Tangible User Interfaces and Metaphors for 3D Navigation

Andrew Blakney

A Thesis

in

The Department

of

Computer Science and Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Computer Science at
Concordia University
Montreal, Quebec, Canada

December 2010

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Andrew Blakney

Entitled: Tangible User Interfaces and Metaphors for 3D Navigation

and submitted in partial fulfillment of the requirements for the degree of

Master of Computer Science

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Sabine Bergler _______________ Chair
Dr. Thiruvengadam Radhakrishnan Examiner
Dr. Peter Grogono _______________ Examiner
Dr. Sudhir Mudur________________ Supervisor

Approved by

Dr. Hovhannes Hrutyunyan______________________
Chair of Department or Graduate Program Director

Robin Drew_____________________________
Dean of Faculty

Date Tuesday, December 14, 2010

ABSTRACT

‘Tangible User Interfaces and Metaphors for 3D Navigation’

Andrew Blakney

The most fundamental and common 3D interaction is the control of the virtual camera or viewpoint, commonly referred to as navigation. The navigational requirements of controlling multiple degrees of freedom and maintaining adequate spatial awareness are big challenges to many users. Many tasks additionally demand large portions of cognitive effort from the user for non-navigational aspects. Therefore, new solutions that are simple and naturally efficient are in high demand. These major challenges to 3D navigation have yet to be satisfactorily addressed, and as a result, there has yet to be a declaration of a suitable unified 3D interaction technique or metaphor.

We present a new domain and task independent 3D navigation metaphor, Navigational Puppetry, which we intend to be a candidate for the navigational portion of a unifying 3D interaction metaphor. The major components of the metaphor - the puppet, puppeteer, stage, and puppet-view - enable a new meta-navigational perspective and provide the user with a graspable navigational avatar, within a multiple-view perspective, that allows them to ‘reach’ within the virtual world and manipulate the viewpoint directly. We position this metaphor as a distinct articulation of the front wave of a puppetry related trend in recent 3D navigation solutions. The metaphor was implemented into a tangible user interface prototype called the Navi-Teer. Two usability studies and a unique spatial audio experiment were completed to observe and demonstrate, respectively, the metaphor’s benefits of tactile intimacy, spatial orientation, easy capture of complex input and support for collaboration.

ACKNLOWEDGEMENTS

I would like to thank my supervisor Dr. Sudhir Mudur for his support and guidance throughout this research. This project was funded through Dr. Mudur via research grants from NSERC, GRAND NCE, and Concordia University’s ENCS Faculty. I would also like to thank Samuel Thulin for his participation in the spatial sound experiment and for all of his valuable expert input and feedback on this research. Thanks also go out to all of the participants of our usability studies - who are too numerous to name individually, but you know who you are - for their time and energy. I would like to also thank the members of the TAG Research Group for their valuable feedback and for helping to arrange the meetings with game level designers. Finally, I would like to thank my family and friends for all of their encouragement.

TABLE OF CONTENTS

LIST OF FIGURES........VIII

1.0 INTRODUCTION........1

1.1 MOTIVATION AND PROBLEM STATEMENT........1

1.2 METHODOLOGY........3

1.3 SUMMARY OF CONTRIBUTIONS........3

1.4 THESIS ORGANIZATION........4

2.0 3D INTERACTION........6

2.1 3D USER INTERFACES........8

2.1.1 3DUI ELEMENT #1 - IT / IM........10

2.1.2 3DUI ELEMENT #2 - INPUT........11

2.1.3 3DUI ELEMENT #3 - OUTPUT........15

2.1.4 3DUI ELEMENT #4 - SYSTEM / VIRTUAL........17

2.2 GENERAL 3D INTERACTION CHALLENGES........18

2.2.1 3D PERCEPTUAL / MOTION CUES........18

2.2.2 STRIVING FOR NATURAL INTERACTION........20

2.2.3 3DUI SOLUTION DESIGN AND IMPLEMENTATION ISSUES........21

3.0 3D NAVIGATION........24

3.1 TRAVEL........26

3.2 WAYFINDING........27

3.3 3DUI NAVIGATION SOLUTIONS........31

3.3.1 INTERACTION TECHNIQUE / INTERACTION METAPHOR DRIVEN SOLUTIONS...32

3.3.1 INPUT DRIVEN SOLUTIONS........36

3.3.2 OUTPUT DRIVEN SOLUTIONS........39

3.3.4 VIRTUAL / SYSTEM DRIVEN SOLUTIONS........45

3.3.5 ENTIRE INTERFACE DRIVEN SOLUTIONS........49

3.4 3D NAVIGATION SOLUTION CHALLENGES........50

3.4.1 GENERAL CHALLENGES........50

3.4.1 VIRTUAL NAVIGATION VS. REAL-WORLD NAVIGATION........52

4.0 A NEW FRONTIER FOR 3D NAVIGATION SOLUTIONS........55

4.1 TANGIBLE USER INTERFACES........56

4.2 TANGIBLE USER INTERFACES FOR 3D NAVIGATION........60

5.0 NAVIGATIONAL PUPPETRY........65

5.1 THE NAVIGATIONAL PUPPETRY METAPHOR........65

5.2 IDENTIFYING THE PUPPETRY TREND........68

5.3 THE NAVI-TEER PROTOTYPE........75

5.3.1 THE NAVI-TEER INTERACTION PROTOCOL........ 78

6.0 STUDYING NAVIGATIONAL PUPPETRY........ 83

6.1 USABILITY STUDY #1 – BASIC NAVIGATIONAL BEHAVIORS........ 83

6.2 USABILITY STUDY #2 – SPATIAL ORIENTATION AND WAYFINDING........ 92

6.3 SPATIAL AUDIO EXPERIMENT WITH EXPERT........ 97

6.4 DISCUSSION WITH GAME LEVEL DESIGNERS........ 109

7.0 CONCLUSIONS AND FUTURE WORK........ 111

7.1 CONCLUSIONS........ 111

7.2 FUTURE WORK........ 112

7.2.1 DEVELOPING NEW NAVIGATIONAL PERSPECTIVES........ 113

7.2.2 CREATING A NEW DEFINITION / TAXONOMY FOR 3D NAVIGATION SOLUTIONS114

7.2.3 DEVELOPING NEW 3D NAVIGATION PERFORMANCE METRICS........ 114

7.2.4 DEVELOPING A NEW 3D NAVIGATION SOLUTION EVALUATION METHOD........ 115

REFERENCES........ 117

APPENDIX A – PRE-EXPERIMENT QUESTIONNAIRE........ 127

LIST OF FIGURES

Figure 1 – The Navigational Puppetry Metaphor ……………………………………65

Figure 2 – Puppet Prototype Version #1 ………………………………………………..76

Figure 3 – Puppet Prototype Version #2 ………………………………………………..77

Figure 4 – Stage-Handle …………………………………………………………………...77

Figure 5 – The Navi-Teer Prototype ……………………………………………........78

Figure 6 – Interacting with the Navi-Teer ……………………………………………....79

Figure 7 – Forward Facing Interaction Slice ……………………………………………82

Figure 8 – Training Environment ………………………………………………………85

Figure 9 – Targeted Motion Experiment ………………………………………………..86

Figure 10 – Relative Positioning Experiment………………………………………………87

Figure 11 – Information Gathering Experiment ……………………………………...88

Figure 12 – Disappearing Landmarks Experiment ……………………………………...93

Figure 13 – City Locations Experiment …………………………………………………94

Figure 14 – Using the Navi-Teer for Spatial Sound ………………………………….....98

Figure 15 – Spatial Sound Experiment ………………………………………………...102

1.0 Introduction

1.1 Motivation and Problem Statement

When interacting in a 3D virtual environment (VE), the most fundamental action is the control of the virtual camera or viewpoint – which is the user's window or eyes into the virtual world. The overall action of 3D navigation decomposes into 3D travel, which is the process of controlling the viewpoint's translation and orientation within the 3D VE, and wayfinding, which is the cognitive process behind determining where to travel. It has proven to be a difficult task for many users and is an important area of research due to a growing demand for efficient solutions in increasingly diverse situations - from gaming to architectural design. This is also due to the persistence of many basic 3D human computer interaction (HCI) problems. Users must maintain a conceptual spatial orientation and awareness of their position and surroundings within the 3D VE while also controlling multiple degrees-of-freedom (DOF) simultaneously. There are still many opportunities for innovation within the field of interfaces for 3D interactions and navigation. For example, there has yet to be a declaration of a unified 3D interaction solution for any of the major 3D interactions – navigation, selection, or manipulation. Bowman et al. point out further absences of standardized progress that suggest that, as a research community, we have yet to adequately define what constitutes a standard 3D interaction, interface or solution since there are yet to be any widely accepted standard input devices, display devices, or prototyping tools for 3D interfaces [4].

We argue that the recent popularity of tangible user interfaces (TUI) and developments in relevant interface technologies – such as touch-sensitive interactive surfaces, projection systems, sensors, optical tracking and haptic feedback devices – create opportunities for imaginative interface designs to yield new interactive experiences and better affordances for existing tasks. TUIs can exploit these technologies by allowing the line between the system and the user to be liberally drawn around sensed physical objects and interactive surfaces. This makes the design of TUI systems particularly unique due to the freedom it allows – innovative conceptually designed interaction techniques or metaphors can be tangibly sculpted with an implementation fidelity limited only by the efficiency with which relevant technologies are applied.

We examine the problem of 3D navigation with a specific focus on the relationship between creating new interaction metaphors and the effect of the design freedoms and interactive experience potentials afforded by TUIs [24] [44] [51]. We present a new interaction metaphor, Navigational Puppetry, and its prototype implementation The Navi-Teer – a TUI 3D navigation solution. The metaphor and prototype demonstrate the interactive benefits of spatial orientation, tactile intimacy, easy capture of complex navigational input and support for collaboration (both in an observational and interactive sense). It inspires a new way of looking at 3D navigation by arranging the human computer interface to allow the user to navigate by reaching into the virtual world and grasping a miniature version of their navigational avatar. We intend our TUI 3D navigation solution be a candidate for the foundation of the navigational portion of a unified 3D interaction metaphor.

1.2 Methodology

The goal of this research was to create and investigate a new domain and task independent 3D navigation metaphor and prototype, which we intend to be a candidate for the navigational portion of a unifying 3D interaction metaphor. We position it as a distinct articulation of the front wave of an observed puppetry trend in 3D navigation solutions that use similar tangible objects and interfaces. Following a thorough literature review of 3D Interaction, 3D Navigation, interface technologies, TUI theory, and existing navigation and TUI solutions, an initial design and prototype of the navigational puppetry metaphor were created. As proof of concept and to get inputs for refining the prototype, two usability studies and a unique spatial audio experiment were designed and completed.

1.3 Summary of Contributions

The most significant contributions of this research are the design and presentation of Navigational Puppetry - a distinctly new interaction metaphor - and its prototype implementation, the Navi-Teer - a TUI 3D navigation solution. In addition, we also present relevant background information, related works (including an articulation of a puppetry trend in navigation solutions), and results from two usability studies.

We also present a demonstration of the interactive benefits of the navigational puppetry metaphor and the Navi-Teer prototype when used for the task of 3D audio soundscape modeling and experience. The contribution of this experiment is a unique audiovisual case study that contributes to the fields of sound related TUIs and 3D navigation. It also suggests a new perspective pertaining to audio manipulation by adding a tangible puppetry link to the sounds and auditory perception point.

1.4 Thesis Organization

We will now describe the actual organization of this thesis. Thus far, we have provided a description of the motivation for this research and a brief introduction to our contributions. The following chapters will provide the necessary background information, describe our contributions in further depth, and will conclude with a re-summarization of our work and articulations of a few options for future work.

Chapters 2 and 3 both describe the essential background information for this research. First, chapter 2 sets the stage by providing a comprehensive overview of the entire field of 3D interaction and interfaces. Chapter 3 then focuses closer on the particular sub-area that our work concentrates on - 3D navigation. Special attention is made to give an adequate understanding of basic navigation theory, common navigation challenges, and of the elemental nature of 3D navigation solutions.

Chapter 4 describes our motivation to investigate the combination of 3D navigation and TUIs. Further background information is provided on the subject of TUIs and their suitability for 3D navigation in analyzed.

Chapter 5 presents the main contributions of our work – the Navigational Puppetry metaphor and the Navi-Teer prototype – and relates them to a description of the puppetry trend that was observed in many existing solutions during the course of our research. This chapter also provides the interaction protocol for the Navi-Teer.

Chapter 6 describes the results of our experimental investigations into the navigational puppetry metaphor via the Navi-Teer prototype. It begins by presenting the experimental designs and results for the two usability studies that were completed. It concludes by detailing the spatial sound experiment that was completed with an audio expert.

This thesis concludes with chapter 7, which begins with a succinct restatement of our major contributions. We then outline four possibilities for future work that were discovered during the course of this research - developing new navigational perspectives, creating a new definition / taxonomy for 3D navigation solutions, developing new 3D navigation performance metrics, and developing a new 3D navigation solution evaluation method.

2.0 3D Interaction

The term 3D interaction refers to a sub-group of the entire spectrum of human-computer interactive experiences that describes a communication between the user(s) and the system that is three dimensional (3D) and spatially oriented in nature, involving varying degrees of both 3D input and 3D output from the system. It is an obvious extension of the ever evolving human-computer communication dynamic that began in the ‘dark ages’ of HCI with punch-cards and keystrokes. The invention of the mouse in the 1960s began an interactive trend that continues to this day towards better and increasingly complex HCI through the simultaneous exploitation of humans’ natural interactive potentials and current interface technologies. 3D interactions remain one of the most important frontiers for HCI research and interface development. It is still an unconquered territory due to the continued absence of a complete and unified 3D interaction technique or interaction metaphor.

We further define the term 3D interaction within this document as a term which also refers to the abstract HCI experience that involves performing actions or tasks that are within the conceptual framework of three distinct dimensions [3] [4] [5]. The abstract nature refers to its independent existence from the physical and technological implementation issues that afford the actual interactive experience. The mere presence of a 3D output display (such as a 3D projector) and / or an input device which captures three or more DOF does not imply a 3D interaction is taking place. For example, Tic-Tac-Toe is a 2D game regardless of how a player sees the board or how they select the square to enter their ‘letter’. Contrarily, architectural design is definitely a 3D task which requires conception and interactions within all three spatial dimensions, but it does not explicitly require entirely 3D input or output (ie: pencils and 2D Blueprints).

The best way to describe 3D interaction in practice is to consider it within the context of a VE, which is essentially an artificially created 3D space, containing any number of 3D objects or structures, within which 3D interactions are carried out by the user. These virtual spaces are obviously meant to mimic environments in reality, but their actual compositions and realistic fidelities vary according to their purposes. One of the simplest possible examples of pure 3D interaction would be a user moving a virtual camera or viewpoint (their view into the VE) in order to visually inspect the surroundings. The next logical thing that a user would want to do in a VE would be to actually interact with the objects within the environment. In addition, both the user's movement and their object interactions may be augmented by employing various system or virtual functionalities. These simple interactive descriptions constitute the four main categories for 3D interactions [3] [4] [5] [31]:

- Navigation – changing the location and orientation of the viewpoint, or user's view into the VE
    - ex: moving from point A in the VE to point B, in order to get a better view of three virtual apples
- Selection – the choosing or picking of objects, or other aspects of the VE, that are of further interest
    - ex: selecting the largest of three virtual apples in order to do X

- Manipulation – the changing of certain VE object parameters such as location, orientation, or scale
    - ex: rotating a virtual apple in order to visually inspect the entire surface's texture
- System Control – any other interaction that involves invoking the actual system or any purely virtual controls
    - ex: clicking a button to save a screen shot of a particular view of the virtual apple

2.1 3D User Interfaces

We will now define the term 3D User Interface (3DUI), as we understand it in relation to our work, and articulate its major elements. As previously stated, the term 3D interaction is an expansive abstract term for interactions that involve performing actions or tasks within three distinct dimensions. We therefore define a 3DUI to be the vehicle for which the solutions to 3D interaction problems or tasks are realized. In its most inclusive sense, a 3DUI solution refers to all design and implementation elements pertaining to solving a given 3D interaction task or affording a particular 3D interaction experience.

An important term relating to 3DUI solutions is an Interaction Technique (IT). It describes the actual interactive procedure that is happening between the user and the system to afford the completion of the given 3D task. In its simplest form, it could just be protocols for the user to follow regarding how they should perform input actions and how they should interpret the output they receive from the system. With that in mind, if the

ITs are ‘like’ some familiar real-world activity, then pre-existing knowledge of it can predict certain aspects of the technique [53]. In this case, the IT could be described as an Interaction Metaphor (IM). The user’s exemplar concept of the metaphor contains certain pieces of information about the interaction it would describe. First it implies that certain interactions are possible and / or easy to do within the metaphor. Secondly, it also gives an idea of the interactive constraints that pertain to the metaphor [53]. A deciding factor for determining when an IM should be used instead of an IT would be if the task can be described easily using a metaphor that is commonly known, or at least easily learned and understood.

We define a complete 3DUI Solution as having all of the following elements, in addition to the actual details of the associated 3D task, user and computer system (which may include any and all technology and software to be used with the 3DUI):

- Interaction Technique (IT) / Interaction Metaphor (IM)
    - the actual descriptive protocol for the interaction
    - ex: VOODOO selection / manipulation technique [38], GO-GO selection technique [39], flying vehicle metaphor [53]
- 3DUI Input Element(s)
    - how the input is captured from the user and detected by the system
    - ex: Joystick, Mouse, Keyboard, etc.
- 3DUI Output Element(s)
    - how the system displays and communicates changes of state with the user

    - ex: LCD screen, Projected Surfaces, etc.
- 3DUI System or Virtual Element(s)
    - all software / virtual aspects of the 3DUI, excluding input and output hardware
    - ex: VE, visual aids, cursors, maps, etc…

2.1.1 3DUI Element #1 - IT / IM

The creation of ITs and IMs requires a thorough analysis of the task for which they are to be facilitating [3] [4] [5] [53]. If only a concise description of the interactive 'do's and do nots' is needed, then a technically written, instruction oriented IT may suffice. One of the major goals of defining a good IT / IM is to minimize the amount of cognitive resources required to use the IT in order to maximize the amount of attention the user retains for actually performing the task.

For example, explorations or searches of VEs are frequently described using the metaphor of moving or controlling a virtual camera. The use of a camera is a concept that is widely known and easy to understand. It is clearly defined at a high-level (excluding the lower-level technological aspects of camera operation). It is a highly appropriate metaphor for exploring a VE because cameras are actually used in reality to explore and document local and remote regions, such as footage from unmanned submarines. The virtual camera metaphor implies very specific things about what kind of interactions are afforded, namely the movement of the camera through 3D space and also changing the camera's orientation (where the lens is pointed) through rotation, pitch, and yaw. It could also suggest other things like zooming and play-back capability or even night-vision. The metaphor also suggests certain constraints to VE exploration based on cameras, such as they must always be moved from point A to point B in order to see point C, or more bluntly, cameras usually do not teleport themselves around.

There have been many 3D interaction metaphors that have been proposed and investigated [3] [4] [5] [31] [53]. The ultimate goal should be to create a unified 3D interaction metaphor or IT that would function similarly to the 2D pointing / desktop metaphor that has come to define the most common form of 2D HCI. To date there has been no unifying 3D interaction metaphor that has gained wide acceptance.

2.1.2 3DUI Element #2 - Input

As previously stated, 3D interaction does not necessarily dictate that 3D input devices are required. The input element of a 3DUI must capture the user's input as the task and underlying system dictate. Looking again at the example of exploring a VE using the metaphor of camera control, it is obvious that the task and system require that the input element of the interface captures enough input to control the camera. If only the camera's position and not its orientation is to be controlled by the user, then the input element must simply capture some kind of data that can be used to inform the three system parameters for its position (ie: x, y, and z coordinates). For example, if a 2D mouse was intended to be the only input element of the 3DUI, then the two DOF of the mouse's planar movement can control two parameters of the camera, while the third could be toggled in and out of control through a state change by clicking one of the mouse buttons.

Alternatively, the 3D task of VE exploration can also be facilitated by using an actual 3D input device, such as 3D Connexion's SpaceNavigator (www.3dconnexion.com).

