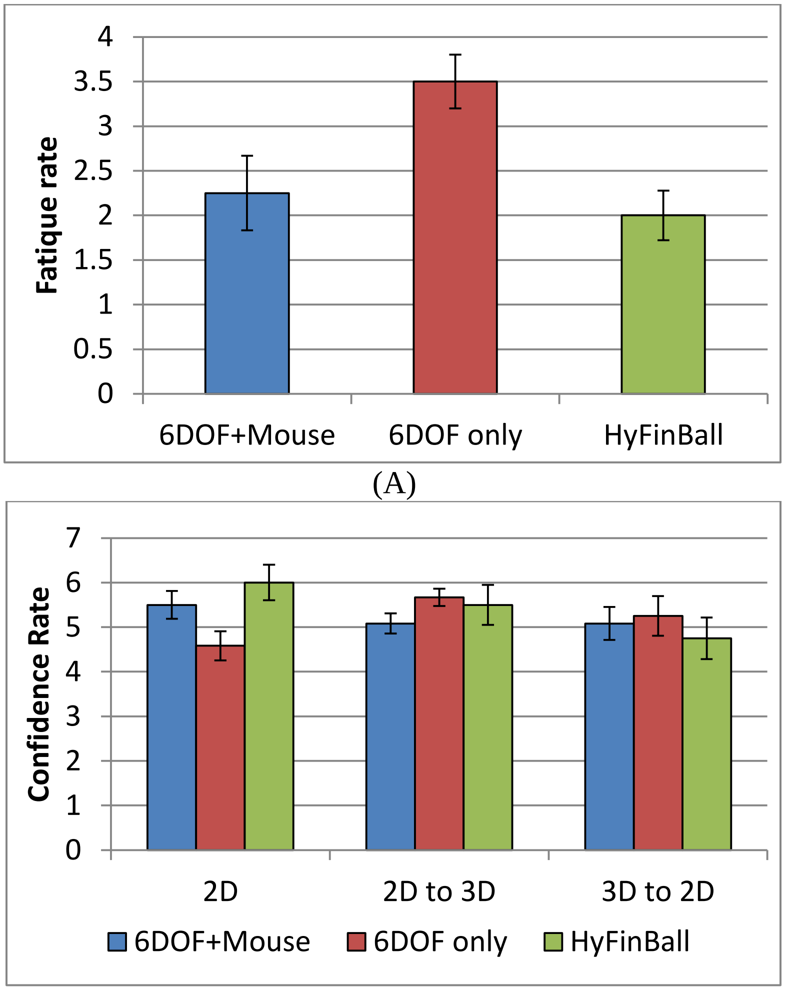
p=.006, η_p^2=.510) (see SM in Figure 48). However, there is no significant difference between 6DOF and planar constraint button balls (p=.517).

The one-way *rm* ANOVA for the SA task indicates a main effect on percentage of wrong button clicks of UI condition (F(1.284,14.127)=11.629, p=.003, η_p^2=.514). Participants pressed wrong buttons with the mouse in fewer trials than both 6DOF (F(1,11)=16.5, p=.002, η_p^2=.600) and planar constraint button balls (F(1,11)=16.09, p=.002, η_p^2=.594) (see SA in Figure 48). In addition, planar constraint button ball has less percentage of pressed wrong buttons than 6DOF button ball (F(1,11)=5.77, p=.035, η_p^2=.344).

There is no simple effect on percentage of wrong button clicks of UI condition in the RS task (p=.294). Interestingly, this is the task which uses a right mouse button. Users may be familiar with the left mouse button because it is the main button for most computer tasks.

The one-way *rm* ANOVA for the BP task shows a main effect on percentage of wrong button clicks of UI condition (F(1.333,14.664)=13.329, p=.001, η_p^2=.548). Participants pressed wrong buttons more with 6DOF button ball than planar constraint button ball (F(1,11)=12.789, p=.004, η_p^2=.538) and the mouse (F(1,11)=16.036, p=.002, η_p^2=.593) (see BP in Figure 48). There is no difference on percentage of wrong button click between the mouse and the planar constraint button ball (p=.139).

Qualitative



(A)

(B)

Figure 48: (A) Fatigue and (B) confidence rate of the user by different UI and tasks.

On a 7-point Likert scale, user rating of arm fatigue (1 *no fatigue* through 7 *very painful*) is significantly different through UI conditions. The one-way ANOVA shows that there is a main effect on rating of arm fatigue of UI condition (F(1.289,14.176)=10.333, p=.004, η_p^2=.484). Participants felt more fatigue with 6DOF only condition (M=3.5, SD=1.45) than both 6DOF+Mouse (M=2.25, SD=.97, p=.033) and HyFinBall (M=2.0, SD=1.04, p=.013) conditions. However, arm fatigue rating is not significantly different between 6DOF+Mouse and HyFinBall conditions (p=.537). Figure 48A illustrates fatigue rates through UI conditions. Possible explanation is that users could not rest their arms and hands with 6DOF UI condition while they are performing 2D tasks.

On a 7-point Likert scale, user rating of how accurately perform 2D task (1 *no accuracy* through 7 *very accuracy*) is significantly different among UI conditions (F(2,22)=4.602, p=.021, η_p^2=.295). Users felt more confidence to perform 2D task more accurate with the HyFinBall UI (M=6.0, SD=1.13) than the 6DOF only UI (M=4.58, SD=1.08, p=.028). Subjective confidence rating of 6DOF+mouse (M=5.5, SD=1.38) is slightly higher than 6DOF only UI but not significant (p=.307). On the same 7-point Likert scale, user rating of how accurately perform 2D to 3D and 3D to 2D tasks, are not different with 5.67 and 5.25 for 6DOF only, 5.08 and 5.08 for 6DOF+Mouse, and 5.5 and 4.75 for HyFinBall (see Figure 48B).

Ten out of twelve participants answered they thought HyFinBall is the easiest UI for overall task, two answered 6DOF+Mouse UI. For only 2D task, eight users answered that mouse is the easiest and four answered planar constraint button ball is the easiest. For 2D task followed by 3D tasks, eight participants answered HyFinBall is the easiest UI, four answered 6DOF only. For 3D task followed by 2D tasks, nine answered HyFinBall is the easiest UI, three answered 6DOF only.

When asked whether how much participants prefer planar constraint button ball or mouse for 2D tasks (-2: strongly prefer mouse, -1: somewhat prefer mouse, 0: neutral, +1: somewhat prefer planar constraint button ball, +2: strongly prefer button ball), four participants answered they strongly preferred planar constraint button ball, five answered they somewhat preferred planar constraint button balls, one answered strongly preferred mouse, and two answered they somewhat preferred mouse input (total score through all twelve participants is +9).

On a 7-point Likert scale regarding whether switching between holding the mouse and holding the button ball affected participant's ability to complete task (*1 not at all* through *7 very much*), the average rating is 4.25. On the same Likert scale rating regarding whether switching between holding the mouse and holding the button ball affected participant's physical comfort while completing the task, the average rating is 3.25. On the same Likert scale rating regarding how frustrating is switching between holding the mouse and holding the button ball, the rating is 3.58.

Seven of twelve participants answered that 6DOF only UI is the most difficult to learn for overall tasks, two answered HyFinBall, three answered 6DOF+Mouse. For 2D tasks, nine answered 6DOF only UI is the most difficult to learn, one answered HyFinBall, and two answered 6DOF+Mouse. For 3D task, five answered 6DOF only is the most difficult, five answered 6DOF+Mouse and two answered no preference.

6.7 Discussion and Conclusion

This chapter presents our two-handed hybrid UI, HyFinBall, for interaction with rich, multi-dimensional visualizations that require coordinated 3D and 2D views in semi-immersive desktop VR. The HyFinBall concept is implemented within a specific visualization tool called DIEM-VR for analyzing terrain meshes. These interaction techniques can be used with not only geospatial data but also with other scientific or medical datasets such as volume datasets with 2D interactive transfer functions or feature spaces representations. Potentially more abstract, less physically based 5-dimensional datasets could be explored by separating the dimensions to into various linked 3D and 2D visualizations. We suggest the HyFinBall UI could be a useful UI approach for multi-dimensional visualizations whose component data dimensions are each best visualized using a different display plus input hardware combination suggesting a cross-dimensional approach.

HyFinBall with multi-touch works robustly, but we have yet to formulate user studies. HyFinBall with 3D hand/finger tracking is not yet robust enough to formally evaluate. Future multi-Kinect configurations may solve the problems with gesture recognition error rate and limited tracking range. Alternatively robust marker based hand and finger tracking could be employed. We believe there is an interesting design space to be explored when combining the HyFinBall button balls with robust 3D finger tracking. Finally, based on the ALCOVE [96] and Houtgast's work in [71], we are in the process of configuring our system into a more seamless L-shaped display that also displays stereo 3D on the horizontal surface. Several of the prior works mentioned in Chapter 2 have begun exploring stereo plus multi-touch, but to our knowledge prior work is limited.

The experiment shows HyFinBall UI performs faster over for cross dimensional tasks then 6DOF only and 6DOF+Mouse UIs. Apparently, the device acquisition time penalty decreases performance of cross dimensional tasks while the overhead of switching input modes does not. In addition, planar constraint button input has similar results for 2D tasks to the mouse input because of benefit of dual mice and ability to do precise tasks. 6DOF button ball input does not provide good performance for 2D tasks. However, there are no differences through UIs for the 3D task. However, the user presses wrong button more with HyFinBall UI than the mouse input for 2D tasks. In subjective questionnaires ten users prefer the HyFinBall UI and two prefer the 6DOF+Mouse UI. Additionally, users prefer the planar constraint button ball input than the mouse input for 2D tasks.

Some participants indicate that the ball shape is not good to use as a mouse input.

They feel more comfortable when they are holding the button ball for the 3D task. Possibly it would be better if the ball has flat back face to be more stable for planar constraint input. Future study is needed to evaluate effects of form factor differences for both 2D and 3D tasks. Some also mentioned about button layouts. In the planar constraint button ball mode, some participants have problems to press the button correctly because they try to press buttons with only thumb rather than index and middle fingers. In addition, resolution of the planar constraint button ball input is not good as one of the mouse input. If the planar button ball has better sensitivity and resolution then 2D performance of the planar constraint button ball input would be significantly better.

Overall, the sequence and duty-cycle of 2D versus 3D tasks is important for the usability of HyFinBall. If the user needs to do many 2D tasks and rare 3D tasks, then the 6DOF+Mouse UI would be better than HyFinBall. However, if the task required frequent mode changes, then HyFinBall would dominate other UIs. Furthermore, the HyFinBall has a lot of potential to enrich interaction for multi-dimensional visualization due to its ability to use finger and hand gestures with advantage of physical 3D and 2D inputs.

CHAPTER 7: DESIGN GUIDELINES

The following design guidelines have been created based on the evaluation of VR displays, user interfaces and interaction techniques. The most important design guidelines that have been determined from this body of research are the following:

- Stereoscopic and head-tracking display can enhance depth perception of 3D volume datasets when the user needs to judge the spatial structure of volume datasets. Stereopsis alone is more effective than head-tracking alone for depth judgment.
- To minimize completion times and button clutches for 7DOF navigation tasks, use a bimanual simultaneous 7DOF interaction technique (Spindle+Wheel). In contrast, for 6DOF navigation tasks, use 6DOF (position + orientation) interaction techniques to minimize completion times. The 7DOF navigation technique should be designed such that the 6DOF IT is available for when the task does not require any scale changes to maximize users' performance.
- Auto-stereo adjustment techniques should be considered for reducing stereoscopic fusion problems in particular in multi-scale virtual environments. In desktop VR, the techniques also can enhance the user's performance for navigation tasks with the one-handed navigation technique when tasks require large amount of scale changes.
- For multi-dimensional visualization, hybrid user interfaces can enhance user's performance when tasks require mode changes between 2D and 3D

tasks frequently.

