Haptic Issues for Virtual Manipulation

A Dissertation

Presented to

the Faculty of the School of Engineering and Applied Science

at the

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy (Computer Science)

by

Ken Hinckley

December 1996

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy (Computer Science)

Ken Hinckley

This dissertation has been read and approved by the Examining Committee:

Randy Pausch (Dissertation Advisor)

Worthy Martin (Committee Chair)

Dennis Proffitt (Minor Representative)

Andrew S. Grimshaw

John C. Knight

John W. Snell

Accepted for the School of Engineering and Applied Science:

Dean Richard W. Miksad
School of Engineering and Applied Science

December 1996

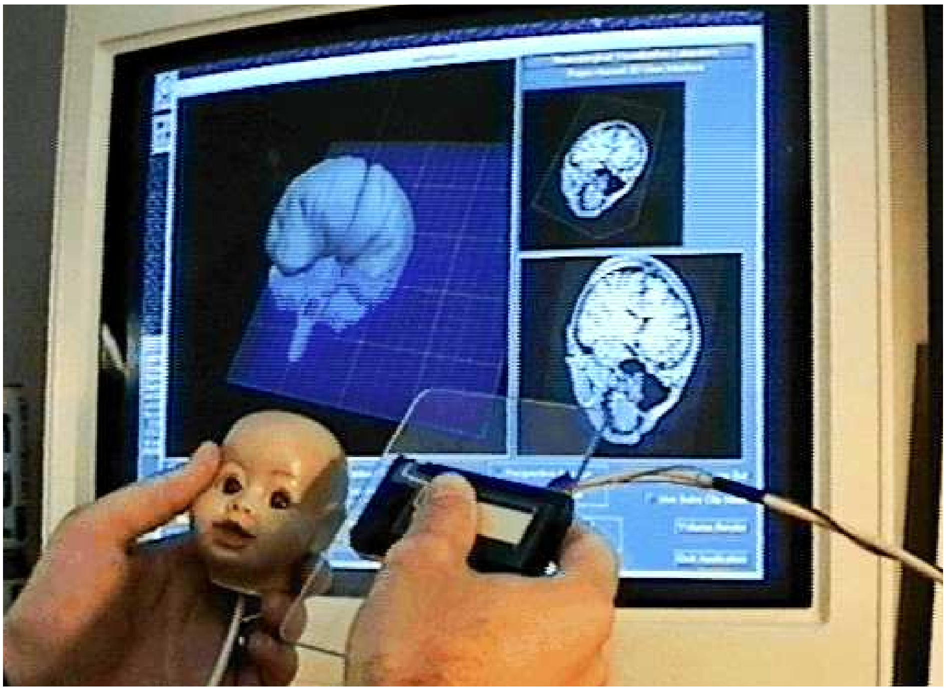
Dedication

I dedicate this work to my father, my mother, and my brother; my grandparents Harold and Mildred, who passed away during the course of my graduate work; and to Kerrie, who has been my anchor for the past two years.

Abstract

The Windows-Icons-Menus-Pointer (WIMP) interface paradigm dominates modern computing systems. Yet these interaction techniques were originally developed for machines that are now 10, 15, or nearly 20 years old. Human-computer interaction currently faces the challenge of getting past this “WIMP plateau” and introducing new techniques which take advantage of the capabilities of today’s computing systems and which more effectively match human capabilities. Two-handed spatial interaction techniques form one possible candidate for the post-WIMP interface in application areas such as scientific visualization, computer aided design, and medical applications.

The literature offers many examples of point design, offering only a description of the thing (what the artifact is) and not the process. But point design only provides a hit-or-miss coverage of the design space and does not tie the multiplicity of efforts into a common understanding of fundamental issues. To get past the WIMP plateau, we need to understand the nature of human-computer interaction as well as the underlying human capabilities.



My research contributes a working system which has undergone extensive informal usability testing in the context of real domain experts doing real work, and it also presents the results of experimental evaluations which illustrate human behavioral principles. Together, these approaches make a decisive statement that using both hands for virtual manipulation can result in improved user productivity.