The main difference between 3D and 2D input devices is the parallel of the main distinction between interactions in 2D and 3D – the use of the device is done conceptually and physically within three spatial dimensions [3] [5]. Most 3D input devices involve either the tracking or sensing of translation and / or orientation changes within a pre-determined interaction space. Many 3D input devices are merely some kind of physical object that is attached to a sensor. For example, the Polhemus 6 DOF sensor became widely used in data gloves, wands, and styli because it could easily provide input data on 3D translation and orientation. In addition to sensors, optical tracking using cameras and markers also allows for 3D input since all the spatial changes of the marker and anything that it is attached to can be dynamically captured.

3D input is usually categorized as either direct or indirect user interaction [31]. With direct interaction, the user's input depends on an extremely transparent and natural mapping between what the user does and how the system reacts and behaves. For example, tracking hand gestures, wand pointing, and gaze direction are all examples of direct user input. An obvious common vein is that direct interaction always either involves a physical device or the user's physical body. Indirect user interaction, on the other hand, usually lacks the blatantly obvious natural mapping between input and system reaction and also does not necessarily need to be tangibly physical in nature. Examples of indirect interaction input could be physical controls like joysticks, steering wheels, buttons, or even purely virtual controls like sliders or dials. The main difference between virtual and physical indirect interactions is the presence of haptic feedback and tactile intimacy that comes with manipulating actual objects.

The design of new devices / objects, or the choice of existing 3D input devices, is dependent on three main factors [11] [12] [13] [14]:

- The device's shape and physical affordances determine the integration or separability of the DOFs under control
    - for example: Masliah et al. have shown that in certain 3D tasks users typically perform translations and rotations separately [29], which implies that devices that separate them into distinct sensors (ie: one 3 DOF sensor for each) may perform better at certain tasks than integrated 6 DOF devices
- The kind of sensors, or combinations of sensors, that are used to capture the user's input:
    - isotonic – captures displacement or movement (ex: standard mouse, Polhemus 6 DOF sensor)
    - elastic – captures applied force, while providing a counterforce (ex: spring-loaded joystick)
    - isometric – captures applied force, with no counterforce (ex: IBM's TrackPoint laptop pointer stick)
    - other (ie: optical tracking, etc…)

- Transfer function – a decision on how to translate the sensor input data to the appropriate virtual behaviours:
    - Position Control – how to interpret the input sensor data for the movement of virtual objects
    - Rate Control – how to interpret the input sensor data for the velocity of virtual object movement

One of the most popular 3D input devices is the wand [3] [5], which has control over all 6-DOF since many instances use a 3D isotonic sensor to continuously track their position and orientation. While the wand is essentially a single-handed device, it seems prop-based devices that afford two-handed interaction are increasing in popularity as 3D input devices [2] [12] [23]. For example, the Cubic Mouse [15] gained attention due to its integration of 12 DOFs into a single bimanual, isotonically sensed device. It consists of a box with 3 orthogonal rods passing through its faces (one rod for each dimension – up / down, side / side, front / back). The entire box is tracked in 6 DOF and each rod can be pushed, pulled and twisted.

Another popular 3D input device is 3DConnexion's SpaceMouse, whose design later inspired 3DConnexion's SpacePilot and SpaceNavigator (http://www.3dconnexion.com/). It is a commercially available 3D input device which is based on single-handed manipulation of an integrated and elastically sensed 6-DOF cylindrical handle. This series of '3D mice' illustrates just how subtle the difference can be between 3D and non-3D devices. The 'cap' of each device is manipulated in a 3D context, but just barely.

Manipulations of the cap - in the form of twists, pushes and pulls – move the cap ever so slightly in the small 3D space within the allowable limits of the cap's elastic sensors. These small movements represent symbolic intentions of larger and more deliberate extensions of the interaction. As another example of this, Fröhlich et al. describe prototypes for the GlobeFish and GlobeMouse 3D input devices, both of which are based on the principle of a trackball that is manipulated with small and indirect 3D 'intentions'. [14].

2.1.3 3DUI Element #3 - Output

The output elements of 3DUI solutions are extremely important and deserve careful attention. As with any kind of interface solution, the output element is primarily how the computer communicates with the user, usually visually and / or aurally. The output element typically has the responsibility of keeping the user informed on the current state of the system, alerting them of state changes, and showing them the results of their inputs and interactions. In the simple case of text-based interactions, the choice of output would only be required to display text in some readable way and the variations of other subtle output parameters has little effect on overall performance. However, in the context of 3D interactions, the issue of output becomes drastically more important and complex. As previously stated, 3D interaction implies performing tasks in a 3D spatial context and, if a 3D device is used, then the input is also performed spatially. The output of a 3DUI must try to display what is happening in a 3D spatial context. This introduces the issue of 3D perceptual cues and how the human brain reacts to various output solutions. The success of a 3DUI solution's output is basically dependent on how well it can allow the user's visual perception to spatially overlap with their conceptual interactive perception. For example, the typical desktop PC output display is a 2D screen that obviously can not adequately portray important 3D cues like depth, perspective, occlusion, relative motion gradient or motion parallax [2] [3].

The issue is further complicated since output solutions may not just fail at approximating a 3D view or keeping the user properly oriented, but in serving their function they may actually induce a form of motion-sickness. When a user is interacting with a 3DUI, there is a certain optical flow or graphical locomotion that is conceptually linked to their internal spatial orientation. If there are unexpected or counter-intuitive output behaviors, such as large delays between movement input and VE view updating, then the user may become nauseous with a form of simulator-sickness [21]. This problem is perhaps most evident in motion simulator theme park rides, such as the 'Star Tours' ride at Disney World's MGM Studios, where movements of the ride's seats are synchronized to a video display of a flight path. Since the movements of the seats can never be completely linked to the changing view, it frequently induces motion sickness and the operators must warn potential riders accordingly.

There are many ways to choose or construct output solutions to be better equipped to handle 3D interactions. One simple way is to increase the display size and / or widen the field of view. Many studies show that increasing both of these factors, commonly done through utilizing projectors and large projected screens, results in better performance in both 3D and 2D tasks [34] [48]. There are also many recent technological developments that increase the amount of 3D perceptual cues that are perceived by the user. Certain head-mounted-displays (HMDs) are capable of displaying 3D graphics due to the separation of output into distinct displays for each eye. Solutions of that nature, which mostly occlude the real-world, are called fully immersive displays - immersion in this case referring to the degree to which the user feels present within the VE [3]. There are also many semi-immersive choices, such as surround screen stereo monitors and 3D projectors. The level of immersion and the size of the display can also have the potential to either enable or disable collaboration during the interaction.

In addition to visual feedback, 3DUIs may also facilitate audio and haptic feedback as a form of output. The addition of audio, provided the interface possesses actual speakers and not just a single set of headphones, also greatly increases the potential for collaborative interaction [5].

2.1.4 3DUI Element #4 - System / Virtual

The final element of 3DUI solutions relates to the system and virtual aspects that are interrelated with, but entirely separate from, the IT / IM, input and output elements. This element exists as its own entity, but it also links all of the other interface elements together. Any element of the entire solution that aids in completing the task that is only afforded by the system or the virtual aspect of the 3DUI falls under this category. For example, in many first-person-shooter (FPS) games the user is often presented with a view of their environment that is augmented with radar, or a small map, and cross-hairs for shooting accuracy. These 3DUI elements, although they are affected by the input device and are displayed by the output element, are functionalities that are being virtually supplied and supported by the system to help the player complete their 3D task. System elements may not be limited to visual aids like cross-hairs and radar, but they may also be more inventive and expansive interaction aids like motion guides or constraints for navigation [49].

2.2 General 3D Interaction Challenges

As previously stated, 3D interaction is difficult for many users and presents unique interactive challenges that are not present with traditional 2D Windows-Icon-Menus-Pointer (WIMP) interactions. People tend to only superficially experience the act of interacting in 3D reality, while their actual understanding of all its subtle aspects is either quite low, completely subconscious, or non-existent. In other words, there is a big difference between experiencing reality in 3D and properly understanding it. Spatial disorientation and cognitive discontinuities - between the user's internal conception, the actual VE, and / or the reality of the 3D input space - are quite common for many novice users.

2.2.1 3D Perceptual / Motion Cues

3D interaction requires a high accuracy of cognitive faculties to successfully carry out activities like mental rotations and spatial relation interpretations. Even though humans live and operate within a 3D world, many people find it hard to conceptualize artificial 3D environments or contemplate certain spatial relationships between virtual objects [5]. When humans are moving and interacting in the real world, there are many environmental and perceptual aids that are being subconsciously utilized. They can be in the form of actual tactile sensations and haptic feedback about the shape, affordances, and constraints of a given object or also in the form of 3D visual perception cues, such as stereo depth.

For example, if a human reaches out to grab an apple that another human has thrown to them, they are actually performing a lot of mental operations about inferring the trajectory of the apple and also the location and orientation of the apple at the point when they should grasp it. There are aspects of this interaction which are quick behind-the-scenes versions of the more deliberate activity of deciding if a certain shape / object will fit through a given slot or hole [23].

The most relevant perceptual and motion cues for 3D interaction are divided into those that depend on mono or binocular vision and those that depend on being static or in dynamic motion. The main cues that are related to allowing the user to perceive the third dimension of depth are mainly static and monocular. The most important of these cues are linear perspective, aerial perspective, relative height and texture gradient perception. However, humans possess two eyes that enable seeing in 3D through stereopsis. An important perceptual cue involving this binocular aspect is motion parallax, which indicates depth based on the movement of the head and both eyes.

The goal of a good solution should be to create a high fidelity relationship between the actual VE and the user’s conception of the VE and their position within it. When the important perceptual cues are not adequately provided by the 3DUI solution, then the user’s experience and performance with the system will suffer. One of the first areas that the user will notice a problem is with their actual spatial orientation; they will feel that their continuous orientation is not adequately strong and / or they will be frequently disoriented or lost.

2.2.2 Striving for Natural Interaction

When compared to other previous forms of HCI, such as text-based interfaces or 2D web-surfing, the concept of 3D interaction is perhaps the most direct virtual equivalent of how human’s naturally interact within 3D reality. For example, thinking about walking towards an apple and picking it up to visually inspect it in reality requires very little translation to conceptualize the same event taking place as a 3D interaction within a VE. As a result, it is an obvious expectation for users, and a very important consideration for researchers and developers, to have the actual interactive experience be as natural as possible.

Perhaps the simplest way to ensure that the interaction can be perceived as natural is to ensure that the 3DUI affords real-time interactivity. For example, large lags between user input and the output display of its result creates an immediate sense of ‘unnaturalness’. The choice of IT / IM and 3DUI input element will also have a significant impact on the solution’s sense of naturalness. Certain ITs / IMs are varying in degrees of being natural - or more specifically in realism - and be quite abstract with little to no connection to reality. However, the actual interaction that results from realistic ITs / IMs may result in less perceived naturalism than more abstract ones. For example, ‘gaze-directed’ navigation metaphors that use head-tracking for input have no realistic counterpart, but navigating from a user’s gaze is an undisputedly simple IM which is easy to explain and understand. As a result, this could end up being more natural to the user than using a joystick and an IM about piloting airplanes, even though this has a direct real-world equivalent. Looking at the input element specifically, it is easy to see how the more graspable or ‘tangible’ the input device is, the more natural the interaction experience is.

2.2.3 3DUI Solution Design and Implementation Issues

The design of 3DUI solutions can be a difficult activity. First of all, this task is extremely multi-disciplinary in nature. 3DUI designers must possess a lot of diverse knowledge that reconciles the principles of interface / interaction design and the technical knowledge of the peculiarities involved with 3D interactions and tasks. 3D UI solution designs actually relate to many diverse areas of study, the most important of which are:

- Perceptual, cognitive, and behavioral sciences – important for understanding how people interact in real and artificial 3D environments (ie: how human’s use depth cues, how humans formulate navigational knowledge, etc…)
- Engineering – knowledge of materials, mechanics and other required areas for input / output element design and construction
- Advanced computer graphics programming – to be able to adequately understand the theoretical, mathematical, and programming fundamentals involved with 3D graphics

- Expertise of the end-users – most 3DUI solutions are designed for domains that are quite unfamiliar to the designers, so expert end-user's are frequently required to give guidance (ie: the designer of a 3DUI for architectural modeling is probably not going to be an architect themselves)
- Expertise on the current state of all relevant 3DUI technologies – important for the designers to be able to reach full technological and interactive potentials

Another important aspect of 3DUI solution design is the paradoxical practice of using non-3D design tools that were, at least initially, intended for traditional 2D desktop-type interaction scenarios and tasks. This perhaps highlights a very important design consequence of the continued non-existence of a unifying 3D IT / IM and standardized design support tools for 3DUI solutions.

The above design issues have obvious consequences for their eventual implementations, but there are also many implementation specific issues for 3DUIs. In general, it is usually a safe assumption that many 3DUI solutions must handle a relatively higher amount and diversity of input data than their 2D counterparts [4]. Many times these large and diverse input data sets must be significantly translated or 'massaged' to become recognizable or usable by the system. This can also be further complicated by the increasingly multimodal nature of many 3DUI input and output elements. Finally, the lack of standardized 3DUI design tools is mirrored by the lack of support for the evaluation of 3DUI solutions. The ad-hoc or case-specific nature of many 3DUI solution designs / implementations leads to a similar approach being used to initially validate prototypes.

Many researchers find themselves literally attempting to create new systems for quantifying both human performance and perceptions for their own 3D tasks [3]. This can become even more complicated when expert non-technical users are required for system evaluation.

3.0 3D Navigation

The most basic and fundamental type of 3D interaction is navigation. It involves the control over both the position and orientation of the user's viewpoint within some form of VE, sometimes in addition to other view specific parameters such as zoom factors and field-of-view scaling [21]. The act of navigation is often compared to the task of controlling a virtual camera's movement path and controlling the camera's orientation through rotations and pitch changes [53]. For example, in FPS games, a player's movement through the 3D world is basically the controlled motion of a virtual camera, which is used to display to a first person perspective view of the game's 3D world. This controlled movement and placement of the viewpoint in turn allows for the other common 3D interactions to take place, namely basic VE exploration and the selection and manipulation of objects located therein [3]. When working in domains such as 3D modeling or architectural design, the ability to easily and precisely place the viewpoint to obtain a specific perspective is integral to more interactively complicated tasks.

The process of controlling a viewpoint's motion through a VE essentially reduces to the control over 6 DOF of the viewpoint's position and orientation - the control of 3 DOF of translational movement, along each of the three dimensional axes, and 3 DOFs of orientation rotations (sometimes referred to as pitch, roll, and yaw). The most general navigational behaviors can be categorized at a high level into four main groups:

- Exploration – hypothesis or goal-free movements around the VE
    - Ex: searching, visual inspections

- Targeted movement – movements around the VE with distinct goals or hypotheses for locations and / or perspectives
    - Ex: docking, following moving objects
- Specific coordinate movement – movements within the VE intended to conclude at a given virtual coordinate location / perspective
    - Ex: moving to a specific location within the VE to activate a feature based on motion detection
- Specific trajectory movement - movements within the VE intended to conform to a given path or trajectory
    - Ex: flying around a given virtual object at a specified distance, height, and perspective

The overall action of 3D navigation can be further decomposed into the related, but separate, activities of travel and wayfinding. Travel is the locomotion related component of navigation and describes the process of the actual controlling of the viewpoint's movement through the 3D VE. Wayfinding is the cognitive process behind determining where to travel and is significantly more conceptually complex. [5]. Although they are distinct activities, it is quite easy to see that they are most often carried out in conjunction with each other and that the method of how one travels has an obvious effect on how one will find their way. For example, navigation for automobiles requires different wayfinding processes than boats on the sea. On the other hand, a certain choice of wayfinding method can also directly inform what kind of travel can or should be used.

For example, if the goal of the current wayfinding strategy is a quick visual exploration of a given area, then an airplane would be a much better travel method than a train.

3.1 Travel

3D travel has proven to be a difficult task for many users to perform. Aside from all the problems related to effects on perceptual cues in different 3D output display solutions, the basic task of simultaneously controlling multiple DOFs for 3D movement is a big challenge to most users [3]. 3DUI IT / IM and input solutions approach the issue by concentrating on how to intelligently restrict / enable the number of DOFs that are controllable at any one time [3] - or, more precisely, how to determine the best integrations or separations of different DOF control combinations [12]. The method of travel can also be considered as the bundle of navigational affordances that are presented to the user through the 3DUI IT / IM and input elements. For example, in most FPS games for the Xbox game system, the method of travel is usually thought of as a moving human body. The Xbox system's input element is a bimanual dual thumb-stick gamepad. The IT / IM, which can usually be slightly modified within each game's control settings, connects the gamepad to the control of the main navigation-related portions of a human body. Typically, one thumb-stick handles feet-related translations and one thumb-stick handles head-related orientation changes.

The act of moving a viewpoint around in a VE is equivalent to moving the VE around the viewpoint [3] [41] [53]. Due to how we travel in reality, the concept of looking at travel through the eyes of the traveler seems most natural; this is referred to as the egocentric view. This idea is facilitated by conceptualizing the motion of a dynamic traveling entity within a static environment. Alternatively, one can consider travel to be occurring if the environment itself becomes dynamic and moves underneath a static traveler. This perspective is called exocentric, since the focus shifts from seeing the traveler as the center of the interaction to the actual environment. The choice between egocentric and exocentric travel perspectives is an important aspect to consider in conjunction with the IT / IM and input elements of the 3DUI solution, since each perspective would require different affordances. Travel must be considered to be simply changes between the point of intersection of the traveler and the environment, and can be achieved by changing the positions of either point in that intersection.

3.2 Wayfinding

Wayfinding is the cognitive counterpart to travel that occurs ‘behind the scenes’ within the mind of the traveler or navigator [5]. It refers to deciding on a travel path through a VE or to all of the activities before, during, and after an intentional relocation from one place within the VE to another [16]. Successful wayfinding depends on the user building up a strong sense of spatial orientation and knowledge about the surrounding environment within the context of the task they are attempting to complete. The importance of wayfinding is best observed when navigation is considered in reality. For example, the efficiency of wayfinding strategies can mean life or death when traveling at sea or deep within remote wilderness areas.

There are four major steps in a typical wayfinding task [16] [34]:

- Orientation – recognition of surrounding environment / objects and an articulation of the interactive goal or desired destination
- Route decision – choosing a path or series of paths to get to a desired location in the environment
- Route monitoring – a constantly occurring determination of the validity of currently traveled paths
- Destination recognition – the realization that the destination or goal has been reached or realized

The above steps reiterate the important aspect of a consecutively built-up sense of spatial knowledge and orientation. In fact, navigators use the spatial knowledge and orientation they receive from traveling and they create and update environmental cognitive maps [16] [34]. These maps, sometimes referred to as mental maps or cognitive models, are basically just an internal conceptual visualization of, in the case of navigation, the spatial knowledge of a given VE.

These internal representations are frequently spatially updated throughout navigation activities. There are typically two main ways that this occurs [48]:

- Piloting
    - The navigator uses the location of external VE landmarks, or easily visible areas of interest, to calculate their own absolute position in the VE.
- Path integration

    - The navigator contemplates their own current travel speed and acceleration, relative to their noted starting point, and then infers their relative position in the VE. This method of spatial updating allows for the recursive integration of separate paths and views of the VE in one all-inclusive cognitive map.

The quality and quantity of the spatial knowledge that these cognitive maps provide are highly dependent on the travel method, IT / IM, and the choice for the input element of the 3DUI. Specifically, the decision on whether navigation will be performed from an egocentric or exocentric perspective is highly influential to how spatial knowledge is received and used. Furthermore, the procurement of strong and usable spatial knowledge comes from navigation solutions that require mostly active participation from the navigators. For example, in most cases it is easier for the driver of an automobile to buildup substantial spatial knowledge of the area they have been actively driving in compared to their passive passenger.