The following are VR display designs that affect depth perception of volumetric data:

- Head-coupled stereoscopic display is best for depth judgment accuracy for volume datasets.
- Stereopsis alone enhances depth perception for volume datasets in both short and unlimited exposure time.
- Head-tracking alone enhances depth perception if data exposure time is unlimited. However, it has no significant effect on depth judgment in short exposure time.

The following are interaction designs that affect navigation task completion times in MSVEs:

- The Spindle+Wheel interaction technique is best for fastest task completion times, when the task requires scale changes.
- Continuous 6DOF interaction techniques are better than a continuous 7DOF IT when the task does not require any scale changes.

The following are user interface designs that affect manipulation task completion times in multi-dimensional visualizations:

- 2D input mode (a mouse or planar constraint button ball) are best for fastest task completion times for pure 2D tasks.
- HyFinBall mode is best for fastest task completion times for combination tasks (i.e. 3D-to-2D and 2D-to-3D tasks)

The following are dynamic adjustment of the stereo view parameter technique designs that affect navigation task completion times in multi-scale virtual environments:

- The dynamic adjustment of the view location technique reduces completion time more than the dynamic adjustment of the stereo view scale factor technique and no dynamic adjustment in desktop VR with the one-handed navigation technique when the task requires large amount of scale changes.
- The dynamic adjustment of the view scale factor technique increases task completion time in a CAVE with the one-handed navigation technique when the task only requires small amount of scale changes.

The following is dynamic adjustment of the stereo view parameter technique designs that affect stereoscopic fusion problems in multi-scale virtual environments:

- Auto-stereo adjustment techniques significantly reduce stereoscopic fusion problems when using the one-handed navigation technique in both desktop VR and a CAVE.

CHAPTER 8: CONCLUSION AND FUTURE WORK

This dissertation has addressed various research challenges for stereoscopic HTDs. Chapter 3 presented effects of stereoscopic and head-tracking displays for human depth perception for 3D volumetric datasets by depth discrimination and depth ordering experiments. The results of both experiments showed benefit of stereo displays which enhances human depth perception. Still future experiments are needed to evaluate motion and stereo with a wider variety of real world volume datasets.

Chapter 4 presented a simultaneous 7DOF object manipulation technique. We conducted two experiments comparing it with a previous two-handed, object manipulation technique. The chapter also compared it against a one-handed technique and showed benefits of the two-handed IT for 7DOF manipulation. In addition, we examined the effect of accidental scale changes and of user's preference between rate control and position control for the one-handed IT. We are currently testing an additional hybrid technique that combines aspects of the unimanual and bimanual ITs for 7 DOF manipulation.

Chapter 5 evaluated stereo auto-adjustment techniques for MSVEs in the desktop VR and CAVE. A user study demonstrated the advantages and disadvantages of two auto-adjustment techniques. The experiments indicated that we need more sophisticated rules for dynamic activation of auto-adjustment techniques in order to maximize their positive effects and minimize their negative effects during varying types of travel tasks. The results also suggest that for multi-screen displays, such as the CAVE, gaze tracking appears necessary for determining which screen should be used for the auto-adjustments.

Chapter 6 presented a bimanual, hybrid user interface for the semi-immersive desktop VR system which combines traditional physical input devices and natural human hand input together for a multi-dimensional visualization application. Our experiments showed advantages of the hybrid user interface for the cross-dimensional tasks in comparison with manually switching between devices and with using only one device. Future study should focus on further hybrid interaction techniques which use human natural hand and finger inputs with traditional tracking devices. While we demonstrated the integration of this technology in our DIEM-VR application, evaluation remains future work.

REFERENCES

[1] 3Gear System. http://www.threegear.com/.

[2] Agrawal, P., Rauschert, I., Inochanon, K., Bolelli, L., Fuhrmann, S., Brewer, I., Cai, G., MacEachren, A., and Sharma, R., "Multimodal interface platform for geographical information systems (GeoMIP) in crisis management," Proceedings of the 6th international conference on Multimodal interfaces, New York, NY, USA (2004), 339-340.

[3] Alencar, M. F. C., Raposo, A. B., and Barbosa, S. D. J., "Composition of HCI evaluation methods for hybrid virtual environments," Proceedings of the 2011 ACM Symposium on Applied Computing, New York, NY, USA (2011), 1237-1244.

[4] Althoff, F., McGlaun, G., Schuller, B., Morguet, P., and Lang, M., "Using multimodal interaction to navigate in arbitrary virtual VRML worlds," Proceedings of the 2001 workshop on Perceptive user interfaces, New York, NY, USA (2001), 1-8.

[5] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B., "Recent advances in augmented reality," Computer Graphics and Applications, IEEE, vol. 21, no. 6 (2001), 34-47.

[6] Bacim, F., Kopper, R., and Bowman, D. A., "Design and evaluation of 3D selection techniques based on progressive refinement," International Journal of Human-Computer Studies, vol. 71, no. 7 (2013), 785-802.

[7] Balakrishnan, R. and Hinckley, K., "Symmetric bimanual interaction," Proceedings of the SIGCHI conference on Human Factors in Computing Systems, New York, NY, USA (2000), 33-40.

[8] Balakrishnan, R. and Kurtenbach, G., "Exploring bimanual camera control and object manipulation in 3D graphics interfaces," Proceedings of the SIGCHI conference on Human Factors in Computing Systems, New York, NY, USA (1999), 56-62.

[9] Benko, H., "Beyond flat surface computing: challenges of depth-aware and curved interfaces," Proceedings of the 17th ACM international conference on Multimedia, New York, NY, USA (2009), 935-944.

[10] Benko, H., Ishak, E. W., and Feiner, S., "Cross-dimensional gestural interaction techniques for hybrid immersive environments," Virtual Reality, 2005. Proceedings. VR 2005. IEEE (2005), 209-216.

[11] Benko, H., Wilson, A. D., and Baudisch, P., "Precise selection techniques for multi-touch screens," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (2006), 1263-1272.

[12] Bergman, L. D., Richardson, J. S., Richardson, D. C., and Brooks, J. . F. P., "VIEW: an exploratory molecular visualization system with user-definable interaction sequences," Proceedings of the 20th annual conference on Computer graphics and interactive techniques, New York, NY, USA (1993), 117-126.

[13] Bianchi, G., Jung, C., Knoerlein, B., Szekely, G., and Harders, M., "High-fidelity visuo-haptic interaction with virtual objects in multi-modal AR systems," Mixed and Augmented Reality, 2006. ISMAR 2006. IEEE/ACM International Symposium on (2006), 187-196.

[14] Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose, T. D., "Toolglass and magic lenses: the see-through interface," Proceedings of the 20th annual conference on Computer graphics and interactive techniques (1993), 73-80.

[15] Bi, X., Grossman, T., Matejka, J., and Fitzmaurice, G., "Magic desk: bringing multi-touch surfaces into desktop work," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (2011), 2511-2520.

[16] Bolt, R. A., "“Put-that-there”: Voice and gesture at the graphics interface," Proceedings of the 7th annual conference on Computer graphics and interactive techniques, New York, NY, USA (1980), 262-270.

[17] Boucheny, C., Bonneau, G.-P., Droulez, J., Thibault, G., and Ploix, S., "A perceptive evaluation of volume rendering techniques," ACM Trans. Appl. Percept., vol. 5, no. 4 (2009), 1-24.

[18] Bouchet, J. and Nigay, L., "ICARE: a component-based approach for the design and development of multimodal interfaces," CHI '04 Extended Abstracts on Human Factors in Computing Systems, New York, NY, USA (2004), 1325-1328.

[19] Bowman, D. A., Coquillart, S., Froehlich, B., Hirose, M., Kitamura, Y., Kiyokawa, K., and Stuerzlinger, W., "3D User Interfaces: New Directions and Perspectives," Computer Graphics and Applications, IEEE, vol. 28, no. 6 (2008), 20-36.

[20] Bowman, D. A. and Hodges, L. F., "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments," Proceedings of the 1997 symposium on Interactive 3D graphics, New York, NY, USA (1997), 35--ff.

[21] Bowman, D. A., Johnson, D. B., and Hodges, L. F., "Testbed evaluation of virtual environment interaction techniques," Proceedings of the ACM symposium on

Virtual reality software and technology, New York, NY, USA (1999), 26-33.

[22] Bowman, D. A., Kruijff, E., LaViola, J. J., and Poupyrev, I., "3D User Interfaces: Theory and Practice," Addison Wesley (2005).

[23] Bowman, D. A. and Wingrave, C. A., "Design and evaluation of menu systems for immersive virtual environments," Virtual Reality, 2001. Proceedings. IEEE (2001), 149-156.

[24] Brandl, P., Forlines, C., Wigdor, D., Haller, M., and Shen, C., "Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces," Proceedings of the working conference on Advanced visual interfaces (2008), 154-161.

[25] Bruckner, S. and Gröller, E., "Enhancing Depth-Perception with Flexible Volumetric Halos," IEEE Transactions on Visualization and Computer Graphics, vol. 13, no. 6 (2007), 1344-1351.

[26] Bruckner, S. and Groller, E., "VolumeShop: an interactive system for direct volume illustration," Visualization, 2005. VIS 05. IEEE (2005), 671-678.

[27] Bruder, G., Steinicke, F., and St, "Effects of Visual Conflicts on 3D Selection Task Performance in Stereoscopic Display Environments," IEEE Symposium on 3D User Interfaces (2013).

[28] Bryson, S., "Virtual reality in scientific visualization," Commun. ACM, vol. 39, no. 5 (1996), 62-71.

[29] Bryson, S. and Levit, C., "The virtual wind tunnel," Computer Graphics and Applications, IEEE, vol. 12, no. 4 (1992), 25-34.

[30] Butkiewicz, T., Chang, R., Ribarsky, W., and Wartell, Z., , Kathleen S. Hornsby, Ed., CRC Press/Taylor and Francis, ch. Visual Analysis of Urban Terrain Dynamics (2007).

[31] Butkiewicz, T. and Ware, C., "Multi-touch 3D exploratory analysis of ocean flow models," OCEANS 2011 (2011), 1-10.

[32] Buxton, W., "Lexical and pragmatic considerations of input structures," SIGGRAPH Comput. Graph., vol. 17, no. 1 (1983), 31-37.

[33] Buxton, W. and Myers, B., "A study in two-handed input," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (1986), 321-326.

[34] Carvalho, F., Trindade, D. R., Dam, P. F., Raposo, A., and Santos, P. I. N., "Dynamic Adjustment of Stereo Parameters for Virtual Reality Tools," Virtual Reality (SVR), 2011 XIII Symposium on (2011), 66-72.

[35] Coffey, D., Malbraaten, N., Le, T., Borazjani, I., Sotiropoulos, F., and Keefe, D. F., "Slice WIM: a multi-surface, multi-touch interface for overview+detail exploration of volume datasets in virtual reality," Symposium on Interactive 3D Graphics and Games, New York, NY, USA (2011), 191-198.

[36] Cohen, P., McGee, D., Oviatt, S., Wu, L., Clow, J., King, R., Julier, S., and Rosenblum, L., "Multimodal interaction for 2D and 3D environments [virtual reality]," Computer Graphics and Applications, IEEE, vol. 19, no. 4 (1999), 10-13.

[37] Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C., "The CAVE: audio visual experience automatic virtual environment," Commun. ACM, vol. 35, no. 6 (1992), 64-72.

[38] Davis, E. T. and Hodges, L. F., "Human Stereopsis, Fusion, and Stereoscopic Virtual Environments," in Virtual Environments and Advanced Interface Design, Woodrow Barfield and Thomas A. Furness III, Eds., Oxford University Press, ch. Human Stereopsis, Fusion, and Stereoscopic Virtual Environments (1995), 145-172.

[39] de Carvalho, F. G., Raposo, A., and Gattass, M., "Designing a hybrid user interface: a case study on an oil and gas application," Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry, New York, NY, USA (2009), 191-196.

[40] Deering, M., "High resolution virtual reality," Proceedings of the 19th annual conference on Computer graphics and interactive techniques, New York, NY, USA (1992), 195-202.