I contribute to the field by:

(1) Showing that virtual manipulation needs to study the feel of the interface, and not just the graphical look of the interface;

(2) Applying virtual manipulation technology to a volume visualization application which has been well received by neurosurgeons;

(3) Demonstrating two-handed virtual manipulation techniques which take advantage of the highly developed motor skill of both hands;

(4) Contributing basic knowledge about how two hands are used;

(5) Showing that two hands are not just faster than one hand, but that two hands together provide information which one hand alone cannot, and can change how users think about a task;

and finally (6) Providing an overall case study for an interdisciplinary approach to the design and evaluation of new human-computer interaction techniques.

Acknowledgments

My research has necessarily been an interdisciplinary effort, and as such, there are many persons to whom I am indebted. In particular, I owe great thanks to the Department of Neurological Surgery for supporting the work and participating in the design effort.

I would like to thank the members of my committee, and in particular, Randy Pausch, Dennis Proffitt, and John Snell, who have all worked closely with me over the past few years.

I also owe thanks to Joe Tullio for assistance with experiments, design, and implementation of testing software, and James Patten, for assistance with running experimental subjects.

I would like to acknowledge external colleagues including Yves Guiard, currently at the Cambridge Applied Psychology Unit, Bill Buxton, of the University of Toronto and Alias | Wavefront, and Rob Jacob, at Tufts University, for enlightening discussions about some of the issues raised by my work.

I have been assisted by the friendly disposition and skill of machinists Bob Frazier and Bob Bryant who helped to design and build the physical devices and apparatus which made the interface tools and the experimental equipment work.

I would like to thank the present members, graduates, and associates of the Neurosurgical Visualization Lab, including Hunter Downs, Sean Graves, Adrian Filipi-Martin, Dave Schlesinger, Dave Moore, Michelle Plantec, Minyan Shi, Greg Harrington, Delia McGarry, Bill Katz, Ted Jackson, Will McClennan, Marc Pilipuf, Jim Bertolina, Dr. Neal Kassell, Dr. Dheerendra Prasad, and Dr. Mark Quigg. Each of these individuals participated in design discussions, assisted implementation efforts, participated in pilot experiments, or provided general support, assistance, and advice in the course of my dissertation work.

I would like to thank the general support and many valuable discussions provided by members of the user interface group, including current members Steve Audia, Tommy Burnette, Matt Conway, Kevin Christiansen, Dennis Cosgrove, Beth Mallory, Kristen Monkaitis, Jeff Pierce, Joe Shochet, Brian Stearns, Chris Sturgill, John Viega, and George Williams; as well as the many graduates with whom I have worked, including Robert Allen, Jonathan Ashton, Brian Cadieux, AC Capehart, Roddy Collins, Thomas Crea, Rob DeLine, Jim Durbin, Rich Gossweiler, Simeon Fitch, Drew Kessler, Shuichi Koga, Jan Hoenisch, Chris Long, Steve Miale, Kimberly Passarella Jones, Anne Shackelford, Richard Stoakley, Jeff White, and Scott Witherell.

I also owe thanks to colleagues in the Biomedical Engineering, Radiology, and other departments who helped along the way, including Jim Brookeman, Sam Dwyer, Linda Young, and Doug Kraus.

I would like to thank members of Dennis Proffit's Perception Lab who have assisted my efforts, shared valuable experimental space with me, and offered advice on working with subjects, including Sarah Creem, Steve Jacquot, Jane Joseph, MJ Wraga, and Tyrone Yang.

Finally, I thank the many students who participated in my experiments.

Contents

3 System Description 40

List of Figures

1 Introduction **1**

2 Related Work **14**

3 System Description **40**

"A major challenge of the post-WIMP interface is to find and characterize appropriate mappings from high degree-of-freedom input devices to high degree-of-freedom tasks."