Spatial knowledge consists of three main types of information that also correspond to the incremental steps in which spatial knowledge is built up [16] [34]:

- Landmark Knowledge
    - The initial sense of the spatial layout of major areas of interest in the environment. It can be acquired relatively easily and quickly from basic exploration.
- Procedural or Route Knowledge

    - When the environment is thoroughly traveled, the navigator can begin to articulate egocentric based travel paths and routes based on realized connections between the relative positions of landmarks. The presence of this aspect of spatial knowledge allows for the beginnings of more quick and thoughtless navigations.
    - Children typically have more difficulty than adults in trying to associate spatial knowledge between landmarks [27].
- Survey Knowledge
    - This form of spatial knowledge begins to form when the navigator has an adequate amount of existing landmark and route knowledge. Interconnections are made that transforms the previous knowledge from discrete unrelated sub-areas into a view of the entire VE. The navigator can now understand the subtle spatial interrelationships between landmarks and routes.

A fourth step is sometimes referenced called ‘chunking’ [16]. This aspect of the spatial knowledge building experience describes how the navigator can begin to see a large environment as smaller interrelated fragments. In this sense, it is almost a reversal of the transformation of route / landmark knowledge into survey knowledge – but the resulting VE fragments are typically larger and more expansive than the separate routes and landmarks themselves.

3.3 3DUI Navigation Solutions

We define the term 3D navigation solution as the entirety of all design and implementation aspects of systems that afford some degree of 6 DOF of navigational control. We further define these solutions as having the following four main components:

- Interaction Technique / Interaction Metaphor (IT / IM)
    - Design – choreography of the interaction that will solve the given navigation task and that will dictate the requirements of the other solution elements during implementation
    - Implementation – instructions on how to interact with the solution implementation to solve the given task
    - Ex: flying vehicle, eyeball in hand, scene in hand [53]
- Input Interface
    - Design – plans for the physical input object(s) and / or device(s), either porting existing devices or creating unique elements, that will afford the capture of the required navigational input
    - Implementation – the physical construction of the input interface elements
    - Cubic Mouse [15], Joystick, 3D Connexion's SpaceNavigator
- Output Interface
    - Design - plans for the physical output device(s), either porting existing devices or creating unique elements, that will afford the display of the required navigational output
    - Implementation - the physical construction of the output interface elements

    - Ex: LCD Projector, Head-Mounted-Display (HMD), audio speakers
- Virtual or System Controls
    - Design – plans for any non-physical or entirely system-handled aspects of the navigation solution
    - Implementation – the inclusion of desired virtual affordances into the functionality of the underlying system
    - Ex: Affordances for teleportation, constrained or guided locomotion, visual navigational cues

These major elements of typical 3D navigation solutions also suggest a method of classifying or categorizing existing and future solutions. The categories relate to analyzing solutions by recognizing that they are usually centered on one (or more) of the main solution elements. These concentrations will dictate certain commonalities among solutions. Furthermore, we can define one more final category that describes complete 3D navigation solutions that concentrate equally on each of the major elements.

3.3.1 Interaction Technique / Interaction Metaphor Driven Solutions

This category of 3D navigation solutions is centered on the interactive protocol of how the user is supposed to complete their navigation task. Typically, these solutions are regarded as possessing an intangible interactive choreography that dictates how the remaining elements of the solution should fit together.

The main difference between ITs and IMs, in the context of 3D navigation, is the association of the navigation protocols to some existing real-world entity in the form of an explicit simile [53]. As previously stated in the IM section for general 3DUI solutions, the best metaphors for describing ITs are those that match the tasks and system and those that are already known or are easily understood to the user. Applying this principle to 3D navigation, an appropriate IM would be a common or simple real-world entity that is either navigational in nature or else could be easily applied to the task of navigation. It should blatantly inform the user about the navigational affordances and constraints that would exist with this solution. For example, connecting 3D navigation to riding a tour-bus should suggest to the user that, while they may be afforded discrete choices of destination locations, their lower-level navigational behaviors may be constrained.

This introduces an interesting aspect to the design of IMs, namely the spectrum of conceptual distance between the act of navigation and the metaphor that is attempting to be applied. Specifically, there is a choice on whether to use what some may consider conceptually similar navigation metaphors – such as linking 3D navigation to driving a car – or whether it may be advantageous to inventively apply or create a conceptually incongruous metaphor – such as linking 3D navigation to the act of drawing with a pencil. Up to this point, many existing 3D navigation solutions have made use of connecting their tasks to real-world navigational strategies, such as vehicular travel [53] and traveling with the aid of a map [47]. It is useful to note cases when metaphors that are not quite navigational in nature, but are very close, are applied to afford navigation. In these cases, the resulting IM becomes a new hybrid entity which is a combination of the pre-existing definition of the chosen metaphor that is modified to fit the model of 3D navigation. In these cases, the actual metaphor will not exist in reality. For example, our IM of Navigational Puppetry is a hybrid metaphor that allows the real-world entity of puppeteering, which has certain navigational elements but is not exclusively navigational in nature, to be considered in the context of navigation.

We have previously outlined the following issues that should completely inform the design and implementation aspects of the IT / IM elements of an entire 3D navigation solution:

- Must decide between an IT or IM?
    - A decision must be made on whether to describe, design, and implement the interactions of the solution as either:
        - IT: a set of non-cohesive protocols to describe the solutions affordances and constraints
        - IM: an explicit simile between 3D navigation and a real-world entity that cohesively explains the affordances and constraints of the solution, making use of either existing user knowledge of the real-world entity or blatant conceptual connections between the task and metaphor

- Other important factors to consider when considering a metaphor candidate:
    - How closely related to navigation is it?
    - Will it require the creation of a new hybrid definition of itself for the IM?

- How well will it compliment the other solution elements?

Examples

One of the most popular and simplest 3D navigation metaphors is flying. Chuck Blanchard of VPL has been famously quoted making the statement "*... nobody walks in VR (virtual reality) – they all fly ...*" [21]. While flight is not the most natural human connection to navigation, it remains quite an easy solution to understand and implement. Many solutions actually get more specific when they use flight as an IM. For example, architectural walkthrough systems have been known to use a joystick and an interaction metaphor that relates to helicopter flight [21]. Ware et al also take the generic flying metaphor in other interesting directions [53]. They have experimented with their IM of a 'flying vehicle', which involves the manual manipulation of a miniature flying vehicle model that was considered to be in locomotion (like a child playing with a toy airplane). In the same vein, they also investigated a slightly more abstract take on the flying metaphor called the 'eyeball in hand'. This version of the IM involves the manual manipulation of a flying eyeball device, which is essentially quite similar to the 'flying vehicle' metaphor, except there is less of a focus on vehicular locomotion and more on 1-to-1 movements of the 'eyeball' as if it were a camera.

Another interesting non-navigational metaphor that has been experimented with involves the idea of manual viewpoint manipulations. These types of IM are used in 3D navigation solutions that concentrate on the user's manual gestures (often captured by data gloves or sensors places on the hands) to complete navigation tasks. For example, there have been studies performed where the user ‘grabs the air’ and pulls themselves forward, as if they are standing on a skateboard and are pulling themselves by some rope that is rigidly attached somewhere off in the distance [4] [5].

It is possible that more realistic metaphors, like map aided car navigation, can produce affordances that result in quite an unnatural navigation experience. For example, an interactive map of VE can be presented to the user and they can ‘teleport’ themselves to any discrete location instantaneously [3] [47]. Alternatively, the user could simply choose discrete locations on the map as milestones of a path or route that they wish to follow.

3.3.1 Input Driven Solutions

This category of 3D navigation solutions is centered on the actual input element(s) of the 3D navigation solution. The most distinct aspect of these solution types is their focus on the user generated side of HCI communication and their detachment from any system output responses. As previously stated, the input elements of general 3D interaction solutions must capture enough input from the user to complete the given task. In the case of 3D navigation, the input element must capture the user’s input that, when combined with the IT / IM and other solution elements, provides the system with an adequate amount of the 6 DOF of navigational control data.

Choosing an appropriate existing input device, or designing and implementing a brand new one, can be a very difficult task to do successfully. As previously stated, careful attention must be made to the physical shape of the input element because that dictates the affordances and constraints that are provided. Equally as important to the outward appearance of the input devices are its inner workings - such as sensor types and transfer functions.

There is a great amount of diversity among possible tasks that require navigation for completion, luckily there is a correspondingly broad spectrum of input affordances and constraint combinations that can be arranged. For example, it has been shown that user's tend to perform rotation and translation actions as separate subsets during the navigationally relevant task of docking [29]. For these types of navigational tasks, solutions that have input elements that separate these actions, either physically through the device itself or logically through how the device is used, should prove to be better than those that integrate those activities. Information gathering potential, which is one of the major performance metrics for 3D navigation solutions outlined by Bowman [3], could be greatly enhanced by allowing the user to move and look around separately. Consider the real life activity of walking through a hallway that has messages written on the walls. The messages can be read by a human being while constantly moving to the end of the hallway, since our heads, flexible necks, and legs allow for a separation of translation and orientation control. However, a bumble-bee whose method of flight does not allow for this separation must keep flying toward each message on the wall separately in order to read all the messages.

Considering the lower-level detail of sensor type exclusively, some may be better suited for navigation. For example, 3D navigation solutions that are meant to mimic vehicular transportation, like a motorcycle, would benefit greatly from input elements that use elastic sensors since they provide a counterforce in the same way as a motorcycle's grip-based acceleration control. However, elastic sensors would be less appropriate for navigation solutions that were meant to be closer to moving a game-piece around a board game. In that case, it would be best to use isotonic sensors since they measure actual displacement and orientation and provide no counter-force.

Entire 3D navigation solutions that are concentrated around the input element(s) must be carefully designed and implemented around the exact specifications of the input interface or device. In these solutions, the choice for input element(s) usually precedes all others. For example, selecting a dual thumb-stick controller (such as the controllers for Xbox360) as the input element for a FPS game dictates that a very specific IT / IM will have to be used in conjunction. Viewpoint translations and orientations are typically mapped to each thumb-stick, respectively. However, this would not be the case if the input device was changed to a generic isometric joystick. There is usually not a sufficient amount of DOF to be controlled at one time so there would be a requirement in the IT / IM to switch states between mapping the joystick to control the translation and when it would control the orientation. An example of the choice for the input element dictating the choice of output is using gaze-directed steering to capture navigational input. In this case, the most appropriate output element to use would be an HMD. The user will be changing their head's orientation frequently to steer their virtual movement in the VE, so it would be very useful to be able to dynamically change the position and orientation of the user's view of the VE (via the output element) to compliment their changing gazes.

Examples

One of the most popular car racing arcade games is 'Cruis'n USA' by Midway Games. At first glance, the most noticeable feature of this arcade is the quasi-realistic car interior format of the entire arcade system – which is mostly made up of the input interface. This is perhaps one of the best examples of an entirely input element driven navigation solution since every aspect of this solution is centered on this car-like interface. Firstly, the car interior interface is designed to very closely mimic a real car and it dictates that only an IT / IM that involves, with some degree of fidelity, the real-world act of driving a car would be appropriate. Secondly, it dictates that the accompanying output element of the interface must also conform to the existing car interior style, and this is done by placing an output display where the windshield would be located. Other output elements can be further complimented to the car interface by the addition of audio speakers for sound effects and also by enabling the actual chair of the interface to provide haptic feedback.

3.3.2 Output Driven Solutions

This category of 3D navigation solutions centers on the output element(s) and is distinguished uniquely from the rest since their sole concern is with communications made to the user from the system. In this category, we refer only to the physical and technological elements involved with displaying output to the user and not on 'what is being displayed'. Immediately, this suggests that these solutions are more focused with exploring or enhancing the less tangible experiential aspects of navigation. Unlike their input-centered counterparts that focus more on satisfying technical input data requirements, these output-devoted solutions focus on things that enhance the perceived incoming sensory experience of the user – visually, aurally, haptically, and occasionally olfactively. While input elements may seek to explicitly afford new or better interactions, output elements typically seek to better compliment the existing interactive affordances. For example, designing a solution around wearing a HMD would not necessarily enable any new interactions to take place, but it would greatly compliment and improve the interactive experience of certain solutions that perhaps require complete occlusion of the real-world or extremely realistic visual depth cues.

This category may overlap in some cases with the input element, if the device itself is bi-directional in terms of HCI communication. For example, Xbox 360 game controllers are capable of rumbling and vibrating to provide haptic feedback during game play. This subtle form of output can add a tactile dimension that deepens the user's interactive experience, which may afford certain overall interactive benefits due to increases in the user's sense of presence and engagement.

The most basic aspect of output centered solutions is the choice for how the system visually communicates with the user. In the simple case of traditional desktop PC interaction, the display screen serves this purpose. With the act of 3D navigation, this display is the user's view or window into the VE in which they are interacting. Changes to how the visual display is achieved will have the most impact on the following major navigation performance metrics [3]:

- Presence / immersion: the user's sense of being actually inside the environment

- spatial awareness / orientation: user’s knowledge of their position and orientation in the environment
- information gathering ability: the degree with which the user can obtain and use information during navigation

The number of possible modifications that can be made when focusing on visual output exclusively is quite large. To improve or change the above mentioned performance metrics, modifications are usually made that affect the user’s sense of immersion and their field of view [3] [48]. This process will start with making the choice for the actual physical technology that will facilitate the visual output. These choices can be summarized into three main categories that are distinguished based on the degree to which they immerse the user [3]:

- Non-Immersive
    - These displays does not provide any increased sense of immersion or presence compared to the traditional HCI paradigm of desktop PCs
    - Ex: LCD display screens, laptop screens, television set
- Semi-Immersive
    - These displays aim to provide an increase in immersion and presence for the user, but in a fashion that does not fully occlude the real-world or the rest of the interface
    - Ex: stereo projectors / glasses, 180 degree surround screens, etc.
- Fully-Immersive

- These displays seeks to immerse the user as much as possible into the environment by occluding the real world and the rest of the interface
- Ex: head mounted displays (HMD), virtual retinal displays, etc.

The first two categories can be further subdivided by examining their actual display sizes and the field of views they afford the user. Many studies [3] [7] [17] [34] [48] have indicated that the magnitude of the available field of view has an important impact on performance in many 3D interaction tasks. As previously stated, the quality of the user's spatial awareness and orientation is highly dependent on the choice for the visual output display. There have been many studies performed that directly link larger display sizes and fields of view to increases in spatial awareness [34] and 3D navigation performance in general [48].

The other sensory aspects of output elements, namely haptic and auditory feedback, also can directly affect the user's sense of presence and even their sense of spatial orientation. 3D audio cues can help the user orient their position in sound-enabled VEs through sound localization and it can also allow for a certain degree of collaboration, since audio can be 'broadcast' to multiple users regardless of their immediate proximity to other interface elements. Haptic feedback can also heighten the user's sense of presence within a VE, especially in solutions that have little tactile intimacy in their input element(s). For example, a 3D navigation solution that utilizes gaze directed steering does not provide the user with any tactile sensation to indicate the beginning, continuation, or end of navigation actions. However, if the user was equipped with a necklace or belt that vibrated in a certain fashion during navigation, then the user is given very useful additional sensory information that will make their entire body feel more present within the VE. In this sense, audio and haptic outputs could be interchangeable, in some cases, since they are each intending to provide an additional mode of sensory information.

3D Navigation solutions that are centered on the output element must be designed differently from all other types. It must almost be done in reverse compared to those that focus on the input element or the IT / IM. In these cases, the problem statement would essentially dictate a desired perceptual / sensory experience. Depending on the exact task specification, either the remaining elements of the solution must be created from scratch around this goal or the output must be designed to afford the desired experience and compliment pre-existing choices for the other elements. For example, the previously mentioned amusement park attraction 'Star Tours' could have conceivably begun as a problem that required an output centered navigation solution. The entire goal and concentration of the project is about giving the riders of this attraction a very specific sensory experience. Obviously, these riders are like passive passengers and do not provide the system with any navigational input. However, the design of this attraction required very careful decisions to be made on how to arrange the various visual and sound displays that must function in a synchronized fashion with the haptically enhanced rider seats.

Perhaps the most important challenge to output centered navigation solutions is the degree to which they can provide an adequate amount of 3D perceptual cues to the user.

The most integral cue for navigation is arguably depth, so the problem becomes how to maximize the degree to which depth can be perceived by the user that also compliments the other elements of the interface. For example, the most realistic method of displaying a 3D scene to a user is for them to wear a 100% reality-occluding 3D HMD. While this may provide the user with a very strong sense of depth, their interactive experience may suffer in other ways since they now would not be able to see their input interface. The same principle applies also to the more general aspect of increasing the display size and / or field of view. The benefits from these changes to the output must be done in ways that compliment the other solution elements.

Examples

An interesting example of an output focused 3D interaction solution that enables a very unique form of audio navigation, is the Pendaphonics project [22]. This interface provides a very unique and physical way of experiencing a virtual 3D soundscape, which is linked to a physical interaction space in reality. The system consists of a large vertical display, providing the view of the soundscape to the user(s), audio speakers and one or more suspended pendulums. As the pendulum is swung by one or more users through the physical interaction space, the sounds it generates changes according to its relative virtual path through the soundscape. The nature of the interaction style - the swinging of a suspended tangible object - affords very interesting user experiences and explorations with sound control.

Another interesting example relates to an entire grouping of solutions that conform to a certain output display protocol called CAVE - Cave Automatic Virtual Environment. First developed by Electronic Visualization Laboratory at University of Illinois at Chicago and presented at the 1992 SIGGRAPH, it has since been a popular output protocol. The basic principle of CAVE is the placement of the user's interaction space in the center of a room or cube whose bounding surfaces (walls, ceiling, and floor) are projected displays. Obviously, a strict output protocol such as this dictates certain things about what remaining interface elements work and don't work. LaViola et al. developed STEP-WIM [28], which is a very interesting interface for 3D navigation, which uses an HMD but was also designed around the output principle of CAVE. Their system allows the user to 'step' through the VE by tracking their feet, through sensed slippers, on top of the floor, which displays an overview of the VE. They also see other horizontal views of the VE on the remaining CAVE walls.

3.3.4 Virtual / System Driven Solutions

This category of 3D navigation solutions describes all system related and non-physical aspects of the interface. Anything purely virtual that can be perceived as contributing to the overall navigation falls into this category. Virtual centered solutions are similar to those focused on the IT / IM element, and they often overlap with each other since they are focused with 'how' the action of navigation is being committed.

The most common type of navigation related system / virtual affordances are visual cues or aids. For example, many navigation solutions use virtual visual aids that are based on real-world wayfinding tools or strategies [47], such as maps, compasses, grids, horizon indicators, directional arrows, or destination demarcations. The inclusion of multiple views of the VE could also be considered as a form of virtual aid. In the same sense, audio cues also fall under this category. It is important to note that the ability or choice to display audio is an output related aspect, but the protocols that govern how the audio is used is purely a virtual element.

The scope of system / virtual elements of navigation solutions may go beyond sensory augmentations into the realm of system-controlled navigation behavior. For example, solutions have been investigated that incorporate either the guidance or constraint of their navigation travel [49]. It can be very useful in some cases to reclaim control over certain simultaneous DOF from the user, such as restricting travel to being planar.

The virtual elements may also go further in their scope to encompass the entire activity of navigation – a virtual counterpart and manifestation of the IT / IM. In these cases, there is a direct overlap with the IT / IM element since it is being entirely afforded by the system. For example, linking 3D navigation to teleportation is an interesting IT / IM that is entirely encompassed and afforded by the virtual aspect of the solution.

The first decision that must be made when addressing the issue of virtual / system related navigational element is the determination of its scope. As previously stated, this element may only be augmenting the existing IT / IM or they may be intended to entirely afford the act of navigation. Solutions that focus on using virtual navigational aspects in a supportive role must primarily concern themselves with reducing as much of the cognitive effort required for the navigation interaction as possible, in order to allow remaining mental resources to be used for whatever task is being supported by the navigation.

If the system is completely affording the IT / IM associated with the solution, then these elements should be designed and arranged together. Care must then be taken to select or design the input element based on the lower-level details of the virtual / system implementation. For example, consider a solution that will accomplish navigation through teleportation and another solution that is intended to be closer to driving an automobile. The input requirements will be discrete for teleportation, rather than continuous for driving. Another important concern for these completely virtual navigation protocols is that they most often will lack any kind of non-visual locomotive feedback. Care must be taken in these cases to select input and / or output elements that may be able to counteract this by providing some sense of motion or haptic indications of locomotive starting and stopping.