[41] Dietz, P. and Leigh, D., "DiamondTouch: a multi-user touch technology," Proceedings of the 14th annual ACM symposium on User interface software and technology, New York, NY, USA (2001), 219-226.

[42] Durlach, N. I. and Mavor, A. S., "Virtual reality: scientific and technological challenges," National Academies Press (1994).

[43] Ebert, D. S. and Shaw, C. D., "Minimally immersive flow visualization," Visualization and Computer Graphics, IEEE Transactions on, vol. 7, no. 4 (2001), 343-350.

[44] Ellis, S. R., "Nature and origins of virtual environments: a bibliographical essay ,"

Computing Systems in Engineering , vol. 2, no. 4 (1991), 321-347.

[45] Ellis, S. R., "What are virtual environments?," Computer Graphics and Applications, IEEE, vol. 14, no. 1 (1994), 17-22.

[46] Ellis, S. R. and Menges, B. M., "Judgments of the distance to nearby virtual objects: Interaction of viewing conditions and accommodative demand," Presence: Teleoperators and Virtual Environments, vol. 6, no. 4 (1997), 452-460.

[47] Englmeier, K. H., Krapichler, C., Haubner, M., Seemann, M., and Reiser, M., "Virtual reality and multimedia human-computer interaction in medicine," Multimedia Signal Processing, 1998 IEEE Second Workshop on (1998), 193-202.

[48] Esenther, A. and Ryall, K., "Fluid DTMouse: better mouse support for touch-based interactions," Proceedings of the working conference on Advanced visual interfaces, New York, NY, USA (2006), 112-115.

[49] Feiner, S. and Shamash, A., "Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers," Proceedings of the 4th annual ACM symposium on User interface software and technology, New York, NY, USA (1991), 9-17.

[50] Finkelstein, S., Nickel, A., Lipps, Z., Barnes, T., Wartell, Z., and Suma, E. A., "Astrojumper: Motivating Exercise with an Immersive Virtual Reality Exergame," Presence: Teleoperators and Virtual Environments, vol. 20, no. 1 (2011), 78-92.

[51] Forlines, C., Wigdor, D., Shen, C., and Balakrishnan, R., "Direct-touch vs. mouse input for tabletop displays," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (2007), 647-656.

[52] Frohlich, B., Plate, J., Wind, J., Wesche, G., and Gobel, M., "Cubic-Mouse-based interaction in virtual environments," Computer Graphics and Applications, IEEE, vol. 20, no. 4 (2000), 12-15.

[53] Furnas, G. W. and Bederson, B. B., "Space-scale diagrams: understanding multiscale interfaces," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (1995), 234-241.

[54] Glueck, M. and Khan, A., "Considering multiscale scenes to elucidate problems encumbering three-dimensional intellection and navigation," Artificial Intelligence for Engineering Design, Analysis and Manufacturing, vol. 25 (2011), 393-407.

[55] Grossman, T. and Balakrishnan, R., "An evaluation of depth perception on volumetric displays," AVI '06: Proceedings of the working conference on Advanced

visual interfaces, New York, NY, USA (2006), 193-200.

[56] Grossman, T., Balakrishnan, R., Kurtenbach, G., Fitzmaurice, G., Khan, A., and Buxton, B., "Interaction techniques for 3D modeling on large displays," Proceedings of the 2001 symposium on Interactive 3D graphics, New York, NY, USA (2001), 17-23.

[57] Grossman, T., Wigdor, D., and Balakrishnan, R., "Multi-finger gestural interaction with 3d volumetric displays," Proceedings of the 17th annual ACM symposium on User interface software and technology, New York, NY, USA (2004), 61-70.

[58] Guiard, Y., "Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model," (1987).

[59] Hachet, M., Bossavit, B., Cohé, A., and de la Rivière, J.-B., "Toucheo: multitouch and stereo combined in a seamless workspace," Proceedings of the 24th annual ACM symposium on User interface software and technology, New York, NY, USA (2011), 587-592.

[60] Haluck, R. S. and Krummel, T. M., "Computers and virtual reality for surgical education in the 21st century," Archives of surgery, vol. 135, no. 7 (2000), 786.

[61] Hancock, D. J., "Distributed volume rendering and stereoscopy display for radiotherapy threatment planning," The University of Manchester, Ph.D. dissertation (2001).

[62] Harrison, L., Butkiewicz, T., Wang, X., Ribarsky, W., and Chang, R., "A linked feature space approach to exploring lidar data," Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 7709 (2010).

[63] Held, R. T. and Banks, M. S., "Misperceptions in stereoscopic displays: a vision science perspective," APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization, New York, NY, USA (2008), 23-32.

[64] Hinckley, K., "Input technologies and techniques," The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications (2002), 151-168.

[65] Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F., "Passive real-world interface props for neurosurgical visualization," Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence (1994), 452-458.

[66] Hinckley, K., Pausch, R., Proffitt, D., and Kassell, N. F., "Two-handed virtual manipulation," ACM Transactions on Computer-Human Interaction (TOCHI), vol.

5, no. 3 (1998), 260-302.

[67] Hinckley, K., Sinclair, M., Hanson, E., Szeliski, R., and Conway, M., "The VideoMouse: a camera-based multi-degree-of-freedom input device," Proceedings of the 12th annual ACM symposium on User interface software and technology, New York, NY, USA (1999), 103-112.

[68] Holliman, N., , John Dakin; Robert G W Brown, Ed., New York : Taylor & Francis, ch. Three-dimensional display systems (2006), 1067–1099.

[69] Holliman, N. S., Mapping perceived depth to regions of interest in stereoscopic images, (2004).

[70] Holzapfel, H., Nickel, K., and Stiefelhagen, R., "Implementation and evaluation of a constraint-based multimodal fusion system for speech and 3D pointing gestures," Proceedings of the 6th international conference on Multimodal interfaces, New York, NY, USA (2004), 175-182.

[71] Houtgast, E., Pfeiffer, O., and Post, F., "Navigation and Interaction in a Multi-Scale Stereoscopic Environment," Proceedings of the 2005 IEEE Conference 2005 on Virtual Reality, Washington, DC, USA (2005), 275-276.

[72] Hubbold, R. J. and Hancock, D. J., "Stereo display of nested 3D volume data using automatic tunnelling," in Stereoscopic Displays and Virtual Reality Systems VI, Proceedings of SPIE 3639 (1999).

[73] Kaufman, A. E., "Volume visualization," ACM Comput. Surv., vol. 28, no. 1 (1996), 165-167.

[74] Kaufmann, H., Schmalstieg, D., and Wagner, M., "Construct3D: A Virtual Reality Application for Mathematics and Geometry Education," Education and Information Technologies, vol. 5, no. 4 (2000), 263-276.

[75] Kersten, M., Stewart, J., Troje, N., and Ellis, R., "Enhancing Depth Perception in Translucent Volumes," IEEE Transactions on Visualization and Computer Graphics, vol. 12, no. 5 (2006), 1117-1124.

[76] Kin, K., Agrawala, M., and DeRose, T., "Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation," Proceedings of Graphics Interface 2009, Toronto, Ont., Canada, Canada (2009), 119-124.

[77] Kniss, J., Kindlmann, G., and Hansen, C., "Multidimensional Transfer Functions for Interactive Volume Rendering," IEEE Transactions on Visualization and Computer Graphics, vol. 8, no. 3 (2002), 270-285.

[78] Kopper, R., Ni, T., Bowman, D. A., and Pinho, M., "Design and evaluation of navigation techniques for multiscale virtual environments," Virtual Reality Conference, 2006 (2006), 175-182.

[79] Krueger, M. W., Gionfriddo, T., and Hinrichsen, K., "VIDEOPLACE-an artificial reality," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (1985), 35-40.

[80] Kruger, J. and Westermann, R., "Acceleration Techniques for GPU-based Volume Rendering," VIS '03: Proceedings of the 14th IEEE Visualization 2003 (VIS'03), Washington, DC, USA (2003), 38.

[81] Kruijff, E., "Unconventional 3D User Interfaces for Virtual Environments," Graz University of Technology, Institute for Computer Graphics and Vision, Ph.D. dissertation (2006).

[82] Lambooij, M. T. M., IJsselsteijn, W. A., and Heynderickx, I., Visual discomfort in stereoscopic displays: a review, (2007).

[83] Leganchuk, A., Zhai, S., and Buxton, W., "Manual and cognitive benefits of two-handed input: an experimental study," ACM Transactions on Computer-Human Interaction (TOCHI), vol. 5, no. 4 (1998), 326-359.

[84] Leigh, J., Johnson, A. E., Vasilakis, C. A., and DeFanti, T. A., "Multi-perspective Collaborative Design in Persistent Networked Virtual Environments," Proceedings of the IEEE Virtual Reality Annual International Symposium (1996), 253-260,271-272.

[85] Ling, F. and Yang, L., "Improved on Maximum Intensity Projection," AICI '09: Proceedings of the 2009 International Conference on Artificial Intelligence and Computational Intelligence, Washington, DC, USA (2009), 491-495.

[86] Lipton, L., "Foundations of the Stereoscopic Cinema," Van Nostrand Reinhold (1982).

[87] Liu, J., Pastoor, S., Seifert, K., and Hurtienne, J., "Three-dimensional PC: toward novel forms of human-computer interaction," In Three-Dimensional Video and Display Devices and Systems SPIE (2000), 5-8.

[88] Logitech, "2D/6D Mouse technical reference manual," (1991).

[89] Maciejewski, R., Choi, S., Ebert, D. S., and Tan, H. Z., "Multi-Modal Perceptualization of Volumetric Data and Its Application to Molecular Docking," WHC '05: Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems,

Washington, DC, USA (2005), 511-514.

[90] MacKenzie, I. S., Soukoreff, R. W., and Pal, C., "A two-ball mouse affords three degrees of freedom," CHI '97 Extended Abstracts on Human Factors in Computing Systems, New York, NY, USA (1997), 303-304.

[91] Mackinlay, J. D., Card, S. K., and Robertson, G. G., "A Semantic Analysis of the Design Space of Input Devices," Human-Computer Interaction, vol. 5 (1990), 145-190.

[92] Mapes, D. P. and Moshell, J. M., "A two-handed interface for object manipulation in virtual environments," Presence: Teleoperators and Virtual Environments, vol. 4, no. 4 (1995), 403-416.

[93] Marquardt, N., Kiemer, J., and Greenberg, S., "What caused that touch?: expressive interaction with a surface through fiduciary-tagged gloves," ACM International Conference on Interactive Tabletops and Surfaces, New York, NY, USA (2010), 139-142.

[94] Massink, M., Duke, D., Smith, S., and Maria, V. S., "Towards Hybrid Interface Specification for Virtual Environments," (1999).

[95] McCrae, J., Mordatch, I., Glueck, M., and Khan, A., "Multiscale 3D navigation," Proceedings of the 2009 symposium on Interactive 3D graphics and games, New York, NY, USA (2009), 7-14.

[96] Meyer, M. and Barr, A., "ALCOVE: design and implementation of an object-centric virtual environment," (1999), 46-52.

[97] Mine, M. R., Brooks, J. . F. P., and Sequin, C. H., "Moving objects in space: exploiting proprioception in virtual-environment interaction," Proceedings of the 24th annual conference on Computer graphics and interactive techniques, New York, NY, USA (1997), 19-26.

[98] Moehring, M. and Froehlich, B., "Effective manipulation of virtual objects within arm's reach," Virtual Reality Conference (VR), 2011 IEEE (2011), 131-138.

[99] Mokka, S., Väätänen, A., Heinilä, J., and Välkkynen, P., "Fitness computer game with a bodily user interface," Proceedings of the second international conference on Entertainment computing, Pittsburgh, PA, USA (2003), 1-3.

[100] Mora, B. and Ebert, D. S., "Instant Volumetric Understanding with Order-Independent Volume Rendering," Computer Graphics Forum, vol. 23, no. 3 (2004), 489-497.

[101] Motion, L. https://www.leapmotion.com/.

[102] Nielson, G. M., Hagen, H., and Muller, H., "Scientific Visualization, Overviews, Methodologies, and Techniques," , Washington, DC, USA (1997).

[103] Oh, J.-Y. and Hua, H., "Usability of multi-scale interfaces for 3d workbench displays," Presence: Teleoper. Virtual Environ., vol. 17, no. 5 (2008), 415-440.

[104] OpenAL. http://connect.creativelabs.com/openal/.

[105] OpenSceneGraph. http://trac.openscenegraph.org/projects/osg/.

[106] osgEarth. http://osgearth.org/.

[107] Oviatt, S., "Multimodal interfaces," The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications (2003), 286-304.

[108] Oviatt, S., "Ten myths of multimodal interaction," Commun. ACM, vol. 42, no. 11 (1999), 74-81.

[109] Oviatt, S. and Cohen, P., "Perceptual user interfaces: multimodal interfaces that process what comes naturally," Communications of the ACM, vol. 43, no. 3 (2000), 45-53.

[110] Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., and Saarikko, P., "It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (2008), 1285-1294.

[111] Pfister, H., Lorensen, B., Bajaj, C., Kindlmann, G., Schroeder, W., Avila, L. S., Martin, K., Machiraju, R., and Lee, J., "The Transfer Function Bake-Off," IEEE Comput. Graph. Appl., vol. 21, no. 3 (2001), 16-22.

[112] Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., and Mine, M. R., "Image plane interaction techniques in 3D immersive environments," Proceedings of the 1997 symposium on Interactive 3D graphics, New York, NY, USA (1997), 39--ff.

[113] Pierce, J. S. and Pausch, R., "Navigation with place representations and visible landmarks," Virtual Reality, 2004. Proceedings. IEEE (2004), 173-288.

[114] Pirchheim, C., Schmalstieg, D., and Bornik, A., "Visual Programming for Hybrid User Interfaces," Proc 2nd International Workshop on Mixed Reality User Interfaces (2007).

[115] Polhemus. http://www.polhemus.com/.

[116] Poupyrev, I., Billinghurst, M., Weghorst, S., and Ichikawa, T., "The go-go interaction technique: non-linear mapping for direct manipulation in VR," Proceedings of the 9th annual ACM symposium on User interface software and technology, New York, NY, USA (1996), 79-80.

[117] Poupyrev, I., Ichikawa, T., Weghorst, S., and Billinghurst, M., "Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques," Computer Graphics Forum, vol. 17, no. 3 (1998), 41-52.

[118] PQ Lab. http://multitouch.com/.

[119] Rauschert, I., Agrawal, P., Sharma, R., Fuhrmann, S., Brewer, I., and MacEachren, A., "Designing a human-centered, multimodal GIS interface to support emergency management," Proceedings of the 10th ACM international symposium on Advances in geographic information systems, New York, NY, USA (2002), 119-124.

[120] Rick, T., von Kapri, A., and Kuhlen, T., "GPU implementation of 3D object selection by conic volume techniques in virtual environments," Virtual Reality Conference (VR), 2010 IEEE (2010), 243-246.

[121] Robertson, G., Czerwinski, M., and van Dantzich, M., "Immersion in desktop virtual reality," Proceedings of the 10th annual ACM symposium on User interface software and technology, New York, NY, USA (1997), 11-19.

[122] Robinett, W. and Holloway, R., "Implementation of flying, scaling and grabbing in virtual worlds," Proceedings of the 1992 symposium on Interactive 3D graphics, New York, NY, USA (1992), 189-192.

[123] Robinett, W. and Holloway, R., "The Visual Display Transformation for Virtual Reality," Chapel Hill, NC, USA, Tech. rep. (1994).

[124] Roettger, S., Guthe, S., Weiskopf, D., Ertl, T., and Strasser, W., "Smart hardware-accelerated volume rendering," VISSYM '03: Proceedings of the symposium on Data visualisation 2003, Aire-la-Ville, Switzerland, Switzerland (2003), 231-238.

[125] Rule, J. T., "The Shape of Stereoscopic Images," Journal of the Optical Society of America, vol. 28 (1940), 313-322.

[126] Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalav\', Encarna\c, Gervautz, M., and Purgathofer, W., "The studierstube augmented reality project," Presence: Teleoper. Virtual Environ., vol. 11, no. 1 (2002), 33-54.

[127] Schultheis, U., Jerald, J., Toledo, F., Yoganandan, A., and Mlyniec, P., "Comparison of a two-handed interface to a wand interface and a mouse interface for fundamental 3D tasks," (March), 117-124.

[128] Segen, J. and Kumar, S., "Fast and accurate 3D gesture recognition interface," Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on, vol. 1 (1998), 86-91 vol.1.

[129] Shaw, C. and Green, M., "Two-handed polygonal surface design," Proceedings of the 7th annual ACM symposium on User interface software and technology, New York, NY, USA (1994), 205-212.

[130] Shibata, T., Kim, J., Hoffman, D. M., and Banks, M. S., "The zone of comfort: Predicting visual discomfort with stereo displays," Journal of Vision, vol. 11, no. 8 (2011).

[131] Shoemake, K., "ARCBALL: a user interface for specifying three-dimensional orientation using a mouse," Graphics Interface, vol. 92 (1992), 151-156.

[132] Song, P., Goh, W. B., Fu, C. W., Meng, Q., and Heng, P. A., "WYSIWYF: exploring and annotating volume data with a tangible handheld device," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (2011), 1333-1342.

[133] Song, P., Goh, W. B., Hutama, W., Fu, C. W., and Liu, X., "A handle bar metaphor for virtual object manipulation with mid-air interaction," Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, New York, NY, USA (2012), 1297-1306.

[134] Spottiswoode, R. and Spottiswoode, N., "The Theory of StTereoscopic Transmission," University of California Press, Berkeley and Los Angeles (1953).

[135] St. Julien, T. U. and Shaw, C. D., "Firefighter command training virtual environment," Proceedings of the 2003 conference on Diversity in computing, New York, NY, USA (2003), 30-33.

[136] Steed, A., "Towards a General Model for Selection in Virtual Environments," 3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on (2006), 103-110.

[137] Stewart, A. J., "Vicinity shading for enhanced perception of volumetric data," Visualization, 2003. VIS 2003. IEEE (2003), 355-362.

[138] Suma, E., osgVirtualEnvironment.

[139] Sutherland, I. E., "A head-mounted three dimensional display," Proceedings of the December 9-11, 1968, fall joint computer conference, part I, New York, NY, USA (1968), 757-764.

[140] Svakhine, N. A., Ebert, D. S., and Andrews, W. M., "Illustration-Inspired Depth Enhanced Volumetric Medical Visualization," IEEE Transactions on Visualization and Computer Graphics, vol. 15, no. 1 (2009), 77-86.

[141] Taylor, I. R. M., Hudson, T. C., Seeger, A., Weber, H., Juliano, J., and Helser, A. T., "VRPN: a device-independent, network-transparent VR peripheral system," Proceedings of the ACM symposium on Virtual reality software and technology, New York, NY, USA (2001), 55-61.

[142] Teather, R. J. and Stuerzlinger, W., "Pointing at 3D targets in a stereo head-tracked virtual environment," (2011), 87-94.

[143] Trindade, D. R. and Raposo, A. B., "Improving 3D navigation in multiscale environments using cubemap-based techniques," Proceedings of the 2011 ACM Symposium on Applied Computing, New York, NY, USA (2011), 1215-1221.

[144] Ulinski, A. C., "Taxonomy and experimental evaluation of two-handed selection techniques for volumetric data," Charlotte, NC, USA, Ph.D. dissertation ISBN: 978-0-549-91826-4, (2008).

[145] Ulinski, A. C., Wartell, Z., Goolkasian, P., Suma, E. A., and Hodges, L. F., "Selection performance based on classes of bimanual actions," 3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on (2009), 51-58.

[146] Ulinski, A., Zanbaka, C., Wartell, Z., Goolkasian, P., and Hodges, L. F., "Two handed selection techniques for volumetric data," 3D User Interfaces, 2007. 3DUI'07. IEEE Symposium on (2007).

[147] Valkov, D., Steinicke, F., Bruder, G., and Hinrichs, K., "2d touching of 3d stereoscopic objects," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (2011), 1353-1362.

[148] Vicon. http://www.vicon.com/.

[149] Wagner, M., "The metric of visual space," Perception \& Psychophysics, vol. 38 (1985), 483-495.

[150] Wang, R., Paris, S., and Popović, J., "6D hands: markerless hand-tracking for computer aided design," Proceedings of the 24th annual ACM symposium on User interface software and technology, New York, NY, USA (2011), 549-558.

[151] Ware, C., "Dynamic stereo displays," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (1995), 310-316.

[152] Ware, C., "Information Visualization: Perception for Design; 2nd Edition," Morgan Kaufmann (2004).

[153] Ware, C., "Using hand position for virtual object placement," Vis. Comput., vol. 6, no. 5 (1990), 245-253.

[154] Ware, C., Arthur, K., and Booth, K. S., "Fish tank virtual reality," Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, New York, NY, USA (1993), 37-42.

[155] Ware, C. and Franck, G., "Evaluating stereo and motion cues for visualizing information nets in three dimensions," ACM Trans. Graph., vol. 15, no. 2 (1996), 121-140.

[156] Ware, C., Gobrecht, C., and Paton, M. A., "Dynamic adjustment of stereo display parameters," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, vol. 28, no. 1 (1998), 56-65.

[157] Ware, C. and Jessome, D. R., "Using the bat: a six-dimensional mouse for object placement," Computer Graphics and Applications, IEEE, vol. 8, no. 6 (1988), 65-70.

[158] Ware, C. and Jessome, D. R., "Using the Bat: A Six-Dimensional Mouse for Object Placement," IEEE Comput. Graph. Appl., vol. 8, no. 6 (1988), 65-70.

[159] Ware, C. and Lowther, K., "Selection using a one-eyed cursor in a fish tank VR environment," ACM Trans. Comput.-Hum. Interact., vol. 4, no. 4 (1997), 309-322.

[160] Ware, C. and Mitchell, P., "Reevaluating stereo and motion cues for visualizing graphs in three dimensions," APGV '05: Proceedings of the 2nd symposium on Applied perception in graphics and visualization, New York, NY, USA (2005), 51-58.

[161] Ware, C. and Mitchell, P., "Visualizing graphs in three dimensions," ACM Trans. Appl. Percept., vol. 5, no. 1 (2008), 1-15.

[162] Wartell, Z., "Stereoscopic Head-Tracked Displays: Analysis and Development of Display Algorithms," Georgia Institute of Technology, (2001).

[163] Wartell, Z., Hodges, L. F., and Ribarsky, W., "An Analytic Comparison of α-False Eye Separation, Image Scaling and Image Shifting in Stereoscopic Displays,"

(2000).

[164] Wartell, Z., Houtgast, E., Pfeiffer, O., Shaw, C., Ribarsky, W., and Post, F., "Interaction Volume Management in a Multi-scale Virtual Environment," in Advances in Information and Intelligent Systems, ZbigniewW. Ras and William Ribarsky, Eds., Springer Berlin Heidelberg, vol. 251 (2009), 327-349.

[165] Wartell, Z., Ribarsky, W., and Hodges, L., "Third-Person Navigation of Whole-Planet Terrain in a Head-tracked Stereoscopic Environment," VR '99: Proceedings of the IEEE Virtual Reality, Washington, DC, USA (1999), 141.

[166] Welch, G. and Foxlin, E., "Motion tracking: no silver bullet, but a respectable arsenal," Computer Graphics and Applications, IEEE, vol. 22, no. 6 (2002), 24-38.

[167] Westerman, W., "Hand tracking, finger identification, and chordic manipulation on a multi-touch surface," University of Delaware, Ph.D. dissertation (1999).

[168] Wickens, C. D., "Virtual reality and education," Systems, Man and Cybernetics, 1992., IEEE International Conference on (1992), 842-847 vol.1.