Examples

There are many examples of research that investigate the use of purely virtual / system controlled visual and auditory aids to support navigation. For instance, Haik et al [20] used simple arrows to guide the user through navigation tasks to get to certain locations. Other solutions have even tried warping the user's view of the VE to overcome occlusion or perspective problems [29]. It is also a common practice, in certain cases, to visually augment the VE terrain with gridlines and directional cues to aid in navigation activities.

There have also been many attempts to investigate the usefulness of varying amounts of system control or guidance over user navigation. For example, Tan et al. [49] experimented with controlling the orientation parameters of the viewpoint with the system based on the travel speed being controlled by the user. Specifically, this allowed for an easy transition between flying for global exploration and for localized visual inspections of sub-regions of interest. They also experimented with providing the user with an auto-pilot orbiting feature to allow for easy object inspections at constant viewing orientations.

Other navigation solutions that relate to varying levels of virtual control can prompt for new ideas on how to define navigational IT / IMs. For example, non-traditional or unrealistic methods of travel have highlighted ways to reconsider how the act of navigation could be viewed by both the user and the system. Examples of this would include discrete target selection from a map – such as the World-in-Miniature (WIM) [47] style of navigation. The basic principle of this approach is that the user is presented with an interactive map view of the VE in addition to their regular view. The user is able to discreetly select spots on the map, and the system automatically moves them there (which could be accomplished either through teleportation or entirely system controlled locomotion).

3.3.5 Entire Interface Driven Solutions

The final category of 3D navigation solution types refers to those that equally concern themselves with all of the interface elements – IT / IM, input, output, and virtual / system – in an effort to design and implement an entire navigation experience. In this case, there would be equal importance placed on the how the navigation would be completed conceptually, how the input data could be collected (including any device specifics), and how the desired sensory experience would be communicated by the system

The primary challenge for entire 3D navigation solutions is the successful reconciliation and complementation of each separate element. When successfully designed and implemented, these solutions become an interactive experience that exists as a separate entity that is greater than the sum of its parts. The intent should be that the actual solution interface elements are never perceived as being more prominent than any others or as being separate from the entire interactive experience, which is obviously never completely possible. However, it is possible to arrange the elements such that each one is peripherally perceived quasi-simultaneously as related portions of a entire cohesive interaction experience.

Examples

We will now use the widely known Nintendo Wii System as an example of a solution that that is equally concerned with all of the interface elements and affords a distinct entire navigation experience. It is possible to concentrate only on the input issues surrounding their controllers and the gestural nature of their ITs / IMs. However, in this instance, we choose to focus on the higher-level aspects of the entire Wii interaction experience, which successfully reconciles gestural IT / IMs, quasi-3D spatial input with unique devices, system-aided functionality, and a synchronized graphical display with audio. When considered as a whole, the interactions with the Wii become a distinctly observable experience that is made possible by the integration of all these separate elements. More examples of entire interface solutions for 3D navigation will be discussed in section 5.2.

3.4 3D Navigation Solution Challenges

3.4.1 General Challenges

In order to design and implement successful 3D navigation solutions, one must first understand the major navigational challenges and then use them as inspirational guides throughout the process. We re-defined the basic definition of a 3D navigation solution to be an interface that includes four major elements - IT / IM, input, output, virtual / system – that together afford the user control of the viewpoint's (or the user's virtual eyes') position and orientation within some form of VE. The main elements of the task of 3D navigation are control over potentially multiple and simultaneous DOFs and the building and maintenance of conceptual references to relative spatial positions and orientations in support of some navigation related task.

Bowman's landmark 3D interaction experiments in travel techniques used the following broad performance metrics, which further highlight the most important high-level areas of the act of navigation [3]:

- Speed

- Accuracy (in terms of steering towards targets, etc.)
- Spatial Awareness
- Ease of Learning
- Ease of Use
- Information Gathering Potential
- Presence

Three of the most common navigation challenges that are of direct concern to our work to be efficient control over multiple DOF of navigational control, maintaining spatial awareness or orientation, and realizing adequate 3D visual perception [3] [5]. For example, care must be taken when specifying the number, integrations, and separations of the DOF of navigation control [31]. There must be a balance between the travel freedoms given by high amounts of DOF control and with the ease-of-use of more restrictive methods. Similarly, the entire solution must be geared towards providing the user with the highest possible potential for acquiring and maintaining strong spatial orientation. In general, people tend to possess a wide range of spatial abilities [5]. There could be a drastic difference between how two users' orientations are affected by the inclusion or exclusion of even 1 DOF of control. Finally, special design attention must be made to provide the user with as many 3D visual aids and perceptual cues as possible to ensure they feel conceptually connected and present in the VE

3.4.1 Virtual Navigation vs. Real-World Navigation

One of the most distinguishing elements of 3D navigation, that also highlights a significant challenge, is the ‘virtual’ aspect. 3D navigation is a virtual expression of the real-world activity of navigation. Drawing comparisons and parallels is unavoidable since virtual navigation was directly inspired by the real-world and one cannot contemplate them as mutually exclusive topics.

Being virtual clearly dictates that the navigation is taking place in a separate reality. The user is not present in this reality, nor is the navigation interface. For all intents and purposes, the ‘virtual – reality’ can be considered to be another dimension that can only be accessed remotely through input interface communications and some form of output sensory feedback. It is a commonly held notion that one of the major drawbacks of VEs is the inherent lack of haptic feedback [31]. This is because the user can never actually touch anything within that environment. Everything must be done by proxy through the affordances granted by the particular interface solution in question. It is precisely this point which highlights the importance of trying to increase the user’s sense of presence during navigation. Solution designers need to try to give the user a sense that they are actually in this other dimension.

In addition to a sense of presence being communicated across the gulf between dimensions, there must also be an attempt to maximize the amount of interactive fidelity or precision that survives the translation to the virtual world [31]. This is where the spatial input and 3D perceptual cues come in to the picture, in that they should be attempting to closely mimic the subtle natural interactive affordances that users exploit in the real world.

Much of the knowledge that was used in the earlier years of 3D navigation research came from studies of real-world human and vehicular navigation. Specifically, literature on geographic knowledge as it cognitively applies to aviation and locomotion is extensive and highly appropriate. For example, studies on how people use maps for orientation in the real-world has direct implications for how designers should go about including map-like elements to their 3D navigation solutions [37] [46]. Part of the known protocol for real-world navigation with the aid of a paper map is that people must translate and re-interpret the map's geocentric spatial information to compliment or align with their current environmental orientation [46]. Many 3D navigation solutions exploit this fact to provide the user with an increased sense of spatial orientation, and more potential to get a stronger awareness at a glance, by dynamically modifying the virtual map's orientation to constantly align with the user's movements. For example, the FPS Xbox game HALO provides the user with a dynamically oriented motion sensing map, which is always centered on the user's current position and synchronized to their vertical axis of orientation.

Another important distinction between real-world and virtual navigation is the breakdown of navigation into travel and wayfinding. While conceptually it is easy to see how these activities are related but distinct, they are usually not perceived that way during the act of real-world navigation. The same can be said of certain 3D navigation solutions to some degree, but it is even easier to see how people usually gather environmental information, plan their movements, and execute them all simultaneously in the real-world. This principle is directly linked to potentials for spatial orientation, which tends to degrade as the actual method of travel becomes increasingly unnatural [3]. For example, travel and wayfinding are essentially the same activity when people walk, run or ride their bicycles, but they become increasingly separate and distinct when people begin driving cars or flying planes, since these acts are more artificial and complex than walking. Therefore, care must be taken to provide a virtual navigation experience that strives to simultaneously enable good spatial orientation and interactively link travel and wayfinding.

4.0 A New Frontier for 3D Navigation Solutions

We believe that the current state of research in the field of 3D navigation can be expressed as an extension of the experiential plateau being observed throughout all 3D interactions and UIs. We believe that at least part of the problem is due to an inability to adequately visualize the true potentials of existing interface styles, such as TUIs, and technological elements, such as optical tracking. We suggest that in order to yield truly useful and innovative knowledge and progress, effort must be made to exploit these ideas to the fullest extent possible.

Specifically, recent advances in input / output technologies, sensors, and even the re-invention of existing technologies for new purposes form a very extensive and diverse ‘toolbox’ for solutions, which should allow for more imagination in designs and interface implementations. For example, optical tracking can be easily accomplished with inexpensive webcams and simple paper markers. This is directly relevant to 3D interfaces since optical tracking can provide a simple method for continuous 6 DOF tracking of an object or even part of the user’s body. In the same vein, the increasing popularity, and constantly improving technology, of projectors and touch sensitive interactive surfaces creates new possibilities for both input and output solutions.

In addition to being able to envision the physical interface possibilities of the current state of relevant technologies, one must also be able to imagine the intangible IT / IMs that will be made possible by the application of these technologies. Although 3D interaction has been around for many years, it is still a relatively undiscovered design territory.

While many specific domains and activities have been made possible or augmented through 3D interaction - gaming, movie special effects, architectural modeling, etc. – many basic HCI activities, like web surfing, remain within the realm of 2D interactions. In general, people are just beginning to see the potential of the experiences that can be realized with 3D interaction. The recent emergence and popularity of quasi-3D gaming interfaces, like Nintendo Wii and Playstation Move, are evidence of this. Considering 3D navigation solution designers specifically, they need to begin to look at the design space of 3D navigation IT / IMs as being a 3D canvas or empty theatre stage that can be quite inventively filled and choreographed. In this sense, it should be obvious that designers must possess a form of 'workshop' that allows them to rapidly design and build prototypes, and then test them in ways that bolster their own personal creativity and do not require excessive external support.

4.1 Tangible User Interfaces

Tangible User Interfaces (TUIs) are a relatively new area of interface research that seek to take advantage of the inherent haptic and spatial interaction skills that humans possess when manipulating physical objects or tools in reality. They afford an interaction experience which connects computational cyberspace and the physical environment. Digital information is essentially given physical form by coupling computational control and information with graspable physical objects [8] [24] [27] [42] [44] [51]. The resulting interface will typically contain one or more physical objects – alternatively called tangibles, props, or graspable devices - which will act both as tangible representations and controls within the interaction. In regard to what they represent and control, the tangibles can either be generic (like a puck [9] or dial [35]) or highly specialized (like a toy action-figure [16]). They provide input to the underlying system through optical tracking, wireless sensors, or even through electromagnetism [35]. In addition to the actual tangible objects, TUIs also usually contain one or more intangible output displays such as a large projected screen or interactive surface.

TUIs can also be described by how they differ from traditional Graphical User Interfaces (GUIs). A traditional GUI – with mouse, keyboard, and typical display screen – contains physical objects for control that are technically tangible. However, their position and state do not represent anything about the underlying system or application state. This results in a clear distinction between the input space and devices and the output space and devices. TUIs attempt to eliminate that boundary by coupling control and representation to physical objects. The boundary is further blurred when intangible output representations, like projected display surfaces, are used in conjunction with the tangible objects to display dynamic graphics directly on or around the object. Furthermore, the TUI philosophy has a profound impact on redefining the boundaries between the user and computer due to its blurring of input and output spaces. A major consequence of this is a greater affordance for collaborative interactions.

TUI research encourages the exploitation of people’s innate spatial interaction skills and abilities, which can be most obviously observed through frequent affordances for bimanual interaction. The fact that most TUI systems involve space-multiplexed tangible objects as controllers is evidence of the many opportunities for two-handed interactions.

Allowing for two-handed interaction means a wider bandwidth of the human-computer communication channel through increased parallel input from the user. It also leads to a richer interactive experience for the user. For example, a TUI system that affords tangible manipulation of two parameters of a simulation simultaneously, using both hands, has many interactive and computational benefits over a system that only permits single parameter interactions.

A very good example which highlights the main principles of TUIs is the SenseTable from MIT's Tangible Media Lab [35]. The SenseTable uses electromagnetism to wirelessly track multiple generic tangible pucks, which also have modifiable dials on top of them. The pucks can be bimanually manipulated on a horizontal display surface which is being projected down upon from overhead. In addition, a small vertical display screen is placed at the upper left edge of the horizontal surface. Graphical representations of an underlying simulation are projected down onto the horizontal surface and extra summary information is displayed on the vertical screen. The pucks become bound (computationally and graphically linked) to a particular underlying graphical entity when placed within a given proximity, and can now be modified manually by interacting with the pucks. The projected graphical information is now directly on top of the bound puck and is dynamically updated as the puck or dial is manipulated. Pucks can interact with each other when in close proximity or can be used to interact and bind with additional information from the vertical screen. The SenseTable was used for modeling system dynamics – the displayed graphical entities on the horizontal surface corresponded to parameters in an underlying model and the information being displayed on the vertical screen is additional information relating to the underlying system model.

The design of a TUI is different from most other interfaces due to the freedom its own definition and philosophy provides. Basically any object(s) can be part of the system - whether on its own or within some interactive surface - and can be mapped to control / represent any number of computational elements. The object itself can be anything from a pre-existing coffee mug, to a specifically designed LEGO structure, or even to a doll representing a miniature real version of a virtual character [30]. The overall structure of the TUI can also be liberally designed in terms of the relationship between output displays and tangible objects. However, before the physical design of the TUI can take place, an appropriate interaction metaphor must be designed (or an existing one changed) specifically for the TUI. One of the leading philosophies behind TUIs – 'making bits tangible' [24] - implies that the interaction metaphors being 'tangibly' implemented can be approached with a similar sense of freedom in their design phase. 'Bits' previously considered inappropriate or too difficult to map for control in more traditional interfaces can now be considered for potential elements of the tangible interface.

An interesting work that highlights the extent to which extremely imaginative IT / IMs can be developed for TUIs is the Pendaphonics project [22]. This interface provides a uniquely physical way of experiencing a virtual 3D soundscape, which is linked to a physical interaction space in reality. The system consists of a large vertical display, providing the view of the soundscape to the user(s), and one or more suspended pendulums. As the pendulum is swung by one or more users through the physical interaction space, the sounds it generates changes according to its relative virtual path through the soundscape. The nature of the interaction style, the swinging of a suspended tangible object, affords very interesting experiences with sound control which are not possible without such an interface.

Progress within the field of TUIs has afforded certain specific improvements to HCI in both the tactile and auditory sensory realms. In fact, there has been a vast amount of research regarding TUIs and sound [6] [22] [36]. A notable contribution to the field of TUIs and sound is the reacTable [26] - an interactive tabletop interface for synthesizer control. Generic tangible handles are manipulated on a graphical surface to modify the sound parameters of a synthesizer. Its interaction experience explores the tactile benefits of tangible sound control and also highlights the performative bonus of observing the synchronization of the visual and graspable aspects of audiovisual TUIs.

4.2 Tangible User Interfaces for 3D Navigation

We propose that TUIs are the best of all possible candidates to be 3D navigation solution ‘design and prototyping workshops’ that allow for quick experimental designs, rapid prototype implementations, affordances for natural interaction and creative output elements, and testing in ways that support and encourage innovation and that also do not require excessive external support. More specifically, we argue that they address many of the previously stated major challenges that face 3D navigation solution designers.

The common issue of maintaining efficient control over multiple DOFs simultaneously can be creatively addressed through the tactile and haptic centered IT / IMs and graspable props inherent to the TUI philosophy. In addition to the previously mentioned aspects of space-multiplexing and bimanual interaction, the basic interactive model of TUIs involves a double interaction loop of passive haptic feedback with every graspable element of the interface. Since the input elements themselves are tangible objects that control and visually represent things, the user is constantly receiving passive tactile information from the grasped devices in addition to the regular digital or visual feedback from the system [24]. We believe that the coincidence of action and perception spaces should create a very engaging navigational interface. Furthermore, the combination of the tactile richness of TUI props and their support for imaginative choreographing of IT / IM results in tremendous opportunities for experimenting with inventive methods for capturing navigational input from users.

The other common issue of spatial orientation can also be addressed through the philosophy of TUIs. Regarding output exclusively, the arrangement of graphical displays and their relation to the actual method of navigation can be liberally arranged to display multiple VE views, maps, and other different visual orientation cues. Furthermore, it is possible to have the main navigational input prop function as both an input controller and also as some kind of visual and tactile indication of the location and orientation of the viewpoint, which should significantly affect the user's sense of spatial awareness. Studies that investigated interfaces with physically graspable objects for robotic control [18], which is a task that requires high amounts of spatial and situation awareness, found that users could get extra spatial information from the orientation and position of physical objects when they are manipulated relative to their surroundings. In the case of navigation, this would allow for the users to get new higher-level forms of navigational information from observing their navigational behaviour during the interaction, which we refer to as meta-navigational information. For example, consider the extra information that could be perceived through parallel parking a car from a removed perspective through the manipulation of a tangible miniature model, versus doing the same from within the car. These meta-navigational pieces of information benefit the act of navigation by giving new data and action-choices to consider during travel. Meta-navigation combines navigating with the 'observation of navigation' and brings into reality aspects of the act that are usually purely conceptual. In addition to alleviating a portion of the cognitive effort required from the user, allowing them to see and grasp their position within the VE will provide an enhanced navigational experience regarding spatial orientation and situation awareness.

The meta-navigational approach to 3D navigation that is made possible by TUIs has a direct benefit to the important challenge of adequately enabling 3D perception. The actual graspable devices themselves will be 3D tangible objects and they can be used within IT / IMs that exploit spatially oriented interactions and gestures. If the user is manually interacting in the same number of dimensions that their perceived navigation behaviour is occurring, then the user will be required to perform less mental gymnastics to reconcile their action and perception spaces. Furthermore, the combination of vertical and horizontal display surfaces with appropriate viewing perspectives can simulate a 3D interaction space for the user even if the displays themselves are 2D. For example, the BUILD-IT system [10] used multiple generic tangible bricks to enable both egocentric and exocentric virtual camera and scene control within a VE. The interface contained both a first person view on a vertical screen and a third person overview from above projected down onto a horizontal interactive surface. This project demonstrates the perceptual benefits of multiple perspectives that act as 2D shadow-like projections of the boundaries of a virtual 3D space.

TUIs are also generally regarded as easy to use interfaces for beginners, mainly due to the naturalness, intuitiveness, and directness of the interaction style. In fact Guzman et al. state explicitly that "*Tangible interfaces have the potential to support learning for non-expert users, ease 3D navigation, and foster collaboration*" [19]. The previously articulated difficulties of efficiently controlling multiple DOFs and maintaining spatial orientation are especially prevalent with novice navigators. Therefore, it appears that a TUI for 3D navigation should cater better to beginners than other interface styles, such as a flight simulator cockpit. At the same time, if a 3D navigation TUI supports beginners then they should also be able to also support complex or irregular navigational input and behaviors. In addition, since TUIs were previously stated to afford collaborative interactions then it follows that they would also support collaborative navigation. As is suggested by the current popularity of global positioning system (GPS) car-navigation tools, TUIs that can afford collaborative 3D navigation could really enable a new way of looking at navigation that is no longer based on the individual navigator archetype.

In general, using TUIs, or their quasi-tangible direct-touch / multi-touch interface counterparts, to investigate 3D navigation is an area producing interesting preliminary results, but which could benefit from further investigation. The Follow-my-Finger [1] 3D camera control solution utilizes two displays, a vertical first person perspective view and a horizontal touch surface depicting the VE as seen from above. Using direct touch on the horizontal surface, in conjunction with a few virtual widgets, basic camera control can be achieved on the horizontal surface. Another interesting tangible navigation experiment was the ProBono System [16], which was designed for elementary school children to learn navigational knowledge. We will discuss these and other TUI 3D navigation solutions in section 5.2.

We argue that applying the principles of TUIs to the task of 3D navigation should be extremely beneficial for certain tasks that require 3D navigation as a means to achieve some larger goal and those that do not require a conceptually strong link between the user's virtual position and how that relates to their position of interaction in reality. For example, determining a virtual camera's path through a 3D Architectural Model Viewer would be a perfect example of a task that does not require such a strong link. On the other hand, a high fidelity flight simulator which is meant to mimic flying a real plane does require a strong association of those two things. Another area that would be very appropriate would be high-level Human Robot Interactions (HRI) and the piloting of Unmanned Vehicles, both unmanned Ground Vehicles (UGV) and Unmanned Air Vehicles (UAV).