[169] Wigdor, D., Williams, S., Cronin, M., Levy, R., White, K., Mazeev, M., and Benko, H., "Ripples: utilizing per-contact visualizations to improve user interaction with touch displays," Proceedings of the 22nd annual ACM symposium on User interface software and technology, New York, NY, USA (2009), 3-12.

[170] Wilson, A. D., "Depth-Sensing Video Cameras for 3D Tangible Tabletop Interaction," Horizontal Interactive Human-Computer Systems, 2007. TABLETOP '07. Second Annual IEEE International Workshop on (2007), 201-204.

[171] Woods, A. J., Docherty, T., and Koch, R., "Image distortions in stereoscopic video systems," , vol. 1915 (1993), 36-48.

[172] Wu, A., Zhang, W., Hu, B., and Zhang, X., "Evaluation of Wayfinding Aids Interface in Virtual Environment," in Human-Computer Interaction. Interaction Platforms and Techniques, JulieA. Jacko, Ed., Springer Berlin Heidelberg, vol. 4551 (2007), 700-709.

[173] Yeh, Y. Y. and Silverstein, L. D., "Limits of fusion and depth judgment in stereoscopic color displays," Hum. Factors, vol. 32 (1990), 45-60.

[174] Youngblut, C., "Educational Uses of Virtual Reality Technology.," DTIC Document, Ph.D. dissertation (1998).

[175] Zeleznik, R. C., Forsberg, A. S., and Strauss, P. S., "Two pointer input for 3D interaction," Proceedings of the 1997 symposium on Interactive 3D graphics

(1997), 115-120.

[176] Zhai, S., "Investigation of feel for 6DOF inputs: isometric and elastic rate control for manipulation in 3D environments," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 37 (1993), 323-327.

[177] Zhai, S., Milgram, P., and Buxton, W., "The influence of muscle groups on performance of multiple degree-of-freedom input," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (1996), 308-315.

[178] Zimmerman, T. G., Lanier, J., Blanchard, C., Bryson, S., and Harvill, Y., "A hand gesture interface device," Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface, New York, NY, USA (1987), 189-192.

[179] Zyda, M., "From visual simulation to virtual reality to games," Computer, vol. 38, no. 9 (2005), 25-32.

APPENDIX A: MATERIALS FROM DEPTH PERCEPTION FOR A VOLUMETRIC DATASET EXPERIMENTS

This appendix contains materials used in depth perception studies for a 3D volume dataset, which was reported in Chapter 3. The following materials are included, listed in order of appearance:

1. The informed consent form
2. Demographics and post questionnaires form
3. Confidence rate form
4. Post questionnaires form

UNC CHARLOTTE
The University of North Carolina at Charlotte

Informed Consent for
Perception of volumetric dataset under virtual environment conditions

Project Purpose

In this study we will test the depth perception of users under virtual environment conditions (stereopsis and head-tracking) with scalar volumetric dataset.

Investigators

Isaac Cho, Computer Science
Zachary Wartell, Computer Science

Eligibility

You may participate in this study if you are above 18 and if you have 20/20 vision or corrected vision to 20/20 and can comfortably communicate in spoken English.

Overall Description of Participation

In the first step, we will demonstrate stereoscopic display and head-tracking to make you familiar with them. Stereopsis is an important binocular depth cue generated by the differences between the two views of a scene seen from a person's two eyes. Head-tracking is a tracking system which tracks the position and the orientation of a participant's head to generate an optimal perspective image. We will train you on how these technologies work by having you view a volumetric medical dataset generated from a CT or MRI scan. Then you will take a simple depth perception test. You will be asked which virtual object appears in front in the display. Also, we will survey you about your experiences with 3D applications (e.g. 3D games and 3D movies) and your familiarity with using a computer.

In the next step, you will be asked about depth ordering of cylinders under specific conditions (combination of stereopsis and head-tracking). There are two experiments. The instructor will tell you in which experiment you will participate. If you participate in Experiment 1, then you will be asked which virtual cylinder is in front (or middle or behind) from a set of 6 virtual cylinders. You can answer by pressing the number keys on the keyboard. This experiment will be repeated 4 times with different conditions. If you participate in Experiment 2, then you will be asked which of two target cylinders is in front of the other. The two target cylinders will be picked out of a set of 6 cylinders. The target pair will be designated by highlighting the pair's intersection in a 2D image of the

6 cylinder set. First you will see this 2D image for 1 second with the highlighted intersection, and then you will see the 3D cylinders for 2 seconds. After the 3D dataset disappears, you can answer which is in front by selecting from a menu using a mouse. You will have 5 seconds to answer this question. This experiment will be repeated 6 times. The whole experiment time is approximately 40 minutes.

Then you will complete a post-experiment questionnaire which focuses on which display conditions were most effective for you for perceiving the depth of the cylinders.

Length of Participation

Participation should take approximately 50-60 minutes.

Risks and Benefits of Participation

While using virtual environment systems, some people experience slight symptoms of disorientation, nausea, or dizziness. These can be similar to "motion sickness" or the feeling experienced in wide-screen movies and theme park rides. We do not expect these to be strong or to last after participants leave the laboratory. You will be reminded often that if you feel uncomfortable and wish to stop the experiment, you are free to do so at any time without penalty by simply announcing you desire to stop.

The primary scientific benefit of the study is determine effectiveness of virtual environment conditions for medical or scientific volumetric images.

There is a benefit for participants who volunteer via the Psychology Department’s on-line subject pool. Participants will receive 1 credit in their general psychology class. Students who do not wish to participate, or who are excluded in the study due to stereoblindness, may leave and they are granted participation credit for 15 minutes or one-half a credit.

Students in ITCS 6125/8125 Virtual Environments course must complete an experment as part of the course reqruiments. Therefore this experiment will be used and students in ITCS 6125/8125 must complete the experiments as part of a course homework assignment. However participation in the research study is voluntary and students will be given the option of having their data not collected for experimental analysis. The course, ITCS 8125/6125, is about virtual environments and virtual reality. The study is a perceptual study regarding stereoscopic and head-tracked virtual reality displays. Students participating in the experiment will benefit by hands-on learning about a current research project in perception in virtual reality systems. The experiment will engage students with an additional virtual reality hardware system beyond those used in their class projects. The informed consent form below has a special section for ITCS 8125/6125 students informing them of this option.

Volunteer Statement

You are a volunteer. The decision to participate in this study is completely up to you. If you decide to participate in the study, you may stop at any time. You will not be treated any differently if you decide not to participate in the study or if you stop once you have started. You can request to withdraw your segment after the testing is complete.

If you are enrolled in course ITCS6125 or ITCS8125, your participation in this study counts as a required homework assignment. However, having your data recorded and used for the actual research study is voluntary.

Confidentiality Statement

Any information about your participation, including your identity, is completely confidential. The following efforts will be taken to protect confidentiality and privacy:

1) The informed consent form will be kept in a locked filing cabinet, separate from the rest of the data.
2) All participants will be assigned a random ID consisting two randomly-generated initials (initials will not correspont to participants' name). The participants will only be referred by assigned alphanumeric codes both in internal communication between researchers or in the form of written reports.
3) The investigator and co-investigators will ask the participants not to mention their name or identify themselves during the recordings. The recording is only for internal use such as transcription and will not be made available to the public. Screenshots from the video recording might be published without disclosing the identify of any participants.
4) All digitally recorded files during the study will be kept in the Charlotte Visualization Center (room 437 in Woodward Hall) on password-protected NOVELL-powered computers. The files will be destroyed after two years by investigators under the guidance of the responsible faculty.

Statement of Fair Treatment and Respect

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the university’s Research Compliance Office (704-687-3309) if you have questions about how you are treated as a study participant. If you have any questions about the actual project or study, please contact Isaac Cho (icho1@uncc.edu) or Dr. Zachary Wartell (zwartell@uncc.edu) at 704-687-8559.

Approval Date

This form was approved for use on *Oct 8, 2009* for use for one year.

Participant Consent

I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principal investigator of this research study.

If you are a student in ITCS 6125/8125, please also complete the following box:

I am a student in ITCS 6125/8125, I understand that while participating in the experiment is required as a homework assignment, the collection of my data for use in the research study is voluntary and my choice on whether to have the data collected will not affect my homework grade.

Please circle one option:

As a ITCS 6125/8125 student, I [do] | [do not] consent to the collection of my data for use in the research study.

Participant Name (PRINT) DATE

Participant Signature

Investigator Signature DATE

1. Your given ID number:

2. Your age:

3. Your gender:

4. Occupational Status: Undergraduate student ___

Master Student ___

PhD Student ___

Research Assistant/Fellow ___

Staff-systems, technical ___

Faculty ___

Administrative Staff ___

Other : __________________

5. Your major:

6. Are you colorblind? : Yes / No

7. Do you have 20/20 eyesight (or corrected 20/20)? Yes / No

8. Are you familiar with using a mouse and keyboard? Yes / No

9. Have you ever felt cyber sickness such as dizziness or nausea before? Yes / No

1. How often do you use a computer in your daily activities?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

2. How often do you play 2D computer games?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

3. How often do you play 3D computer games?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

4. How often do you play computer games on a computer/PC?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

5. How often do you play computer games using a game console, such as Nintendo® , XBox® , Sony Playstation® , other?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

6. How often do play computer games using a game console with a motion capture device, such as XBox Kinect® , Sony Playstation Move® , other?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

7. How often do you play watch 3D movies for which you have to use shutter glasses in the theater?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

8. How often do you play computer games or watch movies using a 3D TV or display with shutter glasses?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

Experiment 1

1. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

2. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Non Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

3. To what extent do you feel that you were able to accurately answer the question about the Non Stereopsis and Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

4. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Non Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

Experiment 2

1. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

2. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Non Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

3. To what extent do you feel that you were able to accurately answer the question about the Non Stereopsis and Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

4. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Non Head Tracking condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

5. To what extent do you feel that you were able to accurately answer the question about the Stereopsis and Simulation condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

6. To what extent do you feel that you were able to accurately answer the question in Stereopsis and Simulation condition?

(Not At All) (A Great Deal)
1 2 3 4 5 6 7

Overall, which combination of conditions was the most effective to make you perceive the depth ordering?

____

____

____

1. Overall, which singular condition was the most effective for your judgment of the depth ordering?

____

____

____

2. Did size of cylinders help or obstruct your depth judgment?

____

____

____

3. Did the texture of cylinders help or obstruct your depth judgment?

____

____

____

4. Did the noise background help or obstruct your depth judgment?

____

____

____

5. Please write any comments if you have something to tell the instructor about this experiment.

____

____

____

APPENDIX B: MATERIALS FROM SPINDLE + WHEEL EXPERIMENTS

This appendix contains materials used in Spindle+Wheel Experiment 1 and 2, which was reported in Chapter 4. The following materials are included, listed in order of appearance:

1. Informed consent form for Experiment 1
2. Demographics and post questionnaires form
3. Post questionnaires form for Experiment 1
4. Informed consent form for Experiment 2
5. Confidence rate form for Experiment 2
6. Post questionnaires form for Experiment 2

UNCCHARLOTTE
The University of North Carolina at Charlotte

Informed Consent for
Evaluating effectiveness of a two-handed 3D object manipulation technique in desktop VR for 3D applications

Project Purpose

In this study we will determine effectiveness of our two-handed 3D object manipulation technique by comparing with other two-handed object manipulation technique in desktop VR, which provides stereoscopic and head-tracking displays, for 3D applications.

Investigators

Isaac Cho, Computer Science
Zachary Wartell, Computer Science

Eligibility

You may participate in this study if you are 18 years of age or higher and if you have 20/20 vision, corrected vision to 20/20 or higher (i.e. you can clearly read text on a computer workstation monitor), can comfortably use your arms, hands and fingers for everyday tasks and you can communicate in spoken English.

Overall Description of Participation

In the first step, we will demonstrate stereoscopic display and head-tracking to make you familiar with them. Stereopsis is an important binocular depth cue generated by the differences between the two views of a scene seen from a person's two eyes. Head-tracking is a tracking system which tracks the position and an orientation of participant's head to generate an optimal perspective image. We will train you on how these technologies work by showing you how to view the 3D application and how to perform a navigation task (rotating, translating and scaling) with one-handed or two-handed 6DOF buttonballs. Also, we will survey your experiences with 3D applications (e.g. 3D games and 3D movies) and your familiarity with using a computer and user interface.

In the next step, you will perform a target-finding task with two-handed buttonballs depending on the object manipulation conditions. There are 2 object manipulation conditions in the experiment: 1. two-handed "Spindle Only" object manipulation mode and 2. Two-handed "Spindle+Wheel" object manipulation mode. The task is to find a colored box on gridded ground and place the box between two boxes in the center of the screen. The box has to be bigger than the red wired box and smaller than the colored outer box. You, also, need to match the colors between the target box and the outer box.

After you finish the experiment, you will take a post-experiment questionnaire.

Length of Participation

Participation should take approximately 50-60 minutes.

Risks and Benefits of Participation

While using virtual environment systems, some people experience slight symptoms of disorientation, nausea, or dizziness. These can be similar to "motion sickness" or the feeling experienced in wide-screen movies and theme park rides. We do not expect these to be strong or to last after participants leave the laboratory. if you feel uncomfortable and wish to stop the experiment, you are free to do so at any time without penalty by simply announcing your desire to stop.

The primary scientific benefit of the study is to determine effectiveness of the new two-handed object manipulation technique for 3D applications in desktop VR.

There is a benefit for participants who volunteer via the psychology department’s on-line subject pool. They will receive 1 credit in their general psychology class. Students who do not wish to participate, or who are excluded in the study due to stereoblindness, or other exclusion criteria, may leave and will be granted participation credit for 15 minutes or one-half a credit.

Volunteer Statement

You are a volunteer. The decision to participate in this study is completely up to you. If you decide to participate in the study, you may stop at any time. You will not be treated any differently if you decide not to participate in the study or if you stop once you have started. You can request to withdraw your segment after the testing is complete.

Confidentiality Statement

Any information about your participation, including your identity, is completely confidential. The following efforts will be taken to protect confidentiality and privacy:

1) The informed consent form will be kept in a locked filing cabinet, separate from the rest of the data.

2) All participants will be assigned a random ID consisting two randomly-generated initials (initials will not correspont to participants' name). The participants will only be referred by assigned alphanumeric codes both in internal communication between researchers or in the form of written reports.

3) The investigator and co-investigators will ask the participants not to mention their name or identify themselves during the recordings. The recording is only for internal use such as transcription and will not be made available to the public.

Screenshots from the video recording might be published without disclosing the identify of any participants.

4) All digitally recorded files during the study will be kept in the Charlotte Visualization Center (room 437 in Woodward Hall) on password-protected computers. The files will be destroyed after two years by investigators under the guidance of the responsible faculty.

Statement of Fair Treatment and Respect

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the university's Research Compliance Office (704-687-3309) if you have questions about how you are treated as a study participant. If you have any questions about the actual project or study, please contact Isaac Cho (icho1@uncc.edu) or Dr. Zachary Wartell (zwartell@uncc.edu) at 704-687-8442.

Approval Date

This form was approved for use on *April 1st, 2013* for use for one year.

Participant Consent

I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principal investigator of this research study.

____ ____________________________
Participant Name (PRINT) DATE

____
Participant Signature

____ ____________________________
Investigator Signature DATE

1. Your given ID number (Instructor only):

2. Your age:

3. Your gender:

4. Occupational Status: Undergraduate student ___

Master Student ___

PhD Student ___

Research Assistant/Fellow ___

Staff-systems, technical ___

Faculty ___

Administrative Staff ___

Other: __________________

5. Your major:

6. Are you colorblind? : Yes / No

7. Do you have any problems viewing the computer screen without it blurring if you sit 30 inches from the screen? Yes / No

8. Do you have any disabilities or injuries that might limit your ability to use either your left or right arm, hand and/or fingers in everyday tasks such as writing, painting, using a computer mouse or advanced game controller? Yes / No

9. Are you familiar with using a mouse and keyboard? Yes / No

10. Have you ever felt motion sick (dizziness or nausea) while playing a computer game or viewing a large, screen movie before? Yes / No

1. How often do you use a computer in your daily activities?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

2. How often do you play 2D computer games?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

3. How often do you play 3D computer games?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

4. How often do you play computer games (of any kind) on a computer/PC?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

5. How often do you play computer games using a game console, such as Nintendo® , XBox® , Sony Playstation® , other?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

6. How often do play computer games using a game console with a motion capture device, such as XBox Kinect® , Sony Playstation Move® , other?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

Definition: Stereoscopic 3D

Stereoscopic 3D refers to a display that creates a true 3D image that appears to pop-out in front of and behind the screen. These displays are found in some movie theaters, television sets and computer monitors. Most stereoscopic 3D display technologies known to consumers require they wear special glasses.

7. How often do you watch stereoscopic 3D movies in the theater?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

8. How often do you play computer games or watch movies on an in-home television using stereoscopic 3D?

(Never)						(A Great Deal)
1	2	3	4	5	6	7

Definition: 3D User Interface

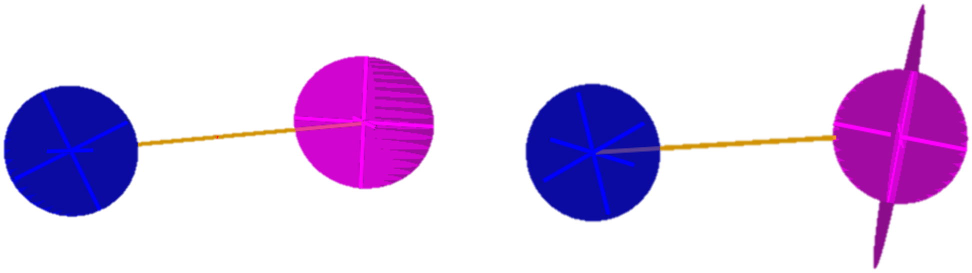
A "3D user interface" is a human-computer interface where the user views 3D computer graphics and interacts with those graphics by traveling through the 3D environment, and/or manipulating and changing the 3d environment. 3D user interfaces may or may not use stereoscopic 3D displays. Also 3D user interfaces may or may not use advanced 3D input devices such as the Microsoft Kinect, PlayStation Move, Nintendo Wii, etc.

9. If you have used any 3D user interfaces before, then describe what 3D user interfaces you have used and mention what type of display and input device technology you used with them.

____

____

____



A. Manipulation Technique 1 and B. Manipulation Technique 2

1. Overall, which object manipulation technique (A or B) was better than the others for the navigation task? Why?

2. Which object manipulation technique was better for rotating? Why?

3. Did the pitch controlling help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

3. How much did the wheel on the 3D cursor to control pitch rotation?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

4. Which one is more helpful to realize that you accomplished the task, sound or color (or both)?

Sound:
(Not At All) (Very Helpful)
1 2 3 4 5 6 7

Box Outline Color:
(Not At All) (Very Helpful)
1 2 3 4 5 6 7

5. Which object manipulation technique was more intuitive to perform the task? Why?

6. Did the line between 3D cursors help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

7. Did the sphere on the line connection between 3D cursors help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

8. Did the box size make the task difficult?

(Not At All) (Very Difficult)
1 2 3 4 5 6 7

9. How much arm fatigue did you feel with the object manipulation 1 (A)?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

10. How much arm fatigue did you feel with the object manipulation 2 (B)?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

Some people experience difficulty in perceiving a clear 3D stereoscopic image on the display. Often if the 3D image extends too far in front or behind the display surface, a person may perceive only two separate double images rather than a single 3D image. In some circumstances, a person may experience eye strain, visual fatigue or headaches when viewing this type of display system. Collectively, these negatives experiences are called “stereoscopic fusion problems”.

10. How frequently did you feel stereoscopic fusion problems with the object manipulation 1 (A)?

(Not At All) (Very Frequently)
1 2 3 4 5 6 7

11. How frequently did you feel stereoscopic fusion problems with the object manipulation 2 (B)?

(Not At All) (Very Frequently)
1 2 3 4 5 6 7

12. If you have any comments for this study, please give us feedback.

UNCCHARLOTTE
The University of North Carolina at Charlotte

Informed Consent for
Evaluating effectiveness of a two-handed 3D navigation technique in desktop VR for 3D applications

Project Purpose

In this study we will determine effectiveness of a two-handed 3D navigation technique by comparing with a one-handed navigation technique in desktop VR, which provides stereoscopic and head-tracking displays, for 3D applications.

Investigators

Isaac Cho, Computer Science
Zachary Wartell, Computer Science

Eligibility

You may participate in this study if you are 18 years of age or higher and if you have 20/20 vision, corrected vision to 20/20 or higher (i.e. you can clearly read text on a computer workstation monitor), can comfortably use your arms, hands and fingers for everyday tasks and you can communicate in spoken English.

Overall Description of Participation

In the first step, we will demonstrate stereoscopic display and head-tracking to make you familiar with them. Stereopsis is an important binocular depth cue generated by the differences between the two views of a scene seen from a person's two eyes. Head-tracking is a tracking system which tracks the position and an orientation of participant's head to generate an optimal perspective image. We will train you on how these technologies work by showing you how to view the 3D application and how to perform a navigation task (rotating, translating and scaling) with one-handed or two-handed 6DOF buttonballs. Also, we will survey your experiences with 3D applications (e.g. 3D games and 3D movies) and your familiarity with using a computer and user interface.

In the next step, you will perform a target-finding task with one-handed or two-handed buttonballs depending on the navigation technique conditions. There are 3 navigation conditions in the experiment: 1. one-handed navigation input mode, 2. two-handed "Spindle+Wheel with separate scale" navigation input mode and 3. Two-handed "Spindle+Wheel" navigation input mode. The task is to find a colored box on gridded ground and place the box between two boxes in the center of the screen. The box has to be bigger than the red wired box and smaller than the colored outer box. You, also, need to match the colors between the target box and the outer box.

After you finish the experiment, you will take a post-experiment questionnaire.

Length of Participation

Participation should take approximately 50-60 minutes.

Risks and Benefits of Participation

While using virtual environment systems, some people experience slight symptoms of disorientation, nausea, or dizziness. These can be similar to "motion sickness" or the feeling experienced in wide-screen movies and theme park rides. We do not expect these to be strong or to last after participants leave the laboratory. If you feel uncomfortable and wish to stop the experiment, you are free to do so at any time without penalty by simply announcing your desire to stop.

The primary scientific benefit of the study is to determine effectiveness of a two-handed navigation technique for 3D applications in desktop VR.

There is a benefit for participants who volunteer via the psychology department’s on-line subject pool. They will receive 1 credit in their general psychology class. Students who do not wish to participate, or who are excluded in the study due to stereoblindness or other exclusion criteria, may leave and will be granted participation credit for 15 minutes or one-half a credit.

Volunteer Statement

You are a volunteer. The decision to participate in this study is completely up to you. If you decide to participate in the study, you may stop at any time. You will not be treated any differently if you decide not to participate in the study or if you stop once you have started. You can request to withdraw your segment after the testing is complete.

Confidentiality Statement

Any information about your participation, including your identity, is completely confidential. The following efforts will be taken to protect confidentiality and privacy:

1) The informed consent form will be kept in a locked filing cabinet, separate from the rest of the data.

2) All participants will be assigned a random ID consisting two randomly-generated initials (initials will not correspont to participants' name). The participants will only be referred by assigned alphanumeric codes both in internal communication

between researchers or in the form of written reports.

3) The investigator and co-investigators will ask the participants not to mention their name or identify themselves during the recordings. The recording is only for internal use such as transcription and will not be made available to the public. Screenshots from the video recording might be published without disclosing the identify of any participants.