5.0 Navigational Puppetry

5.1 The Navigational Puppetry Metaphor

The Navigational Puppetry metaphor implies a distinctly new way of looking at 3D navigation by redefining it as the puppeteering of a graspable navigational avatar, within a multiple view perspective, that allows users to essentially reach within the VE and manipulate the viewpoint directly. This metaphor stimulates a unique superposition of egocentric and exocentric navigation perspectives and a blending of navigational action and perception spaces. It affords an interaction experience that can best be described as navigational puppeteering – a form of meta-navigational interaction.

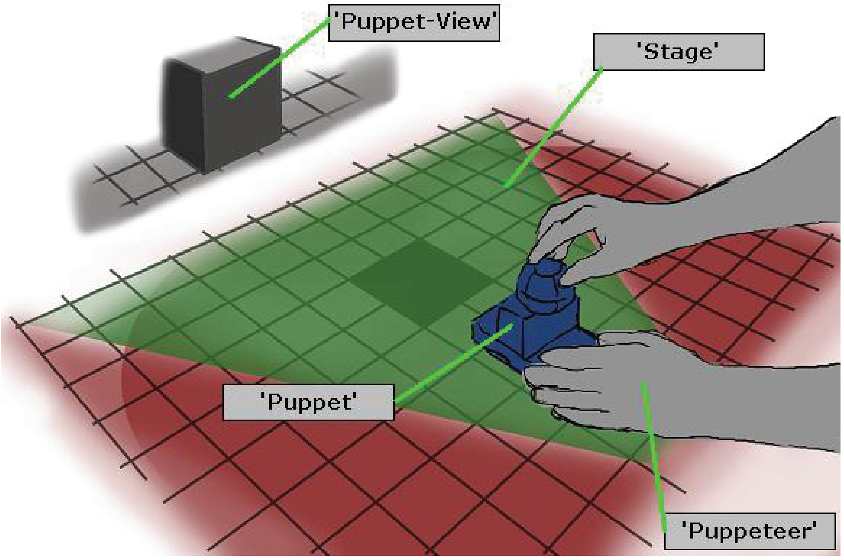

Figure 1 – The Navigational Puppetry Metaphor

The metaphor has four main elements: the 'puppeteer', 'puppet', 'stage', and the 'puppet-view' (see Figure 1). The metaphor positions the user as a 'puppeteer' standing over a 'stage' containing a 'puppet', whose field of vision is displayed through the ‘puppet-view’. It can be conceptually scaled in size from as small as a laptop computer to the size of a war room or movie theatre. The puppet is a tangible prop that affords complete 6 DOF translation and orientation control in a structure that resembles a simplified miniature humanoid – a translatable base / body and a movable head / eye for orientation. The stage is a bird's eye-view of the VE, represented as a graphical interactive surface on which the puppet is manipulated. The traditional notion of the puppeteer is further extended by giving the user a first person perspective of what their puppet sees on the stage. Also, to support the blending of exocentric and egocentric navigation, the puppeteer can also perform translations and rotations of the stage.

The act of puppeteering is something that most people have a general familiarity with - the natural manipulation of a miniature entity in an imagined surrounding environment, relying on natural spatial skills and naïve physics. Complex and / or irregular navigational behaviors would be very easy to capture since the puppet manipulations are so natural and intuitive that they do not need to be significantly translated before they produce the desired behavior.

The metaphor allows for new areas of the 3D navigational experience to be explored, which we previously defined as meta-navigational affordances. It combines navigation with the 'observation of navigation' and brings into reality aspects of the interaction that are usually purely conceptual. For example, the previously mentioned act of parallel parking an automobile can be a difficult task to accomplish quickly on a busy city street.

Consider the changes to that activity, both to the driver and anyone observing it, if they could perform it from a navigational puppetry perspective. All relative distances and turning angles would no longer be required to be conceptually calculated dynamically by the driver; they could be automatically sensed and performed naturally with the same innate spatial skills that humans exploit when manipulating everyday objects. Furthermore, since the automobile is now being ‘parked’ by puppeteering methods, it ceases to be an entirely black-box type of task and can portray many subtle navigational specifics to any observers.

The navigational puppetry metaphor was designed to possess interactive benefits that apply directly to the major 3D navigation challenges. The manipulation of a tangible ‘puppet’ would result in greater tactile intimacy due to the double interaction loop that occurs in TUIs [24] – the ‘puppeteer’ is receiving immediate passive haptic feedback from the ‘puppet’ and also visual feedback from the position of the ‘puppet’ on the ‘stage’. Allowing the user to see and grasp their position in such a way will also provide an enhanced navigational experience regarding spatial awareness, which is partially supported and inspired by other investigations into 3D interactions within multiple perspective display surfaces [1] [10] [16]. Complex and irregular navigational behaviors would also be very easy to capture since the ‘puppet’ manipulations are so direct and intuitive. For example, let us assume that a 3D animator is looking to achieve a specific virtual camera path that requires a particular trajectory with an unconventional pattern of alternating speeds and perspectives. This can be easily captured and perfected with navigational puppetry using the same amount of effort as it would take the animator’s boss to demonstrate the desired behavior with a pen over his desk. Navigational puppetry also affords collaboration, both interactively and in an observational sense.

We argue that the most applicable 3D navigation tasks for the puppetry metaphor would be those that require navigation as a means to achieve a larger goal, as opposed to the actual navigation experience being the focus.

5.2 Identifying the Puppetry Trend

We position our navigational puppetry metaphor as a distinct articulation of the front wave of an entire trend in recent 3D navigation solutions. This trend involves solutions that increasingly possess interface elements, in varying combinations, that afford interactive experiences which closely resemble puppetry. The major elements of the navigational puppetry metaphor – the puppet, puppeteer, stage, and puppet-view – outline the basic observable components of this trend, which has an obvious relationship to TUIs and prop based direct interactions. Its origins can be traced back to the turn of the millennium and the rise of TUIs in the field of 3D interactions [52].

In order to identify the trend's characteristics in existing solutions, one must be able to associate each of the specific elements of the navigational puppetry metaphor to a complimentary interface element, or combination of elements, of the solution in question. Identifying this evolving interactive idea should give credit to notable previous research that inspired our work and also provide observable and practical case-based justifications for the appropriateness of the metaphor as being an interactive ideal that is being inarticulately stumbled toward.

The most general trend characteristic relates to how the navigation interface presents the VE to the user. The combination of the stage and puppet-view elements of the metaphor form the basic definition of how the output elements of existing solutions can form a multiple perspective display that tends towards puppetry. Any solution that provides a stage-like horizontal interactive surface in conjunction with an accompanying first person view of the VE falls within the puppetry trend. The puppet and puppeteer aspects of the metaphor dictate the next part of the trend definition. Basically, a solution must contain a graspable input object or device that is used within an IT / IM that resembles manipulations of marionettes, or other miniature objects, whose movements are meant to inform navigational behaviors. To be considered even closer to actual navigational puppetry, the solutions must afford the manipulations of the input object to occur within the previously mentioned dual display arrangement, with the input object's position and orientation being graphically synchronized on top of the horizontal surface. The actual physical form of the input objects or devices, and their associated IT / IMs, will also further determine their place in the puppetry trend. The input devices could be anything from generic bricks to specially shaped toy-miniatures. Regarding the IT / IMs specifically, they can also be observed as existing in varied states of evolution towards puppetry. Certain IT / IMs are more related to egocentric steering, while others lean towards more direct bodily character control. Furthermore, any affordances for exocentric control over the stage's position or orientation would place the solution even closer to navigational puppetry.

For example, the increasingly frequent presence of both a first person perspective display and a bird's eye view as an interactive surface, provides evidence of the natural tendency of both designers and users towards a puppetry approach to navigation interactions. If the multiple mirror arrangement on modern automobiles is any indication, having the ability to see one's position and surroundings from multiple perspectives can be incredibly valuable. As we have previously stated, there have been many contributions to the field of 3D navigation that involve the use of multiple perspectives [1] [10] [16]. We argue that this indicates an evolving trend for these kinds of interactive perspectives which has yet to be properly articulated and fully realized.

Perhaps the earliest tangible indications of the puppetry trend are the inventive 3D navigation metaphors and prop-based interface prototypes of Ware et al. [53]. They proposed three different IMs that are directly related to puppetry in different aspects. Their 'eye-ball in hand' metaphor and prototype revolves around providing the user with a spherical physical input object that controls a virtual-eye. The user moves the eye-ball around in real space in an imagined miniature representation of the VE. This idea is taken to the next step with their 'flying vehicle' IM and prototype. The eye-ball input object is changed to a 'bat-like' miniature vehicular object, and it is manipulated like a virtual-eye that has the locomotive functionality of a plane. They also touch on the exocentric side of puppetry type interactions through their 'scene in hand' IM and prototype. This third variation affords navigational control by making the entire VE a graspable entity by connecting the manipulations of the vehicular bat object to the position and orientation of the VE relative to a stationary viewpoint. While this early work may not possess the puppetry display characteristics of a graphically synchronized stage and puppet-view, these elements are present in a conceptual capacity. This research successfully scratched the surface of what could be possible in future puppetry related 3D navigation solutions when IMs and prototypes are creatively designed in conjunction with technologically exploitive prototype implementations.

The ‘Follow my Finger’ solution also utilizes two displays, a vertical first person perspective view and a horizontal touch surface depicting the VE and the position of a virtual camera as seen from above [1]. Using direct touch on the horizontal surface, in conjunction with a few widgets, basic camera control can be achieved on the horizontal surface. This project’s output elements are very close to those of the puppetry metaphor, and they also demonstrates the spatial benefits of perceiving the VE through multiple perspectives. However, while the touch surface gives more directness to the manual interaction, it also highlights a lacking of tactile intimacy and haptic feedback that would be present with a tangible object, like a puppet. Furthermore, the interaction technique of ‘Follow my Finger’ requires additional widgets to achieve complete navigation, which is obviously unrelated to puppetry and also suggests a limited appropriateness of the IT / IM of ‘finger pointing’ for 3D navigation.

Shimizu et al. [45] demonstrate an extremely literal interpretation of puppetry as a form of robotic control interface. Although this particular research is mainly concerned with robotic-bodily control and not exclusively with navigation, it does show the haptic and tactile intimacy benefits that could be achieved through direct and natural puppetry-type manipulations of miniature humanoid input objects. The toy-based TUI for robotic control of Guo et al. [18] takes their solution one step closer to navigational puppetry. They afforded navigational robotic control through the direct manipulation of toys (that are miniature robot representations) on a graphically synchronized horizontal interaction surface, which is essentially a puppet on a stage. Their solution demonstrates the spatial advantages that come with puppetry related interactions. Remotely controlling the movement of robots around real world environments is a main concern for many robotic control tasks, so it is easy to see how these solutions could benefit from changing other interface elements to more closely conform to the navigational puppetry metaphor. For example, in the case of robotic unmanned aerial vehicles (UAVs), the robot-puppet could be manipulated on a stage which is actual satellite imagery of a target region and the puppet-view is actual live video from a camera on the nose of the vehicle.

The Danish toy company LEGO has even begun using various puppetry related tactics to allow shoppers to engage with their line of BIONICLE action figures [32]. They have created a unique interface / kiosk type structure, which uses Virtools and reactTIVision software to display a 3D VE and various characters. The kiosk consists of a vertical output screen and a sensed (yet blank) horizontal interaction surface. When shoppers place a BIONICLE toy-package on the horizontal surface, it acts as a navigational puppetry input device to afford certain basic behavioral controls and movements of an associated graphical avatar of the toy in the VE.

There is also another interesting group of 3D interface solutions that use very literal applications of puppetry to inform their interactions. Mazalek et al.'s tangible interfaces for real-time 3D virtual environments [30], Paiva et al.'s SenToy [33] and MIT Media Lab's "sympathetic" metaphor for toy controlled game characters [25] are all examples of solutions that use plush toys as machinima interfaces. They capture user input as gestures and manipulations of plush toys to control the bodily actions of virtual characters within a game. The link between these interaction styles and puppetry is far more literal than our interaction metaphor. These interfaces are other examples of the trend towards puppetry based solutions, but the focus of these interactions is control over subtle bodily or performative behaviors of the character and much less with navigation. For example, in the work by MIT's Media Lab, the navigation involved is one of many possible behaviors that the toy interface affords, and is accomplished through the combination of specific gestures and system-controlled locomotion.

The ProBono System [16] was designed to investigate if elementary school children could learn real-world survey knowledge from virtual navigation. It consisted of a horizontal hard copy layout of a supermarket, as viewed from above, on which there was a tangible shopper figurine with a cart. The figurine represented the user's virtual self and their first person view was displayed on a traditional output display. Movement and rotation of the shopper would translate and change the orientation of the viewpoint and the corresponding view being displayed on the screen. Once again, this project demonstrates the spatial benefits of multiple perspectives that greatly resemble navigational puppetry (although the horizontal bird's eye view is merely a hardcopy cutout). It also depicts the tactile benefits of using a specially designed tangible controller within an IT / IM that resembles puppetry. However, the figurine only affords moving and turning around on the horizontal surface to provide 2 DOF of translation and 1 DOF of rotation control. The underlying IT / IM is actually similar to puppetry, but only in a simplified superficial sense.

There is also an interesting manifestation of the puppetry trend within the domain of autonomous and semi-autonomous vehicular piloting, which is an intriguing task that bridges virtual and real 3D navigation with robotic control. For example, Quigley et al. [40] experiment with using miniature plane-like models as physical controllers for flying UAVs. Manipulations of the sensed model are translated into roll and pitch commands that are transmitted to the UAV. Although their actual output displays do not necessarily conform to the puppetry trend per se, they do graphically synchronize the input model with real-time flight footage from a camera mounted on the UAV. The camera footage is translated to compliment the position and orientation of the model, while two graphical plane images are overlapped onto the actual footage. One of the plane images is blue and corresponds to the models current state and the other red plane image represents the actual real-world position and orientation of the UAV - which is usually just about to synchronize with the model.

Our metaphor is directly inspired by the underlying trend in these works and others, but our research builds on these previous projects by articulating a domain independent and complete navigation metaphor that affords 6 DOF of control, without the need for widgets. Through our specially designed tangible puppet, the support for complex input and tactile intimacy are facilitated in a distinctly simple and natural fashion. The resulting solution prototype, the Navi-Teer, was used to explore many subtleties of the 3D navigation experience which, to the best of our knowledge, no other system can currently provide.

5.3 The Navi-Teer Prototype

The navigational puppetry metaphor was implemented into the Navi-Teer - a fully functional TUI 3D navigation prototype. For the purposes of our work, we chose to create the prototype at a slightly smaller than room-sized scale, which would enable easy observation of the users as they experimented with the system. In addition, research in the area of display size and its affect on navigational performance has suggested that larger display sizes (at least 76” x 57”) can increase users' sense of presence and they can also increase performance at 3D navigation tasks [34] [48].

The puppet prototype went through two design iterations. They were both made out of a LEGO frame equipped with 3 wireless function buttons from a wireless mouse. In the first prototype design (see Figure 2), which was used for the first usability study, the graspable puppet head was a hollow foam sphere that rested on top of the LEGO frame.

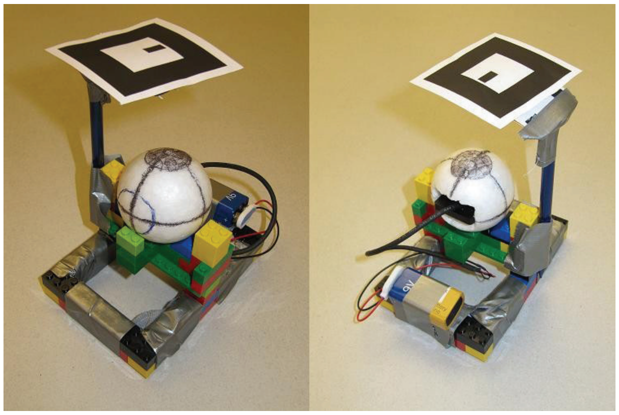

Figure 2 – Puppet Prototype Version 1

The head contained an Intersense InertiaCube 3 wireless 3D orientation sensor. The puppet was outfitted with an optical tracking marker on top of a pole attached to the corner of the body. After the first usability study, the puppet head was changed to a hollow plastic sphere that was attached to a movable ball and socket connection (originally intended to support a sink faucet spray nozzle). This change was in response to certain complaints in the first study about the puppet head feeling disconnected and fragile (see Figure 3). Before the second usability study, a stage-handle for exocentric interactions was also created from LEGO as a vertical graspable column and was outfitted with a second optical tracking marker (see Figure 4).

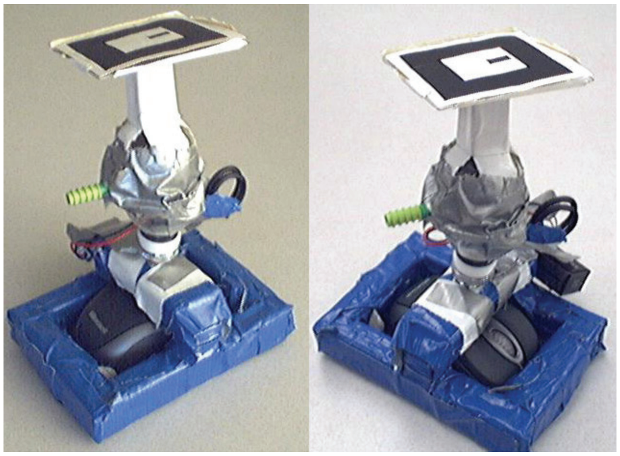

Figure 3 – Puppet Prototype Version 2

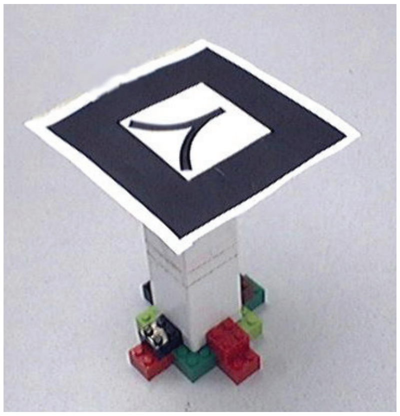

Figure 4 – Stage-Handle

The rest of the TUI is composed of two NEC VT491 projectors; one projecting the puppet-view onto a large vertical rear-projection screen and one projecting a bird's eye view of the VE down onto a horizontal table-top interaction surface, which is used for the stage (see Figure 5). The downward projector constantly displays a bounding box and a look-direction arrow dynamically on top of the puppet. Mounted on the downward facing projector is a Logitech Quickcam 4000 web-camera which tracks the puppet and stage-handle as they move around the stage. The large vertical display is placed behind the horizontal surface so that the users may position themselves as puppeteers directly in front of the stage and facing the puppet-view. The underlying application and VEs for the first two user studies were programmed in C++ using OpenGL (www.opengl.org) and the ARToolkit (www.hitl.washington.edu/artoolkit/). The audio functionality for the spatial audio experiment was programmed in OpenAL (www.opengl.org).

Figure 5 – The Navi-Teer Prototype

5.3.1 The Navi-teer Interaction Protocol

The interaction protocol for the Navi-Teer was choreographed to be as simple, natural, and intuitive as possible (see Figure 6). The basic nature of the navigational puppetry interaction metaphor can best be explained by examining the hypothetical interaction experience of a young child playing with a miniature toy car on a play-mat inscribed with an overview of a town's roads and parking lots. Adding an unrealistic element to the scenario, consider that the child is able see, at will, what his car's driver sees out the front windshield as he drives around. However, this driver-view is in addition to his inherent ability to naturally see himself playing with the car. Furthermore, he is able to literally pick up the town, via the play-mat, and change its position and orientation relative to the car. This playful scenario suggests a puppetry-like interactive experience that approaches navigation both egocentrically and exocentrically.