4) All digitally recorded files during the study will be kept in the Charlotte Visualization Center (room 437 in Woodward Hall) on password-protected computers. The files will be destroyed after two years by investigators under the guidance of the responsible faculty.

Statement of Fair Treatment and Respect

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the university's Research Compliance Office (704-687-3309) if you have questions about how you are treated as a study participant. If you have any questions about the actual project or study, please contact Isaac Cho (icho1@uncc.edu) or Dr. Zachary Wartell (zwartell@uncc.edu) at 704-687-8442.

Approval Date

This form was approved for use on *May 8th, 2013* for use for one year.

Participant Consent

I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principal investigator of this research study.

_______________________________________ ________________________
Participant Name (PRINT) DATE

____
Participant Signature

_______________________________________ ________________________
Investigator Signature DATE

Your given ID number (instructor only):

1. How much arm fatigue did you feel with the one-handed navigation technique?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

2. How much arm fatigue did you feel with the Two-handed with separate scale navigation technique?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

3. How much arm fatigue did you feel with the Two-handed with integrated scale navigation technique?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

Some people experience difficulty in perceiving a clear 3D stereoscopic image on the display. Often if the 3D image extends too far in front or behind the display surface, a person may perceive only two separate double images rather than a single 3D image. In some circumstances, a person may experience eye strain, visual fatigue or headaches when viewing this type of display system. Collectively, these negatives experiences are called “stereoscopic fusion problems”.

1. How frequently did you feel stereoscopic fusion problems with the One-handed navigation technique?

(Not At All) (Very Frequently)
1 2 3 4 5 6 7

2. How frequently did you feel stereoscopic fusion problems with the Two-handed with separate scale navigation technique?

(Not At All) (Very Frequently)
1 2 3 4 5 6 7

3. How frequently did you feel stereoscopic fusion problems with the Two-handed with integrated scale navigation technique?

(Not At All) (Very Frequently)
1 2 3 4 5 6 7

Your given ID number (Instructor only):

1. Overall, which object manipulation technique (One-handed, Two-handed with separate scale, or Two-handed with integrated scale) was the best for the navigation task? Why?

2. Which object manipulation technique (One-handed or Two-handed) was better than other for rotating? Why?

3. Which object manipulation technique (One-handed or Two-handed) was better than other for translating? Why?

4. Which object manipulation technique (One-handed, Two-handed with separate scale, or Two-handed with integrated scale) was better for scaling? Why?

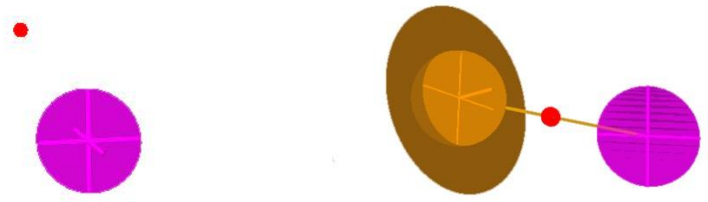
5. In the one-handed mode, which scale technique (rate control or position control) did you prefer?

6. In the two-handed techniques, did the pitch control (rotating the ball) help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

7. When comparing the two-handed with separate scale to two-handed with integrated scale, did having scale separated (via a separate button for scaling) help you accomplish the docking task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7



(A) (B)

8. In the two-handed techniques, how much did the wheel visual feedback on the 3D cursor help to control pitch rotation?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

9. Which object manipulation technique was the most intuitive to perform the task, One-handed, Two-handed with separate scale or Two-handed with integrated scale? Why?

10. When the docking task is complete a sound is played and the box's color changes. Which is more helpful to indicate that you accomplished the task?

Sound:
(Not At All) (Very Helpful)
1 2 3 4 5 6 7
Color:
(Not At All) (Very Helpful)
1 2 3 4 5 6 7

11. In the two-handed techniques, did presence of the orange line between the 3D cursors help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

12-1. In the one-handed technique, did the red sphere, which shows the center of the scale, help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

12-2. In the two-handed techniques, did the red sphere in the middle of the orange line connecting the 3D cursors help you to accomplish the task?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

13. When the target box size differs from the center box size, how much does this add to the task difficulty (compared to when the sizes are the same)?

(Not At All) (Very Difficult)
1 2 3 4 5 6 7

14. Did the orientation difference between the center box and the target box make the task difficult to accomplish?

(Not At All) (Very Difficult)
1 2 3 4 5 6 7

15. If you have any comments on this study, the display technologies, or the interaction techniques, please give us feedback.

____

____

____

APPENDIX C: MATERIALS FROM AUTO-STEREO ADJUSTMENT TECHNIQUE EXPERIMENTS

This appendix contains materials used in Auto-adjustment technique Experiment 1 and 2, which was reported in Chapter 5. The following materials are included, listed in order of appearance:

1. The informed consent form
2. Confidence rate form
3. Post questionnaires form

UNC CHARLOTTE
The University of North Carolina at Charlotte

Informed Consent for
Evaluation of a two-handed 3D navigation technique and auto-adjustment of a scale factor for a Multi-Scale Virtual Environment

Project Purpose

This study examines the effectiveness of a several virtual travel techniques in interactive 3D computer applications. The display systems include two virtual reality systems which provide stereoscopic, head-tracked display: a desktop VR system and a CAVE system. The navigation techniques include a two-handed 3D navigation technique, a one-hand technique and an option for auto-adjustment of the view scale factor.

Investigators

Isaac Cho, Computer Science
Zachary Wartell, Computer Science

Eligibility

You may participate in this study if you are 18 years of age or higher and if you have 20/20 vision, corrected vision to 20/20 or higher (i.e. you can clearly read text on a computer workstation monitor), can comfortably use your arms, hands and fingers for everyday tasks and you can communicate in spoken English.

Overall Description of Participation

In the first step, we will demonstrate the stereoscopic, head-tracked display technology to familiarize you with them. Stereopsis is an important binocular depth cue generated by the differences between the two views of a scene seen from a person's two eyes. Head-tracking is a tracking system which tracks the position and orientation of a person's head to generate an optimal perspective image.

We will demonstrate how to navigate through an example 3D virtual environment using several input device technologies. The input devices act like 3-dimensional mice where the computer can track their position and orientation in 3-dimensional space. The devices are shaped like small balls with several buttons on them and are called "buttonballs."

We will survey your experiences with computers, 3D applications and various display technologies.

Stereoscopic fusion problem:

Some people experience difficulty in perceiving a clear 3D stereoscopic image on 3D displays. Often if the 3D image extends too far in front or behind the display surface, a person may perceive only two separate double 2D images rather than a single 3D image. In some circumstances, a person may experience eye strain, visual fatigue or headaches when viewing this type of display system. Collectively, these negatives experiences are called “stereoscopic fusion problems”.

In the experimental trials, you will perform a target-finding task in which you virtually travel through a 3D environment to a target location. Specifically, the task is to find a small, colored box on a 3D globe and to travel so that this target box is aligned with a second box which always remains centered on the display screen.

You will use four different interaction techniques to perform the travel operation. Two of the techniques will use a single buttonball while two will use two buttonballs, one in each hand. Additionally, two of the techniques will automatically adjust the 3D view scale while in the other two techniques you will manually control the 3D scale factor. Combining these two factors, the four techniques are: 1) one-handed buttonball input with auto scaling, 2) one-handed buttonball input with manual scaling, 3) two-handed buttonball input with auto-scaling, and 4) two-handed buttonball input with manual scaling.

After you finish the experiment trials, you will answer a post-experiment questionnaire.

Length of Participation

Participation should take approximately 50-60 minutes.

Risks and Benefits of Participation

When using virtual environment systems, some people experience slight symptoms of disorientation, nausea, or dizziness. These can be similar to "motion sickness" or the feeling experienced in wide-screen movies and theme park rides. In this experiment, we do not expect these to be strong or to last after participants leave the laboratory. If you feel uncomfortable and wish to stop the experiment, you are free to do so at any time without penalty by simply announcing your desire to stop.

The primary scientific benefit of the study is to determine effectiveness of a two-handed navigation technique and automatic view scaling for 3D applications in virtual reality applications.

There is a benefit for participants who volunteer via the psychology department’s on-line subject pool. They will receive 1 credit in their general psychology class. Students who do not wish to participate, or who are excluded in the study due to stereoblindness or other exclusion criteria, may leave and will be granted participation credit for 15 minutes or one-half a credit.

Volunteer Statement

You are a volunteer. The decision to participate in this study is completely up to you. If you decide to participate in the study, you may stop at any time. You will not be treated any differently if you decide not to participate in the study or if you stop once you have started. You can request to withdraw from the experiment even after you complete the experiment, in which case your performance data will be deleted and cannot be used to support our study.

Confidentiality Statement

Any information about your participation, including your identity, is completely confidential. The following efforts will be taken to protect confidentiality and privacy:

1) The informed consent form will be kept in a locked filing cabinet, separate from the rest of the data.

2) All participants will be assigned a random ID consisting of two randomly-generated initials (initials will not correspond to participants' name). The participants will only be referred by assigned alphanumeric codes both in internal communication between researchers or in the form of written reports.

3) The investigator and co-investigators will ask the participants not to mention their name or identify themselves during any recordings. The recording is only for internal use such as transcription and will not be made available to the public. Screenshots from the video recording might be published without disclosing the identity of any participants.

5) All digitally recorded files during the study will be kept in the Charlotte Visualization Center (room 437 in Woodward Hall) on password-protected computers. The files will be destroyed after two years by investigators under the guidance of the responsible faculty.

Statement of Fair Treatment and Respect

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the university’s Research Compliance Office (704-687-3309) if you have questions about how you are treated as a study participant. If you have any questions about the actual project or study, please contact Isaac Cho (icho1@uncc.edu) or Dr. Zachary Wartell (zwartell@uncc.edu) at 704-687-8442.

Approval Date

This form was approved for use on *May 24th, 2013* for use for one year.

Participant Consent

I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principal investigator of this research study.

____________________________________ ______________________
Participant Name (PRINT) DATE

____
Participant Signature

____________________________________ ______________________
Investigator Signature DATE

Your given ID number (Instructor only):

Stereoscopic fusion Problem:

Some people experience difficulty in perceiving a clear 3D stereoscopic image on the display. Often if the 3D image extends too far in front or behind the display surface, a person may perceive only two separate double images rather than a single 3D image. In some circumstances, a person may experience eye strain, visual fatigue or headaches when viewing this type of display system. Collectively, these negatives experiences are called “stereoscopic fusion problems”.

Task1: Finding a target box
Task2: looking inside of surrounding boxes

<u>Condition: No auto-adjustment mode</u>

1. How much arm fatigue did you feel now?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

1. Overall, how frequently did you feel stereoscopic fusion problems during trials?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

2. How frequently did you feel stereoscopic fusion problem when you were working on the task1?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

3. How frequently did you feel stereoscopic fusion problem when you were working on the task2?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

Condition: Auto-adjustment of the view scale factor

1. How much arm fatigue did you feel now?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

2. Overall, how frequently did you feel stereoscopic fusion problems during trials?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

3. How frequently did you feel stereoscopic fusion problem when you were working on the task1?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

4. How frequently did you feel stereoscopic fusion problem when you were working on the task2?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

Condition: Auto-adjustment of the fusion distance

1. How much arm fatigue did you feel now?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

2. Overall, how frequently did you feel stereoscopic fusion problems during trials?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

3. How frequently did you feel stereoscopic fusion problem when you were working on the task1?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

4. How frequently did you feel stereoscopic fusion problem when you were working on the task2?

   (Not At All) (Very Painful)
   1 2 3 4 5 6 7

Your given ID number (Instructor only):

1. How frequently did you feel stereoscopic fusion problems without auto-adjustment modes?

(Not At All) (Very Frequently)
1 2 3 4 5 6 7

2. How much did the auto-adjustment of the view scale factor help to accomplish the task and to prevent stereoscopic fusion problems?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

3. How much did the auto-adjustment of the fusion distance help to accomplish the task and to prevent stereoscopic fusion problems?