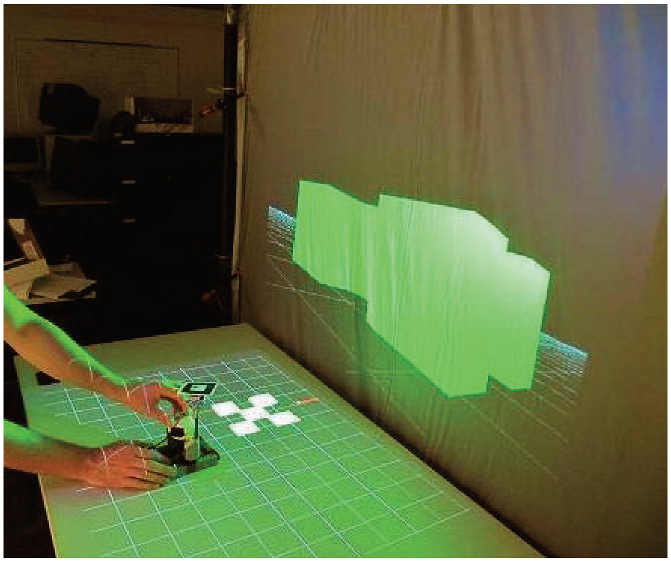

Figure 6 – Interacting with the Navi-Teer

Egocentric navigation with the puppet is the natural default interaction state. In this mode, grasping the puppet and moving it around on the stage translates the viewpoint within the VE. In the default mode, the movement of the puppet on the stage is optically tracked in 2 DOF and results in an equivalent 1-1 movement of the viewpoint on a plane in the VE corresponding to the current height of the puppet – this is called planar travel mode. Alternatively, clicking the appropriate function button on the puppet body will allow for look-directed motion that is not restricted to being planar. In this case, movement of the puppet on the stage results in a 1-1 movement of the viewpoint in all 3 translational DOF within the VE, with the vertical direction coming from where the puppet is looking– this is called look-directed travel mode.

If motion across a greater distance is desired, perhaps to an area of the VE beyond what is currently visible on the stage, then a movement circle can be deployed by clicking the second function button to allow the puppet to ‘fly’. It is merely a circle projected onto the stage around the puppet. When the puppet is moved in any direction to break the boundary of the circle, the viewpoint begins to ‘fly’ in the direction of where the puppet was moved relative to the centre of the movement circle. The speed of the viewpoint’s motion is proportional to how far outside the circle the puppet ends up being pushed (ie: the greater the distance outside the circle, the greater the speed). If the puppet is currently engaged in the planar travel mode, then the travel will be restricted to the current height plane. If the puppet is currently engaged in the look-directed travel mode, then the travel will not be restricted to the current height plane. To stop the motion of the viewpoint, the puppet must simply be moved back inside the circle. Alternatively, the movement circle function button can be pressed again to disable the circle and all motion. At any point, the orientation of the viewpoint can be modified by manipulating the puppet, which suggests asymmetric bimanual use - one hand could perform overall translations, while the other hand could perform orientation changes.

Exocentric navigations represent the other major interaction state. To enable exocentric stage manipulations, the stage-handle is grasped and moved within close enough proximity to the puppet to graphically and computationally link the two interface elements. The stage-handle is now visually connected to the puppet by a green line – this means the stage-handle is in the neutral-engaged mode. When the third function button on the puppet body is pressed, it changes the color of the line to yellow to indicate that the stage-handle is now in translation mode. The stage-handle now functions as if it has been stuck into the stage, which enables the user to translate the stage in 2 DOF in a 1-1 fashion beneath the puppet. If the appropriate function button is pressed once more, then the stage-handle is now in rotation mode. The connecting line between the stage-handle and the puppet now turns red and a red horizontal circle appears with its centre at the current puppet position and a radius that extends to the position of the stage-handle. Manipulating the stage-handle on the interactive surface, at this point, results in the stage being rotated horizontally around the current position of the puppet. To disengage the stage-handle, the function button is pressed again to return to neutral-engaged mode. The stage-handle can then be unlinked by again moving it close enough to the puppet for the green line to disappear.

The main benefit of affording exocentric manipulations is to always allow the user to correctly position virtual objects and areas of interest in order to maintain conceptual coherence between what they can see on the ‘stage’ and what they can see in the ‘puppet-view’. We found that during travel most users require keeping the puppet facing in a 45 degree slice in front of the user in order to avoid disorientation or a conceptual disconnect between the multiple perspectives (see Figure 7). For example, we found that if the user had the puppet facing more towards the far sides or even back toward themselves (outside the forward facing 45 degree slice) then they would have a hard time making sense of what they saw in the puppet-view during travel. Using the stage-handle, they are very able to simply rotate the stage to put their target into the forward facing 45 degree slice.

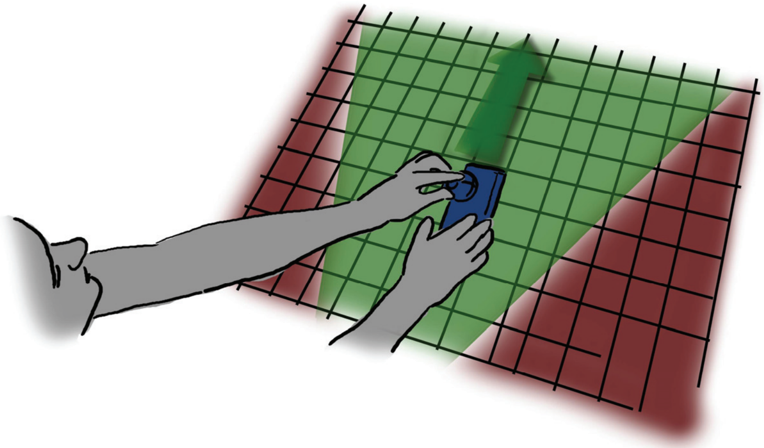

Figure 7 – Forward Facing Interaction Slice

6.0 Studying Navigational Puppetry

Two usability studies and a unique spatial audio experiment are now reported that were designed and completed to prove our IM’s concept and to get inputs for further refining the prototype. The results of these experiments demonstrate that the navigational puppetry metaphor has the benefits of spatial orientation, tactile intimacy, easy capture of complex input and support for collaboration provided by this metaphor.

6.1 Usability Study #1 – Basic Navigational Behaviors

A simple 3D experimental environment was developed to facilitate the first usability study of the Navi-Teer and the navigational puppetry metaphor. Its main purpose was to investigate the system and IM’s basic qualitative usability in some common 3D navigation tasks, as compared to using a 3D mouse (3DConnexion's SpaceNavigator http://www.3dconnexion.com) or using the traditional mouse and keyboard. The experimental environment was a planar grid on top of which the viewpoint moves and all VE objects are located. Three common 3D navigation tasks were developed to closely mimic the basic categories of navigational behaviors of general movement (exploratory), targeted movement, specified coordinate movement and specified trajectory movement.

The resulting experimental tasks were:

- Absolute Motion
- Relative Positioning
- Information-Gathering.

The absolute motion testing environment consisted of target cubes, surrounded by a circle on the horizontal plane, with a flag sticking out their tops. After the target cube appeared in the environment, the user was asked to locate it and travel toward it until the puppet breached the surrounding circle. At that point, the flag drops and the circle changes color.

The relative positioning testing environment consisted of a large green geometric structure which contained four vertical white faces that were only visible when viewed from an almost perpendicular perspective. On each face there was a sequence of numbers. The user was presented with the geometric structure and a floating arrow pointing at one of the faces. Users were asked to position themselves relative to the indicated face in order to read the series of numbers.

The information gathering testing environment consisted of a long twisting and turning planar corridor. At various locations on the walls inside, there were plaques containing short common words and images. The user had to travel through the corridors while attempting to remember the location of the words and images on the walls.

In addition to the testing environments, there was also a training environment which contained examples of every structure described above in addition to a randomly moving stick man that was intended to be followed to get a sense of directed motion (see Figure 8). This is the first environment the users were exposed to so that they become familiar with the device and its usage within this setting before moving on to the actual tests.

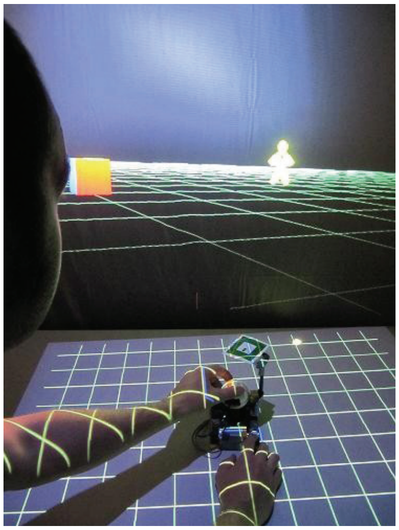

Figure 8 – Training Environment

A step by step breakdown of each task and environment is as follows:

- Targeted motion – navigating to a specific location or object:
    - Test environment: planar grid populated by colored target cubes which are surrounded by a circle and have a flag sticking out of the top (see Figure 9)
    - Task: a target cube appears in the environment the user is asked to locate it and translate toward it until the 'puppet' breaches the surrounding circle when the cube's circle is breached, the flag drops and the circle changes color.

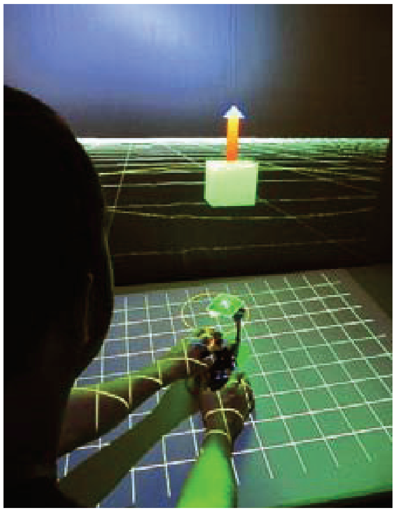

Figure 9 – Targeted Motion Experiment

- Relative positioning – navigating to gain a specific position and perspective relative to an object of interest:
    - Test environment: a large geometric structure which contained four vertical faces only visible when viewed from almost perpendicular angles on each face there is a sequence of numbers (see Figure 10)
    - Task: a floating arrow appears and points at one of the four vertical faces user had to position themselves relative to the indicated face in order to read the series of numbers

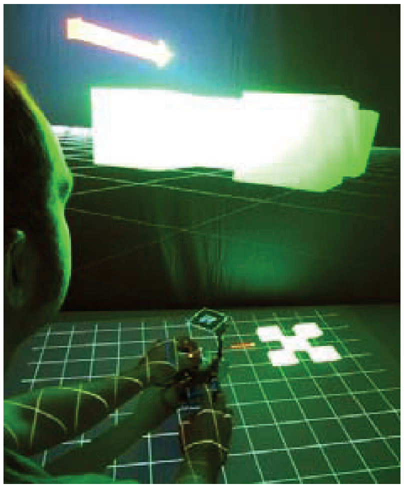

Figure 10 – Relative Positioning Experiment

- Information gathering – navigating through an information laden area attempting to retain as much information as possible:
    - Test environment: a long twisting and turning planar corridor on the walls inside the corridor were plaques containing short common words or images (see Figure 11)
    - Task: the user had to travel through the corridors while attempting to remember the location of the words and images on the walls they were prompted to recall and mark locations on a hardcopy printout of the corridor

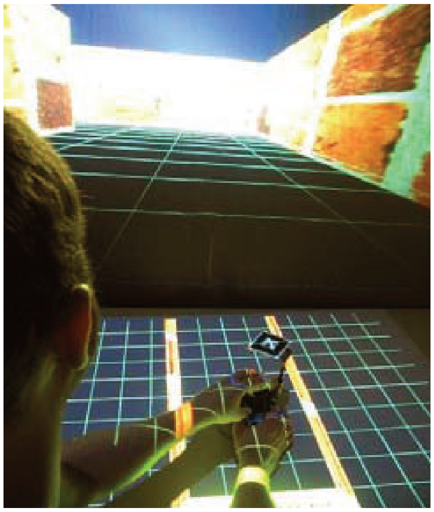

Figure 11 – Information Gathering Experiment

The first usability study consisted of ten users ranging in technical background from PhD and Master's level computer scientists and engineers to non-academic computer hardware and device technicians – all having varying experience with 3D navigation and 3D devices. Before the experiment began, each user read a brief information document giving them sufficient background knowledge on the experiment as well as interaction instructions for the Navi-Teer system. Users were then asked to fill out a pre-experiment questionnaire (see Appendix A), which gathered information about their technical background, previous experience with 3D navigation, or experience with any 3D devices. By way of random number generation, the user was assigned an experimental ordering for using the three possible navigation solutions: mouse and keyboard, 3D mouse, and the Navi-Teer.

The mouse and keyboard functionality used was 'PC Game' traditional, in the sense that the mouse provided toggled 2 DOF of viewpoint orientation control while keyboard buttons provided control of forward/backward translational locomotion and side-to-side strafing. The 3D mouse used was 3DConnexion's SpaceNavigator (http://www.3dconnexion.com). It is a uni-manual 6 DOF 3D navigation device which consists of an elastically sensed joystick cap which is pushed, pulled, raised, and rotated to control the translation and orientation modifications of the viewpoint. When using either the mouse or the 3D mouse solutions, the user was presented with a miniature window of the graphical output for the Navi-Teer's stage just below their main first person perspective window. The user was allowed to get used to each navigation solution in the test environment before each experimental trial began. The entire experiment was audio-visually recorded, with users' consent, and each user was encouraged to 'think-aloud' about their thought processes through the interactions. After the last trial was completed, the users were asked to fill out a free form post-experiment questionnaire to gather information about their general usability opinions about each navigation solution, for each navigation task, and to give their preferred solution for each task.

Results

The results pertaining to the usability of the navigational puppetry metaphor and the Navi-Teer system were extremely positive, especially in the area of overall spatial orientation: eight out of the ten users preferred the Navi-Teer over the other two navigational solutions for all three experimental tasks, one user had no real preference (although this user very much enjoyed using the Navi-Teer and suggested combining aspects of the other two solutions into it), and the remaining user preferred using the 3D mouse for all tasks. It is interesting to note that, based on pre-experiment questionnaire responses, all so-called novice 3D navigation users preferred using the Navi-Teer – most stating specifically that it was very easy to use with no prior experience.

The positive feedback for the Navi-Teer, from the videotape footage and post-experiment questionnaire, came in two main forms. Firstly, almost every user made a positive generalized comment about their experience using the system: *"...makes the most sense in moving in a 3D environment ...", "...gave a real world impression that made it more easy and comfortable...", "...it is more realistic [than the other solutions]...", "...felt very easy to move and look around ...", "...easy and comfortable ...", "...more fun [to use than other solutions] ...", "...gives you more control in the way that the human beings are used to have...", "...closer to how we perform in real world..."*, etc. The comments about the naturalism and realism of the interaction are very interesting since, as previously stated, the system makes use of a unique binary superposition of navigation perspectives. Although it should be very alien to users, the intuitive and playful nature of the Navi-Teer's design appears to succeed in providing a sense of realistic familiarity in an admittedly unfamiliar interface.

Secondly, the other form of feedback describes how the system's architecture and IM provide an increased capacity for acquiring and maintaining spatial knowledge and a great sense of presence – which can be considered together to lead to significant improvement in users' spatial orientation and overall navigation ability: *"...the [horizontal interaction surface] helped me a lot to see where I am...", "...having the tangible puppet on the [horizontal interaction surface] also helped with moving around* *objects and towards them...”, “...though unfamiliar, the correspondence to natural navigation with the [horizontal interaction surface] made this much easier to navigate...* (this comment is again particularly interesting since it acknowledges that user's will initially be unfamiliar with the binary superposition of navigation perspectives, but that it ultimately results in a better interaction experience)*”, “...the [horizontal interaction surface] adds the vital advantage of tracking the things in vicinity very easily ... also helps predict the exact degree of rotation needed...”, “...[the entire system] an ideal choice for exploring/searching things in the vicinity...”,* etc. Compare these comments with those regarding the mouse and keyboard solution and the spatial benefits become obvious: *“...it was a little bit difficult to move around and see the pictures [in the information gathering task]...”, “... its not easy to navigate the 3d surface using mouse and keyboard...”, “...The movement was not easy...”, etc. The same was true, but to a lesser degree, with the 3D mouse solution: “... it was easier then the mouse but still I felt difficulties...”, “... things in the vicinity are not as obvious [compared to Navi-Teer]...”*, etc. Other comments even got more specific, like how being able to move the puppet's head around while in locomotion is a great way to handle viewpoint orientation control – particularly in the information gathering task: *“...especially in the last task [information gathering] ... the ability to turn the eye so naturally was a great asset...”, “...very easy way to set your viewpoint...”, “...have the feeling of full control of your looking direction...”,* etc.

The only critical comments received in regard to the Navi-Teer system were more related to the prototypical nature of the system: some complained of arm fatigue due to the size of the horizontal interaction surface (users made suggestions about being able to customize the size), some complained about the lack of firmness of the foam eyeball (users made suggestions that it should be attached, yet still be free moving and able to be locked in a particular looking direction), and users also complained about the placement of the function buttons. All of these comments were taken into account and changes were made to the Navi-Teer before the second usability study was carried out. Specifically, the puppet's head and buttons were redesigned and also exocentric functionality was added through the addition of the stage-handle device.

6.2 Usability Study #2 – Spatial Orientation and Wayfinding

The second study was then designed to investigate spatial orientation and wayfinding issues of the Navi-Teer, in comparison with the traditional mouse and keyboard interface solution as the first usability study. The 3D mouse option was not included in all experimental trials due to time constraints, so this solution option was ignored during the analysis. The area of concentration for this second round of experimental tasks was based on the observed benefits from the first usability study. An effort was made to take those benefits and use them as inspirations to construct two more complex experimental tasks that focus specifically on spatial orientation and wayfinding. Also, in additions to the changes made to the Navi-Teer after the first study, the basic functionality of all three solution variations were augmented to allow for a simplified form of ray-casting to be afforded. Specifically, the position and direction of the rays were made to be recorded for use during this study. Although this study involved the collection of actual quantitative experimental data, the intent was to informally observe and interview user's about the qualitative details of their interactive experience within formally designed and executed experimental trials.

The first task involved investigating spatial orientation and the build-up of wayfinding knowledge through experimenting with disappearing landmarks within a certain walled sub-region of the VE. The environment consisted of colored geometric shapes and various cubes with large images displayed on each face (see Figure 12). The user is sent around the environment to visually locate and pass through a series of landmarks and at some point all of the landmarks disappear. The user is then asked to point their ray in the direction of where they remember certain landmarks having existed. The pointing error of each ray cast was recorded, but this was only intended to provide a general idea of comparative accuracy and not for any conclusive quantitative claims.

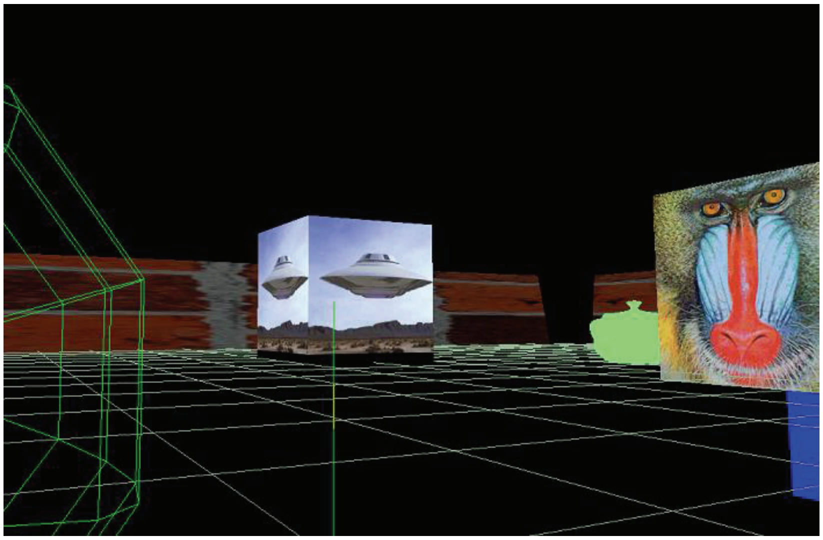

Figure 12 – Disappearing Landmark Experiment

The second task investigated similar issues, but on a much larger scale. The environment was now a large city-like structure with many buildings and streets. A few of the buildings have images on their walls and roofs (see Figure13). The user is allowed to explore the city until they have come across each of the image buildings once. The user is then taken back to a certain point on the outskirt of the city and asked to find their way to a certain image building. From that building, they are then asked to go to another image building, and so on. At their last destination, the user is also asked to cast a ray in the general relative direction of remaining image buildings. The travel times to each destination and also the pointing errors of each ray cast were recorded, but only to provide a general idea of comparative accuracy and times and not for any conclusive quantitative claims.

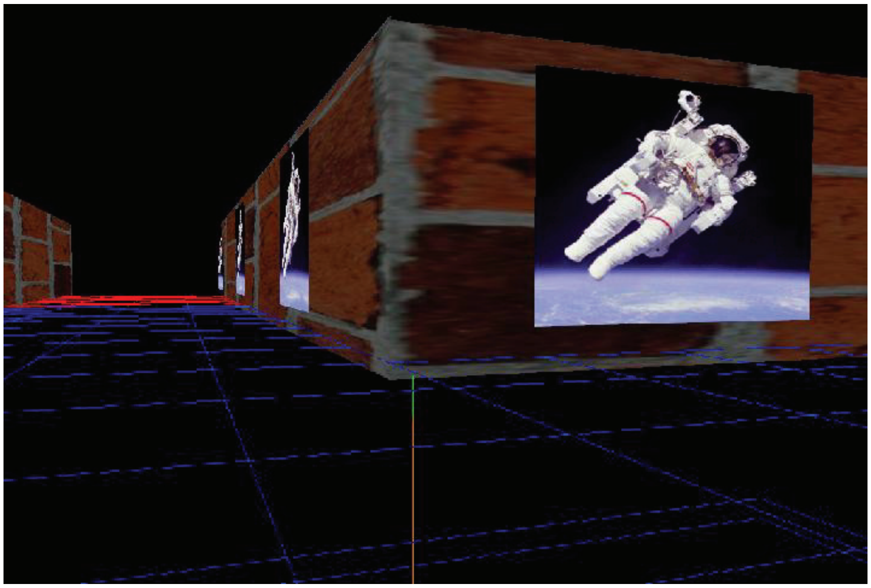

Figure 13 – City Locations Experiment

The step by step breakdown of each task and environment is as follows:

- Disappearing landmarks:
    - Test environment: a large room populated with specific pictures displayed on all the sides of cubes and colored geometric shapes
- Task: the user is asked to locate a few specific shapes and images all images and objects disappear then user was required to cast a ray, which originated from the ‘puppet’ head and pointing in the look-direction, and position it through the previous locations of certain objects and images. The direction of the ray and also the user’s certainty about their guesses were recorded.

- City locations:
    - Test environment: a large city-like structure with equidistant buildings and connecting streets. Certain equally spaced buildings had images on their walls and roofs.
- Task: user was asked to find their way from one particular image-building to another. Once they had visited all the image-buildings, they were asked to travel back to one specific image location as fast they could from that location. They were then asked to again cast a ray in the direction of where they thought certain remaining image-buildings were located

The experiment consisted of ten subjects from diverse academic backgrounds: a mixture of PhD and Master's level Computer Science, Communications, Physics and Biology. Before the experiment began, all participants completed a pre-experiment questionnaire (see Appendix A).

Results

The results of the spatial orientation and survey knowledge study were somewhat mixed compared to the first experiment and we now had both qualitative and quantitative data to consider. While the Navi-Teer was easily and successfully used by all of the participants for each task, the initial observations and qualitative data did not suggest that users objectively performed significantly better with it. We believe that the nature of these experimental tasks are such that any observable quantitative differences in the solutions would be related to very low level and technical aspects of how travel is accomplished, and less with higher level affordances such as those for spatial orientation. For this reason, we decided to only focus on the qualitative user experiences rather than to further analyze the numerical data from this usability study.

Listening to the users describe their experiences using the Navi-Teer seemed to show that the manual interactions with the puppet on the stage led to a very good conceptual sense of where things were located and greatly facilitated the building of strong wayfinding knowledge. However, quantitatively, the actual time differences for the second task did not prove any one of the two solutions to be superior. In hindsight, it also became obvious from observing each subject and analyzing their final feedback that real-time reactions to the Navi-Teer's interaction technique and the spatial / survey aspects of each task both competed for the subjects' attentions. In other words, it was hard to distinguish observations and feedback about the prototype and interaction metaphor in general from those regarding how it specifically supported the spatial and survey aspects of the tasks compared to the other solutions. It became obvious that the tasks for this study were still inadequately designed and not complex enough to yield enough usable data. This led to the decision that an existing, more complex experimental task needed to be chosen (or created) that focused more on higher level navigation, tactile intimacy and spatial orientation. The new task must not be purely navigational in nature; rather, it should use 3D navigation as a means to achieve some larger goal. In that case, the navigation solution can be examined in the context of how it helps achieve the overall goal and not in terms restricted to its own features and they apply to the act of navigation only.

6.3 Spatial Audio Experiment with Expert

The decision was made to look into finding or creating a task that involved navigation to support 3D audio design and experience. We decided that the sensory experience similarities between basic visual navigation and the auditory exploration of a soundscape made tasks related 3D audio design the perfect venue for an experiment with the Navi-Teer interface. This led us to involve a spatial music expert / artist in our work, who would use the Navi-Teer interface as a design and performance tool for creating and experiencing a musical soundscape. A unique task was devised, based on their work, which would require a modified version of the Navi-Teer 3D navigation interface. In the case of navigation for spatial sound, the puppet must be modified to possess the user's virtual ears and the stage must embody the audiovisual soundscape through which the puppet-ears explore and perceive sound.

To enable basic soundscape modeling, the Navi-Teer was modified to afford a simple ray-casting method to grab, position, and orient sound objects in 3D space. Sound objects were represented within the VE as solid color geometric shapes - either spheres (sounds with no directional orientation) or cones (sounds with directional orientation), whose sizes are relative to their volumes (see Figure 14). These sound-spheres and sound-cones were imported into the virtual soundscape by bringing them out of a sound-repository cube, through ray-casting, in a predetermined order. The sound objects could be set to activate when the puppet was within certain proximities and they could also have been given their own dynamic paths to move around while the puppet experiences the soundscape. Furthermore, all sounds could be simultaneously played / stopped in unison, or each sound could be turned on / off independently by ray-selection. At any time, the path of the puppet through the soundscape could be recorded and played back to repeat a desired soundscape experience.

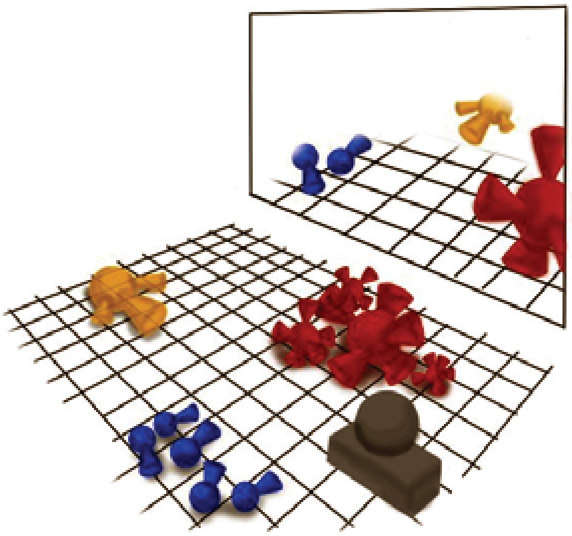

Figure 14 – Using the Navi-Teer for Spatial Sound

Expert's Background

The expert's current work relates to investigating the relationships "between music, place, and mobility" [50]. They examine the experience of an individual listening to music on a personal music device as they move through parts of the city and perceive the existing ambient soundscape of city-noise. In an attempt to explore the interactions between "the real and the imagined; the real and the mediated; the soundscape and music; everyday life and art; and the senses" [50], they have pioneered a certain kind of spatial music travel experience.

A certain travel route is outlined, for example: walking from here to there, then taking a bus, then taking an underground subway. The ambient soundcape of each travel region is recorded using a portable digital recorder and stereo condenser microphone. The raw field recordings are then modified and mixed into actual musical compositions, or musical-movements, using existing sound manipulation software: Adobe Audition for multi-tracking and sound editing (http://www.adobe.com/products/audition/), Propellerhead Reason for sequencing and midi control (http://www.propellerheads.se/products/reason/), and Ableton Live for live looping and sequencing (http://www.ableton.com/live). The resulting creations are arranged to form a soundtrack for traveling through one of the three sections of the route. In a practical experiment, participants physically traveled through the route's existing soundscapes while they listened to the musical-movements made from sounds recorded within that space.

The objective of this investigation into music and mobility was primarily to reframe how we look at the act of travel with a personal sound device. Travel, specifically walking and using public transportation, is usually regarded as a devalued activity. People listen to audio through their personal devices to restore enjoyment during these periods, and the act is usually regarded as an attempt to separate the listener from their surroundings or situation, at least in an auditory sense. Our expert's approach re-addresses the issue by blatantly drawing attention to the fact that one is listening to music and moving in an existing audio filled soundscape, rather than considering them separate or opposing elements of the experience. Furthermore, the musical compositions themselves play with the boundaries between music and environment and provoke moments of confusion as to the actual source of certain sound elements [50].

Pre-Experiment Interview with Expert

Before the expert used the Navi-Teer interface for 3D audio, they experimented with it within the existing basic 3D navigation experimental environment, which had been used previously for a preliminary usability study. Afterwards, an interview was conducted to determine their basic reaction to the interface and how it would function as a front-end interaction for spatial audio design.

The discussion began with the apparent benefits of such an interaction and how they would benefit and shape the 3D audio design process. The tactile blurring of input and output spaces and the dual perspective of egocentric and exocentric outputs was of great interest to the expert, both in terms of basic interactive novelty and also as being very appropriate for audio design. The intuitive tactile intimacy of interacting with the puppet was discussed as being of great benefit to non-technical users, like an artist or director. They could demonstrate every aspect of the navigation or experience, even some excessive aspects which may be difficult to capture otherwise (ie: style, personality, etc...).

The expert then started to brainstorm on how the system would be used as a tool for their spatial audio creations. First, in terms of the basic interactive experience, the interface provides a unique opportunity to be tangibly immersed within the sounds one is attempting to manipulate, rather than being removed to viewing 2D waveforms from the side and grasping at knobs and faders. Second, from a sound manipulation task affordance perspective, the interface provides an extremely accessible tangible method for easily achieving very complex audio manipulations – such as irregular and/or improvisational 3D panning or easier ways to mimic dictated listening experiences. Thirdly, the expert discussed the interface as adding a performative or observational significance to the act of audio manipulation or design, both in terms of the observed manual interactions and the accompanying audiovisual outputs.

The expert also talked about their work specifically and how having a tangible soundscape to manually grasp and explore would greatly benefit the process by tangibly realizing the conceptual soundscape and experience they would be trying to build. Having completed a series of these movements previously, they had an existing repository of both raw field recordings and prepared tracks. Since they had completed this task previously using conventional sound design tools, they were in a good position to comment on the differences in the design process when compared to the Navi-Teer interface.

The Spatial Sound Experiment

The expert used the Navi-Teer interface as a final 3D mixing tool for a particular portion of one of their musical-movements intended for riding the bus. He arrived with nine prepared audio .wav files. Certain files corresponded to elements of what would become the rhythm or percussive elements of the composition – made up of various sounds from raw field recordings taken while riding the bus. Other sound files would become the melody elements of the movement, such as various hums and ambient noises, which when arranged in a specific manner produced interesting melodic riffs.

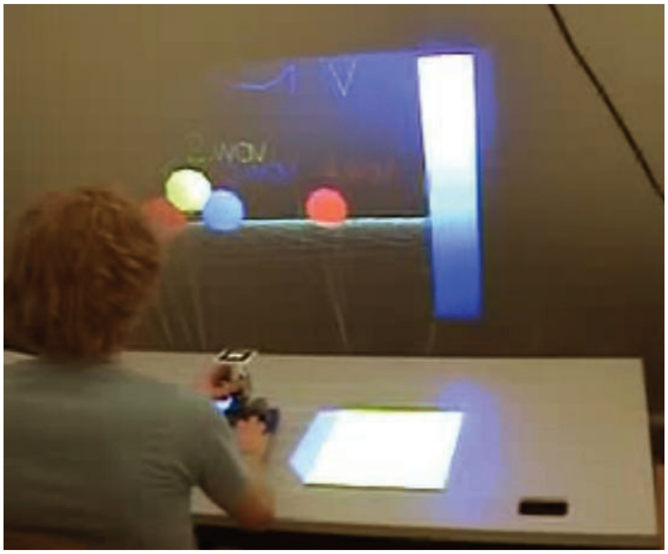

Figure 15 – Spatial Sound Experiment

Prior to the day of the experiment, it was discussed whether or not any existing VE should be present before actual soundscape modeling would take place. For example, having a virtual representation of at least the most superficial aspects of the bus path and any pertinent auditory landmarks (ie: other vehicles, construction areas, people, etc…). However, it was decided to leave the existing environment empty to make it easier to focus strictly on the spatial audio interactions. This left the ability to later discuss the potentials of incorporating more complex visual VE elements.

The first portion of the experiment related to the expert trying to recreate an introductory section of the musical-movement. He wanted to simulate the beginning of the bus ride with an approach of a specific rhythm that would then blend into another rhythm and melody combination. To accomplish this, he first positioned a sound-sphere, representing the track for the introductory rhythm (sphere-A), in an open space next to the sound-repository-cube. Then he positioned five other sound-spheres ahead of the already placed sphere. Two rhythm spheres whose auditory interaction – depending on which rhythm is panned to more than the other – form the desired rhythm to be experienced after the introductory portion is completed. These spheres were placed far from each other, on both the left side and right side relative to the previously placed sphere (sphere-B and sphere-C, respectively). The three remaining spheres, relating to another rhythm element and a basic melody, were arranged to form a triangular structure with one sphere resting above an open space (sphere-D) between two other spheres on the horizontal plane (sphere-E and sphere-F).

There was now a simple soundscape which could be experienced with the Navi-Teer interface (see Figure 15). After some epistemic experimentation with exact placements and timings for sound activations, they were ready to experience the soundscape to obtain the desired portion of the musical-movement. He began by positioning themselves far below sphere-A, with their view pointing towards the other spheres. He activated the sound spheres, one by one, to achieve certain auditory interactions that occur at regularly spaced intervals. Once all the spheres were ‘singing’ properly, he deployed a movement circle and began to slowly travel towards sphere-A. As the expert traveled towards the remaining soundscape, the rhythm and melody of spheres B to F began to fill the auditory void as the departing introductory beat slowly faded behind the puppet. The expert then continued traveling to the open space within the triangular structure containing spheres D, E, and F. From that position in the soundscape, the desired melody and beat interaction was achieved. It is interesting to note that from that position within the soundscape, the expert was presented with a very unique visual representation of what they were hearing.

The next thing that the expert experimented with was a way to achieve an interesting revolving melody and rhythm auditory interaction that they thought would work well for the actual preliminary riding of the bus. He tried various manual puppet paths with no movement circle deployed. He wove in and around the sound-sphere layout and achieved some very interesting sound interactions. After a period of experimentation, he decided on a particular path which captured a sound experience which would work perfectly for the first portion of a bus ride. The interaction was a figure eight style loop or swoop of the puppet – beginning in the center of the triangular structure and going out and around one of the outer spheres, then passing back through the center of the triangular structure, then out and around the other sphere. The auditory result was a very peculiar merry-go-round of repeating rhythms and melody.

This interaction was quite interesting for two main reasons. Firstly, the complexity and irregularity of the interaction in terms of sound mixing and panning was very easily perfected and captured after a minimal period of epistemic experimentation, while simultaneously tangibly realizing what would have been previously a conceptual black-box process to any observer. Secondly, compared to the literal nature of the first interaction – moving past sphere-A to 'get on the bus' - where there is a direct spatial correlation between movement in the soundscape and reality, the figure eight style sound perception portrays almost a totally abstract or excessive type of soundscape interaction which really has no relationship to reality. It seemed as if there were a host of new potential audio manipulations that could be done in the virtual soundscape that, while not having any direct connection to the real-world, could be used to augment the musical-movement while still remaining appropriately connected to the real world. Specifically, if the soundscape was set-up to represent a miniature outline of the path to be traveled in the real-world, then portions of the journey – like the bus ride – can be represented and extended by excessive and complex interactions around a sub-section of the entire soundscape.

The final aspect that was investigated by the expert was related to the capturing of an interactive performance using the Navi-Teer. Part of the existing functionality of the spatial audio interface allows the entirety of an interactive session to be recorded and played back - both the manual interactions and the resulting audio. The soundscape which was used previously was modified to include some dynamic and independent movements from the actual sound-spheres. Certain spheres were made to rotate in a circular fashion, both vertically and horizontally, about separate points of origin. The expert then experimented with different navigation paths through the soundscape, which now had several moving parts. Once a favorable path was determined, they began trying to achieve the optimal sound experience by trying to pass through the moving structures at particular points in time while increasing and decreasing their travel speed. With the recording aspect turned on, and once they felt they achieved the 'best' pass, the interaction was captured and saved for playback. There were two main results: first, the actual audio output was saved for playback or exportation. Secondly, the particulars of the manual interaction were captured and played back as an audiovisual presentation. Hence, an incredibly complex, and definitely difficult to explain, spatial audio interaction can be captured and then played back to visually illustrate exactly what is happening to any observers or collaborators.

Post-Experiment Interview with Expert

After the experiment was concluded, the expert gave an extended interview on their experience with the Navi-Teer interface. The resulting feedback contained general qualitative comments on the interface and experience as a whole, a comparison to previous solutions for the same audio design task (his previous musical movement design and creation process), and comments about new spatial audio actions and performances that may be possible and worth experimenting with.

One of the first comments was on the entirety of the interactive experience of tangible navigation as it relates to 3D audio. They stated that the Navi-Teer's unique linking of time, sound, and space makes for an interesting and immersive interaction experience. The extra focus on space and movement, combined with the usual temporal aspects of sound and music manipulation, changes the entire audio interactive experience. In essence, they said, there is far less distance between the individual and 'the sounds' due to the blurring of input and output spaces. There was a noticeable increase in the tactile immediacy of actions and responses, which is actually a more natural and realistic way to work with sound. They remarked that the interface was reminiscent of older and more analog methods of sound manipulation, even going as far as characterizing the puppet as being instrument-like itself due to its tactile interactive nature and its illusory disconnected state.

The expert stated that the basic nature of the interaction metaphor afforded an increased sense of freedom towards epistemic actions. This allowed them to experiment with many different ideas on the spot, and also to perfect certain specific movements if so desired. In addition, they also commented on how this epistemic freedom allowed their imagination to run freely with new ideas for different kinds of design tasks and auditory interactions. They made it very clear that this aspect would be incredibly useful in the design process of any kind of spatially oriented audio task.

When the Navi-Teer's interactive process was compared with what would was done previously to create the musical-movements using traditional systems (ie: Adobe Audition, Propellerhead Reason, Abelton Live), one main aspect was glaringly obvious. A substantial portion of their previous spatial sound design methodology remained purely conceptual. Certain audio manipulations were being done to simulate imagined effects or scenarios, which would appear as a totally black-box process to any outside observers or collaborators. However, approaching similar spatial sound tasks from a navigational puppetry perspective allows for a tangible realization of these previously conceptual ideas. The underlying soundscape of the musical compositions was now literally presented to be grasped and explored.

The expert also discussed new possibilities and ideas for the interface and what it could be used for within the general domain of spatial sound or music. First, they discussed possibilities related to augmenting the graphical fidelity of the visual sound representations with the soundscape. Specifically, if the sounds could be represented by real-time 3D audio waveforms, then the puppet's perspective view of the audiovisual environment would become a very useful immersive 3D wave view of the surrounding sounds that will be heard. Having the ability to tangible explore these waveforms would be very interesting interactively, both in a visual and tactile sense, but it would also increase the performative or observational value immensely. The same principle would also apply to any general graphical augmentations to the underlying VE of the soundscape.

Finally, the expert discussed the potential of using the Navi-Teer interface as an instrument, which was an idea that they had been cultivating throughout the experiment. Basically, the suggestion was made to use the interface as a way to perform improvisational musical interactions between various sound waves. Their basic idea was to populate the soundscape with various sound objects containing simple but varying repeating sound wave patterns. Moving the puppet, or the sounds themselves, closer or farther away to other sound objects would produce very interesting auditory interactions. Also, if the sound object visual fidelity could be improved, as previously mentioned, then the resulting visual experience could also be very interesting as the waves can be visibly seen interacting with each other while the resulting auditory blend is heard.