(Not At All) (Very Helpful)
1 2 3 4 5 6 7

4. Which auto-adjustment mode (the view scale (A) or the fusion distance (B)) was better to accomplish the tasks and to prevent the stereoscopic fusion problem? Why?

5. Did the box size (4 different box sizes and visible/invisible without an arrow) affect your ability to accomplish the task?

If you have any comments for this study, please give us feedback.

APPENDIX D: MATERIALS FROM HYFINBALL EXPERIMENT

This appendix contains materials used in HyFinBall Experiment, which was reported in Chapter 6. The following materials are included, listed in order of appearance:

1. The informed consent form
2. Confidence rate form
3. Post questionnaires form

UNC CHARLOTTE
The University of North Carolina at Charlotte

Informed Consent for
Evaluating Usability of Hybrid User Interface in Desktop VR for a Visual Analytics Application.

Project Purpose

In this study we will determine usability of a hybrid user interface for combinations of 2D and 3D computer interaction tasks using a virtual reality computer display system for a scientific visualization computer application that displays terrain data.

Investigators

Isaac Cho, Computer Science
Zachary Wartell, Computer Science

Eligibility

You may participate in this study if you are above 18 and if you have 20/20 vision or corrected vision to 20/20, can comfortably use your arms and fingers and communicate in spoken English.

Overall Description of Participation

In the first step, we will demonstrate how to use a desktop virtual reality computer system. This system uses stereoscopic display and head-tracking to display 3D computer graphics. Stereopsis is an important depth cue generated by the differences between the two views of a scene seen from a person’s two eyes. Head-tracking is a tracking system which tracks the position and orientation of person’s head to generate an optimal perspective 3D image. We will demonstrate how these technologies can be used with a scientific visualization computer program. We will show you how to interact with the program’s 2D and 3D visualizations using several different input devices including a computer mouse and a pair of 3D mouse-like devices called 6DOF (“degree-of-freedom”) buttonballs.

We will survey your past experience with 3D computer programs and 3D media (e.g. 3D games and 3D movies) and your familiarity with using a computer and various user interfaces.

In next step, you will perform a series of tasks using the visualization software. These tasks will require interacting with either 2D only graphics, 3D only graphics or a combination of 2D and 3D graphics. 2D task examples include interacting with 2D menus and manipulating a scatter-plot by repositioning the plot, selecting points in the plot, or changing the data displayed in the plot. 3D task examples include manipulating the view of a 3D object and creating and adjusting a 3D selection box.

You will perform tasks with one of four different input devices. One condition uses a single computer mouse. The second condition uses a pair of 6DOF buttonballs each of which behaves like a 3D mouse. You hold each device in one hand in the air and can freely move or rotate it 3D space. The third condition uses both 6DOF buttonballs, but you will hold them while resting yours hands on a desk similar to using two regular computer mice. This condition is called “planar-3DOF mode” because your hands move on a planar surface (the desk) and each buttonball can be moved in two directions and also rotated. In the fourth condition you can switch between using each buttonball in either planar-3DOF mode (hand resting on the desk) or 6DOF mode (hand in the air). The program will detect which mode is activated and alter how the cursors are displayed and how the user interface works.

Participants will be divided in two groups for the experiment and each group will perform different types of tasks. Group 1 will perform 2D and combined 2D plus 3D tasks. Group 2, will perform 3D and combined 3D plus 2D tasks.

After you finish the experiment, you will take a post-experiment questionnaire.

Length of Participation

Participation should take approximately 50-60 minutes.

Risks and Benefits of Participation

While using virtual reality display systems, some people experience slight symptoms of disorientation, nausea, or dizziness. These can be similar to "motion sickness" or the feeling experienced in wide-screen movies and theme park rides. We do not expect these to be strong or to last after participants leave the laboratory. If you feel uncomfortable and wish to stop the experiment, you are free to do so at any time without penalty by simply announcing their desire to stop.

The primary scientific benefit of the study is to compare the usability of several user interfaces for a scientific visualization computer program which displays 2D and 3D graphical objects and requires 2D and 3D interactions with these graphics.

There is a benefit for participants who volunteer via the psychology department’s on-line subject pool. They will receive 1 credit in their general psychology class. Students who do not wish to participate, or who are excluded in the study due to stereoblindness, may leave and they are granted participation credit for 15 minutes or one-half a credit.

There is a benefit for participants who are students in ITCS 6125/8125 Virtual

Environments. This course is about virtual reality computer software and hardware. This study studies a user interface that uses virtual reality display and input devices. Students participating in the experiment will benefit by hands-on learning about a current research project in virtual reality. The experiment will engage students with an additional virtual reality system beyond those used in their class projects. Students will be required either to participate as homework credit but will be given the option of having their data not collected for experimental analysis. This meets IRB protocol requirements for course homeworks that require participation in an experiment. The informed consent form below has a special section for ITCS 6125/8125 students informing them of this option.

Volunteer Statement

You are a volunteer. The decision to participate in this study is completely up to you. If y ou decide to participate in the study, you may stop at any time. You will not be treated an y differently if you decide not to participate in the study or if you stop once you have start ed. You can request to withdraw your segment after the testing is complete.

If you are enrolled in course ITCS 6125 or ITCS 8125, your participation in this study counts as a required homework assignment. However, having your data recorded and used for the actual research study is voluntary.

Confidentiality Statement

Any information about your participation, including your identity, is completely confidential. The following efforts will be taken to protect confidentiality and privacy:

1) The informed consent form will be kept in a locked filing cabinet, separate from the rest of the data.

2) All participants will be assigned a random ID consisting two randomly-generated initials (initials will not correspont to participants' name). The participants will only be referred by assigned alphanumeric codes both in internal communication between researchers or in the form of written reports.

3) The investigator and co-investigators will ask the participants not to mention their name or identify themselves during the recordings. The recording is only for internal use such as transcription and will not be made available to the public. Screenshots from the video recording might be published without disclosing the identify of any participants.

4) All digitally recorded files during the study will be kept in the Charlotte Visualization Center (room 437 in Woodward Hall) on password-protected computers. The files will be destroyed after two years by investigators under the guidance of the responsible faculty.

Statement of Fair Treatment and Respect

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the university's Research Compliance Office (704-687-3309) if you have questions about how you are treated as a study participant. If you have any questions about the actual project or study, please contact Isaac Cho (icho1@uncc.edu) or Dr. Zachary Wartell (zwartell@uncc.edu) at 704-687-8442.

Approval Date

This form was approved for use on *May 8th, 2013* for use for one year.

Participant Consent

I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principal investigator of this research study.

_______________________________________ ________________________
Participant Name (PRINT) DATE

____
Participant Signature

_______________________________________ ________________________
Investigator Signature DATE

Your Given ID (Instructor only):

<u>6DOF Buttonballs only:</u>

Fatigue Rate

1. How much arm fatigue did you feel with the 6DOF buttonballs only input?

(Not At All) (Very Painful)
1 2 3 4 5 6 7

2D tasks

2. To what extent do you feel that you were able to accurately perform the 2D tasks by the 6DOF buttonballs input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

2D to 3D tasks

3. To what extent do you feel that you were able to accurately perform the 2D to 3D tasks by the 6DOF buttonballs only input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

3D to 2D tasks

4. To what extent do you feel that you were able to accurately perform the 3D to 2D tasks with the 6DOF buttonballs only input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

Occlusion Problem

Sometime the occlusion problem that 2D (a scatterplot or menu icons) and 3D (a terrain or a selection box) objects overlap each other increases task difficulty.

5. How much did you have the occlusion problem with the 6 DOF buttonballs only input?

(Not At All) (Very frequently)
1 2 3 4 5 6 7

<u>6DOF and planar constraint DOF buttonballs:</u>

Fatigue Rate

1. How much arm fatigue did you feel with the 6DOF and planar constraint DOF buttonballs input?

(Not At All) (Very painful)
1 2 3 4 5 6 7

2D tasks

2. To what extent do you feel that you were able to accurately perform the 2D tasks by the planar constraint DOF buttonballs input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

2D to 3D tasks

3. To what extent do you feel that you were able to accurately perform the 2D to 3D tasks by the 6DOF and planar constraint DOF buttonballs input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

3D to 2D tasks

4. To what extent do you feel that you were able to accurately perform the 3D to 2D tasks with the 6DOF and planar constraint DOF buttonballs input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

Occlusion Problem

Sometime the occlusion problem that 2D (a scatterplot or menu icons) and 3D (a terrain or a selection box) objects overlap each other increases task difficulty.

5. How much did you have the occlusion problem with the 6DOF and planar constraint DOF buttonballs input?

(Not At All) (Very frequently)
1 2 3 4 5 6 7

6DOF buttonball and mouse:

Fatigue Rate

1. How much arm fatigue did you feel with the 6 DOF buttonballs with mouse input?

(Not At All) (Very painful)
1 2 3 4 5 6 7

2D tasks

2. To what extent do you feel that you were able to accurately perform the 2D tasks by the mouse input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

2D to 3D tasks

3. To what extent do you feel that you were able to accurately perform the 2D to 3D tasks by the 6 DOF buttonballs with mouse input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

3D to 2D tasks

4. To what extent do you feel that you were able to accurately perform the 3D to 2D tasks with the 6 DOF buttonballs with mouse input?

(Not At All) (A great deal)
1 2 3 4 5 6 7

Occlusion Problem

Sometime the occlusion problem that 2D (a scatterplot or menu icons) and 3D (a terrain or a selection box) objects overlap each other increases task difficulty.

5. How much did you have the occlusion problem with the 6 DOF buttonballs with mouse input?

(Not At All) (Very frequently)
1 2 3 4 5 6 7

Your given ID number (Instructor only):

Use the list below for reference when answering the following questions. You may write just the letter (A-C) that labels the input mode(s) you want to refer to in your answers.

Conditions:
6 DOF buttonballs only
6 DOF + table constrained buttonballs input
6 DOF buttonballs + mouse input

In the questions below, the term 2D tasks refers to the tasks such as scatter-plot manipulation or menu selection, while 3D tasks refers to camera navigation and 3D selection box manipulation.

1. Which input mode (A to C) was the easiest to accomplish the overall tasks? Why?

2. Of the input modes, A,B and C, which is the easiest for performing 2D tasks? Why?

3. Of all input modes (A to C) which one was the easiest to perform a pair of tasks where a 2D task followed by a 3D task? Why?

4. Of all input modes (A to C) which one was the easiest to perform a pair of tasks where a 3D task followed by a 2D task? Why?

5. Overall, when comparing your experience using the mouse compared to using the pair of table constrained buttonballs, which do you prefer:

Strongly Prefer Mouse	Somewhat Prefer Mouse	Neutral	Somewhat Prefer Buttonball	Strongly Prefer Buttonball
-2	-1	0	+1	+2

5-1. Regarding your above answer, what reasons led to your preference?

6. Did the height difference (above or between the terrain) of the target 3D selection box effect on your ability to accomplish the 3D task?

(Not At All) (Somewhat) (Very Much)
1 2 3 4 5 6 7

7. When using the combination 6DOF plus the mouse, to what degree do you think the time spent switching between holding the mouse and holding the buttonball affected your ability to quickly complete the task?

(Not At All) (Somewhat) (Very Much)
1 2 3 4 5 6 7

8. When using the combination 6DOF plus the mouse, to what degree do you think switching between holding the mouse and holding the buttonball affected your physical comfort while completing the task?

(Not At All) (Somewhat) (Very Much)
1 2 3 4 5 6 7

9. When using the combination 6DOF plus the mouse, how frustrating was switching between holding the mouse and holding the buttonball?

(Not At All) (Somewhat) (Very Much)
1 2 3 4 5 6 7

Learnability

10-1 Of all input modes (A to C) which mode is the most difficult to learn when performing both 2D and 3D interaction techniques?

10-2 Of all input modes (A to C) which mode is the most difficult to learn the 2D interaction techniques?

10-3 Of all input modes (A to C) which mode is the most difficult to learn the 3D interaction techniques?

If you have any comments for this study, please give us feedback.