6.4 Discussion with Game Level Designers

After completing two usability studies and a unique spatial sound experiment, we hosted multiple demo sessions for level designers from the local game design industry. Two avenues for future work were identified based on those meetings.

The first involved the possibility of using the Navi-Teer interface as a rapid level-design and experience tool to aid in some of the logistical bottlenecks that occur during iterative level design. The affordances for capturing complex and irregular input (from possibly non-technical individuals), support for quick epistemic actions and possibilities for collaboration would be invaluable during a game level-design process.

The second possibility for future work involves developing an underlying game specifically designed to accent the important benefits of the Navi-Teer prototype and the navigational puppetry metaphor. At this point we are in the design process of a navigational puppetry version of a FPS game. The puppet tangible controller will be augmented with a trigger on the head which will fire projectiles in the direction of where the puppet is looking. We are also considering adding a second puppet and a second puppet-view to enable head-to-head gaming on a shared stage.

7.0 Conclusions and Future Work

7.1 Conclusions

The goal of this research was to create and investigate a new domain and task independent 3D navigation metaphor and prototype, which we intend to be a candidate for the navigational portion of a unified 3D interaction metaphor. We have presented the navigational puppetry metaphor and positioned it as a distinct articulation of the front wave of an observed puppetry trend in 3D interaction solutions, which was gained from an extensive literature review. Two usability studies and a unique spatial audio experiment were also reported that observed and demonstrated, respectively, the benefits of spatial orientation, tactile intimacy, easy capture of complex input and support for collaboration.

Toward the achievement of a unified 3D interaction solution, this work is a case study of the successful exploitation of available interface technologies and interactive potential of TUIs to create a unique interaction metaphor and 3D navigation solution. This project was completed independently from design to implementation, with the exception of the input on the sound experiment from our audio expert. This particular aspect should portray the potential value in forcing interface designers to address all aspects of solution design and prototype implementation – we have shown that this is one possible way to assure that the previously mentioned exploitation occurs. The metaphor articulates, in distinct clarity, properties of a related interface and interaction trend that we have demonstrated to possess extremely relevant benefits to the act of 3D navigation. Our recognition of this puppetry trend makes our navigation metaphor unique in an incremental context, but this step is distinctly important due to our harnessing of this interactive idea and redirecting it into a complete 3D navigation metaphor. We now have a solution type that prescribes a new path towards a unified 'puppetry' 3D interaction solution that has yet to be explored. Through our spatial sound experiment, we illustrated how its advantages can be appropriate in new and emerging 3D interaction task domains.

The main contribution of this research is an incrementally inspired, yet unique, 3D navigation metaphor and prototype – Navigational Puppetry and the Navi-Teer. Of equal importance is the connection of this metaphor to a never before articulated interactive trend, existing in many different types of 3D interaction solutions, that we have described as approaching the principle of puppetry. The recognition of this trend is an important contribution because it justifies the design ideas of our metaphor and also inspires a new avenue for potential unified 3D interaction solutions that are directly connected to common aspects of existing solutions. The unique audio task that we devised for our experiment is also a case study contribution to the field of TUIs as audio interfaces. It successfully demonstrated the general appropriateness and usability of the navigational puppetry metaphor for a previously unexplored 3D navigation task, which should be a property of any candidate for a unified solution.

7.2 Future Work

Throughout the course of this research, there have been several discoveries of possible areas for future work. We have already mentioned two possibilities that came as a result of demos and discussions with game level designers - the first involved the possibility of using the Navi-Teer interface as a rapid level-design and experience tool and the second was developing a navigational puppetry version of a FPS game. We will now present the other major opportunities for future work.

7.2.1 Developing New Navigational Perspectives

The issue of egocentric vs. exocentric navigation perspectives has only been investigated at a very superficial level. There have been many works that define these terms and experiment with them as distinct solution options. There have also been works, like ours and many other TUI based 3D navigation solutions (ie: Follow-My-Finger [1], BUILD-IT [10]), that have begun to experiment with combining and blurring the boundaries between these two options. It is entirely possible that these are not the only possible choices for navigational perspectives; they could just be mere articulations of two opposing ends of a spectrum. An interesting task for future work would be to investigate what distinct and useful perspectives may lie in the middle of that spectrum and also what other completely different perspective types or dimensions may exist. For example, it is easy to see the connection of egocentric and exocentric perspectives to real-world navigation, or more accurately, to the navigation of an individual or vehicle. Given that restrictive navigational paradigm, it is hard to conceive of other possibilities. However, there may be fundamentally new ways to look at navigational perspectives that are afforded by connecting 3D navigation to such unlikely paradigms as:

- quantum mechanics / particle superposition
- platonic solids
- meteorological patterns

- aquatic marine life (individual and group movement characteristics)

7.2.2 Creating a New Definition / Taxonomy for 3D Navigation Solutions

Many 3D navigation solutions are described in ad-hoc and case-specific ways. There have also been various formal categorization protocols and taxonomies proposed (ie: Bowman [3], Tan et al. [49]) that have intended to described solutions by breaking them down into their lower-level components and connecting them to the major elements of their own definition of ‘what 3D navigation is’. The main goals of these efforts are to be able to make formal comparisons to other similarly categorized solutions and to be able to find opportunities for new designs. There is a real danger for unintentionally biasing solutions and limiting design potentials towards existing navigational theory. This whole approach supports a type of solution creation called ‘guided design’, which is characterized by a higher concentration on existing standards rather than innovation. There could be great value in the creation of a new unified method for describing 3D navigation solutions that enables formal comparisons and the observance of new design options, while also encouraging a liberal design process towards new interactive experiences.

7.2.3 Developing New 3D Navigation Performance Metrics

There have been 3D navigation performance metrics that have been proposed that are tangible and quantitative in nature (ie: time, accuracy, etc.) and those that are more abstract and qualitative (ease of use, presence, comfort, etc.). It appears that many traditional quantitative metrics suffer from an overall objective inapplicability and many traditional qualitative metrics suffer from either being subjectively biased or merely being inappropriate or too simple to apply to any large claims. For example, it is arguable whether a metric like ‘sense of presence’ or ‘sense of immersion’ are well suited to adequately describe the complexities of the extent to which a user conceptually connects with a VE during navigation.

There is a real need to develop a new standardized set of complex performance metrics that represent the many tangible and intangible particularities inherent in 3D navigation and 3D interaction. For example, consider the common navigation solution output element challenge of trying to adequately display vital 3D perceptual cues to the user. There could be real value in the creation and utilization of a complex metric, which could have both quantitative and qualitative aspects, which attempts to measure the extent to which the user is able to visually and conceptually perceive three distinct dimensions.

7.2.4 Developing a New 3D Navigation Solution Evaluation Method

Evaluations are usually intended to be the basis for comparisons, to determine particular task suitability, or in our case – to determine the appropriateness of a given solution as a candidate for a unified IT / IM. However, it can be difficult to both discover and prove new interactive benefits due to many possible perceived inadequacies in particular experimental designs. There is currently little guidance to offer towards a better or more standardized option. The creation of a new and distinctly standardized evaluation strategy could be extremely valuable to the research community.

In order to design such an evaluation procedure, there must be an articulation of a set of accepted experimental tasks and designs. The issues would become what sort of experimental tasks best represent the true nature of 3D navigation. Many current experimental tasks are biased towards pre-existing ideas on the fundamental and historical aspects of navigational behavior, which may be quite different from the true interactive potential of unified 3D navigation. Another experimental design decision that must be made is the choice between informal evaluations or being formal and statistically valid. Informal studies are better at getting high-level observational data that describes many potentially unknown aspects of the interaction experience. However, they can be easily attacked for their lack of strict provability and may be less useful for comparative analyses. Conversely, formal experiments can provide easy to use quantitative results that can be used to calculate provable and statistically valid conclusions. A great research challenge would be the creation of a set of adequately complex experimental tasks, which reconcile our current navigational understanding with the interactive potentials we envision that navigation is headed toward, that produces a useful combination of formal and informal results for analysis.

REFERENCES

1. Ajaj, R., Vernier, F., and Jacquemin, C. Follow My Finger Navigation. In *Proceedings of the 12th IFIP TC 13 international Conference on Human-Computer interaction: Part II* (2009).

2. Balakrishnan, R. and Kurtenbach, G. Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces. In *Proceedings of the CHI*'99. ACM, (1999).

3. Bowman, D. Interaction Techniques for Common Tasks in Immersive Virtual Environments: Design, Evaluation, and Application (Unpublished Doctoral Dissertation). Georgia University of Technology, (1999).

4. Bowman, D., Fröhlich, B, Kitamura, Y., and Stuerzlinger, W. New Directions in 3D User Interfaces. In *Proceedings of the 2005 IEEE Conference 2005 on Virtual Reality*. IEEE Computer Society. (2005), 312-342.

5. Bowman, D., Kruijff, E., LaViola, J., and Poupyrev, I. An Introduction to 3D User Interface Design. *Presence: Teleoperators and Virtual Environments*. *10*(1), (2001), 96-108.

6. Crevoisier, A. and Polotti, P. Tangible acoustic interfaces and their applications for the design of new musical instruments. In *Proceedings of the 2005 Conference on*

New interfaces For Musical Expression. New Interfaces for Musical Expression. National University of Singapore, (2005), 97-100.

7. Czerwinski, M., Tan, D.S., and Robertson, G.G. Women take a wider view. *Proceedings of CHI 2002*. CHI, (2002), 195-202.

8. Fishkin, K.P. A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, *8*(5), (2004), 347-358.

9. Fitzmaurice, G.W. and Buxton, W. An Empirical Evaluation of Graspable User Interfaces: Towards Specialized, Space-Multiplexed Input. *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, ACM, (1997), 43-50.

10. Fjeld, M., Ironmonger, N., Schar, S.G. and Krueger, H. Design and Evaluation of four AR Navigation Tools Using Scene and Viewpoint Handling. In *Proceedings of INTERACT 2001 8th IFIP TC.13 Conference on Human-Computer Interaction*. (2001), 167–174.

11. Fredriksson, J., Berg Ryen, S., and Fjeld, M. Softly Elastic Device for 6DOF Input. In *Proceedings of ACM NordiCHI*. ACM, (2008), id5-id6.

12. Fröhlich, B. The quest for intuitive 3d input devices. In *HCI International*, (2005).

13. Frohlich, B., Hochstrate, J., Kulik, A., and Huckauf, A. On 3D Input Devices. *IEEE Computer Graphics and Applications*. *26*(2), (2006), 15–19.

14. Frohlich, B., Hochstrate, J., Skuk, V., and Huckauf, A. The GlobeFish and the GlobeMouse: Two New Six Degree of Freedom Input Devices for Graphics Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, (2006), 191-199.

15. Fröhlich, B. and Plate, J. The Cubic Mouse: A New Device for Three-Dimensional Input, In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, (2000), 526-531.

16. Göttel, T. ProBoNO: transferring knowledge of virtual environments to real world situations. In *Proceedings of the 6th international Conference on interaction Design and Children*. IDC '07. ACM, (2007), 81-88.

17. Grabe, M.E., Lombard, M., Reich, R.D., Bracken, C.C., and Ditton, T.B. The role of screen size in viewer experiences of media content. *Visual Communication Quarterly*, 6, (1999), 4-9.

18. Guo, C. and Sharlin, E. Exploring the use of tangible user interfaces for human-robot interaction: a comparative study. In *Proceeding of the Twenty-Sixth Annual SIGCHI*

Conference on Human Factors in Computing Systems. CHI '08. ACM, (2008), 121-130.

19. Guzman, E., Wai-ling Ho-Ching, F., Matthews, T., Rattenbury, T., Back, M., and Harrison, S. EEWWW!!: tangible interfaces for navigating into the human body. In *CHI '03 extended abstracts on Human factors in computing systems*. ACM, (2003), 806-807.

20. Haik, E., Barker, T., Sapsford, J., and Trainis, S. Investigation into effective navigation in desktop virtual interfaces. In *Proceedings of the Seventh international Conference on 3D Web Technology*. Web3D '02. (2002), 59-66.

21. Hand, C. “A Survey of 3D Interaction Techniques.” *Science*. 16(5), (1997), 269-281.

22. Hansen, A. S., Overholt, D., Burleson, W., and Jensen, C. N. Pendaphonics: a tangible pendulum-based sonic interaction experience. In *Proceedings of the 3rd international Conference on Tangible and Embedded interaction*. TEI '09. ACM, (2009), 153-160.

23. Hinckley, K. and Pausch, R., Goble, J.C., and Kassell, N.F. Passive Real-World Interface Props for Neurosurgical Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence*, ACM, (1994), 452-458.

24. Ishii, H. Tangible Bits: Beyond Pixels. In *Proceedings of the TEI'08*. ACM, (2008), 15-25.

25. Johnson, M.P., Wilson, A., Blumberg, B., Kline, C. and Bobick, A. Sympathetic interfaces: using a plush toy to direct synthetic characters. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*. ACM, (1999), 152-158

26. Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international Conference on Tangible and Embedded interaction*. TEI '07. ACM, (2007), 139-146.

27. Koleva, B., Benford, S., Hui Ng, K., and Rodden T. A Framework for Tangible User Interfaces. In *Proceedings of the Real World User Interfaces Workshop at the 5th International Symposium on Human–Computer Interaction with Mobile Devices and Services,* Springer London, (2003), 359-369.

28. LaViola, J. J., Feliz, D. A., Keefe, D. F., and Zeleznik, R. C. Hands-free multi-scale navigation in virtual environments. In *Proceedings of the 2001 Symposium on interactive 3D Graphics* I3D '01. ACM, (2001), 9-15.

29. Masliah, M. and Milgram P. Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In *Proc. of CHI 00: ACM Conference on Human Factors in Computing Systems.* ACM, (2000), 25-32.

30. Mazalek, A. and Nitsche, M. Tangible interfaces for real-time 3D virtual environments. In *Proceedings of the international conference on Advances in computer entertainment technology*. ACM, (2007), 155-162.

31. Mine, M. Virtual Environment Interaction Techniques, University of North Carolina at Chapel Hill. Chapel Hill, NC, (1995).

32. Nielsen, R., Fritsch, J., Halskov, K. and Brynskov, M. Out of the box: exploring the richness of children's use of an interactive table. In *Proceedings of the 8th International Conference on Interaction Design and Children*. ACM, (2009), 61-69.

33. Paiva, A., Andersson, G., Hook, K., Mourao, D., Costa, M. and Martinho, C. SenToy in FantasyA: Designing an Affective Sympathetic Interface to a Computer Game. *Personal Ubiquitous Comput.* 6 (5-6), (2002), 378-389.

34. Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J.A., Verratti, T., and Chiselko, G. Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, (2000), 478-485.

35. Patten, J., Ishii, H., Hines, J., and Pangaro, G. Sensetable: a Wireless Object Tracking Platform for Tangible User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, (2001), 253-260.

36. Pedersen, E. W. and Hornbæk, K. mixiTUI: a tangible sequencer for electronic live performances. In *Proceedings of the 3rd international Conference on Tangible and Embedded interaction*. TEI '09. ACM, (2009), 223-230.

37. Pielot, M., Henze, N., and Boll, S. Supporting map-based wayfinding with tactile cues. In *Proceedings of the 11th international Conference on Human-Compute interaction with Mobile Devices and Services*. MobileHCI '09. ACM, (2009), 1-10.

38. Pierce, J.S., Stearns, B.C., and Pausch, R. Voodoo dolls: seamless interaction at multiple scales in virtual environments. In *Proceedings of the 1999 symposium on Interactive 3D graphics* (I3D '99). ACM, (1999), 141-145.

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

40. Quigely, M., Goodrich, M., Beard, R. Semi-Autonomous Human-UAV Interfaces for Fixed-Wing Mini-UAVs. In *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)*. (2004).

41. Santos, C.R., Gros, P., Abel, P., Loisel, D., Trichaud, N., and Paris, J.P. Metaphor-Aware 3D Navigation, In *Proceedings of the IEEE Symposium on Information Visualization.* IEEE, (2000), 155-165.

42. Shaer, O., Leland, N., Calvillo-Gamez, E.H., and Jacob, R.J.K. The TAC paradigm: Specifying Tangible User Interfaces. *Personal and Ubiquitous Computing*, *8*(5), (2004), 359-369.

43. Sharlin, E., Watson, B., Sutphen, S., Liu, L., Lederer, R., and Frazer, J. A tangible user interface for assessing cognitive mapping ability. *Int. J. Hum.-Comput. Stud.* 67, 3 (2009), 269-278.

44. Sharlin, E. et al. On Tangible User Interfaces, Humans and Spatiality. *Personal and Ubiquitous Computin*g. *8*(5), (2004), 338-346.

45. Shimizu, N., Koizumi, N., Sugimoto, M., Nii, H., Sekiguchi, D. and Inami, M. A teddy-bear-based robotic user interface. *Comput. Entertain.* 4, 3, Article 8, (2006).

46. Smets, N. J. J. M., Brake, G.M., Neerincx, M.A., and Lindenberg, J. Effects of mobile map orientation and tactile feedback on navigation speed and situation awareness. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, ACM, (2008), 73-80.

47. Stoakley, R., Conway, M.J., Pausch, R. Virtual Reality on a WIM: Interactive Worlds in Miniature, In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, (1995), 265-272.

48. Tan, D.S., Gergle, D., Scupelli, P.G., and Pausch, R. Physically Large Displays Improve Path Integration in 3D Virtual Navigation Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, (2004), 439-446.

49. Tan, D. S., Robertson, G. G., and Czerwinski, M. 2001. Exploring 3D navigation: combining speed-coupled flying with orbiting. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. ACM, (2001), 418-425.

50. Thulin, S. *There to Hear: Placing Mobile Music.* Master's Thesis. Concordia University. (2009).

51. Ullmer, B. and Ishii, H. Emerging Frameworks for Tangible User Interfaces. in Carrol, J.M. (Ed.), *Human–Computer Interaction in the New Millennium*. Addison-Wesley, (2001), 579–601.

52. Ullmer, B. and Ishii, H. Tangible Bits: Towards Seemless Interfaces Between People, Bits and Atoms. In *Proceedings of the CHI'97*. ACM, (1997), 234–241.

53. Ware, C. and Osborne, S. Exploration and virtual camera control in virtual three dimensional environments. *SIGGRAPH Comput. Graph.* ACM, 24 (2), (1990), 175-183.

APPENDIX A – Pre-Experiment Questionnaire

Pre – Experiment User Background Questionnaire

All responses will remain anonymous and results will be used solely for research purposes.

1. Age: ___________ years

2. Gender: ☐ Male ☐ Female

3. Are you a left-handed person? ☐ Yes ☐ No

4. Do you wear glasses? ☐ Yes ☐ No

5. Are you Color Blind? (i.e.: difficulty in distinguishing certain colors)

☐ No ☐ Yes (Please specify): ____

6. What is your academic level and field of study **OR** Professional Background? (i.e.: Master's of Computer Science)

____

7. Please list any relevant academic/professional background related to 3D Interaction or 3D Virtual Environments:

____

____

8. How would you rate your knowledge / experience with Computers in general?

☐ None ☐ Basic ☐ Intermediate ☐ Advanced

9. How would you rate your knowledge / experience with 3D Interaction or 3D Virtual Environments (VE) on a Computer?

☐ None ☐ Basic ☐ Intermediate ☐ Advanced

10. How would you rate your knowledge / experience with the task of 3D Navigation (games, etc.)?

☐ None ☐ Basic ☐ Intermediate ☐ Advanced

11. Have you ever used any 3D Devices?

☐ No ☐ Yes If yes, please state which one(s) and for what purpose:

____

____

____

12. Have you ever used any Tangible User Interfaces?

☐ No ☐ Yes If yes, please state which one(s) and for what purpose:

____

____

____

13. Have you ever used any Video Game Controllers (ie: Xbox, Playstation or Wii) to move/interact within a 3D VE?

☐ No ☐ Yes If yes, please state which one(s) and for what purpose:

____

____

____