Stu Card

Chapter 1

Introduction

This thesis originated with a user interface I designed for three-dimensional neurosurgical visualization (*fig. 1.1*). The interface design was well received both by neurosurgeons and by the community of human-computer interface (HCI) designers. By itself, the interface contributes a point design which demonstrates techniques that allow neurosurgeons to effectively view and cross-section volumetric data. A primary goal for this dissertation is to move beyond point design and to introduce some careful scientific measurement of behavioral principles which were suggested by the original system implementation. Based on the synergy of (1) an interface design which has undergone extensive informal usability testing and (2) formal experimental evaluation, I make some general points about interface design and human behavior, so that some of the lessons learned in the neurosurgery application can be applied to future user interface designs.

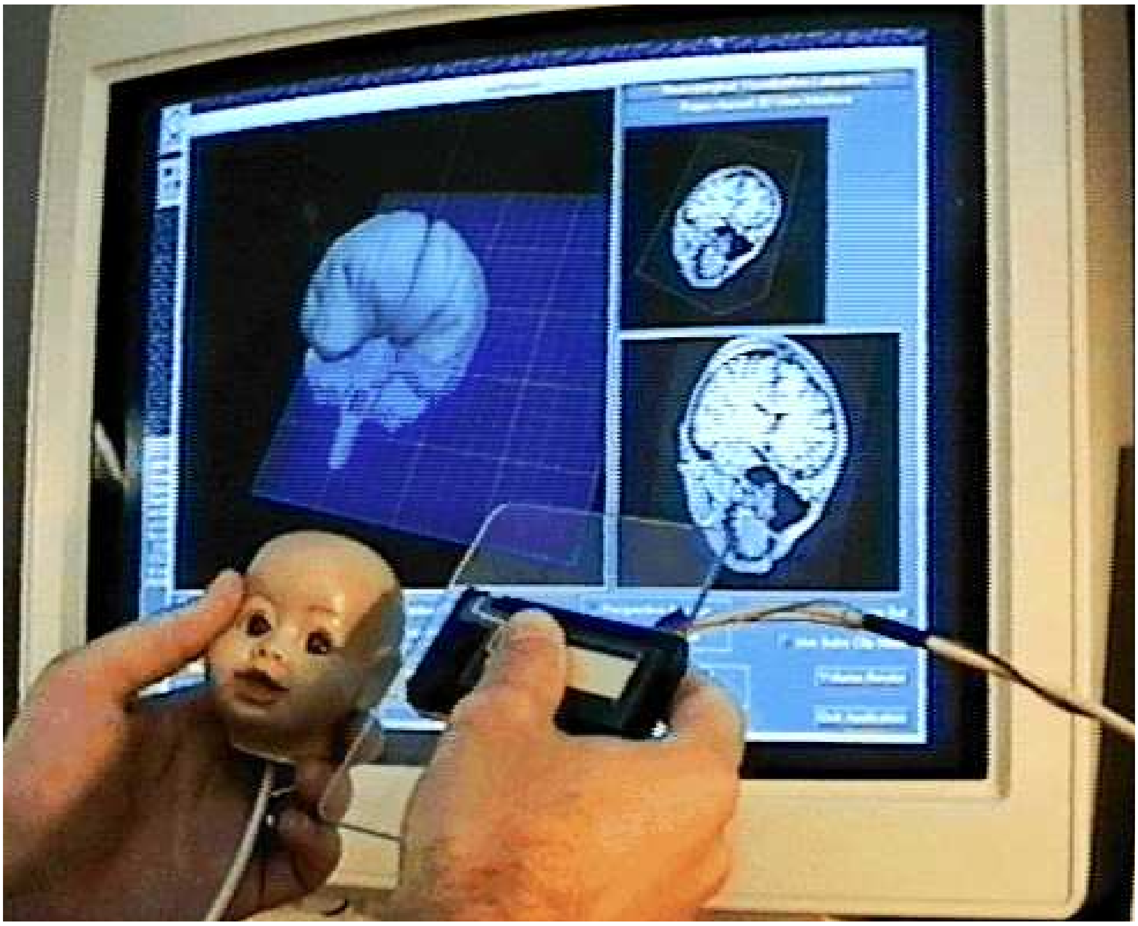


Figure 1.1 The props-based interface for neurosurgical visualization.

1.1 Problem motivation

Figure 1.2 illustrates plateaus in the quality and facility of user interaction over the history of computing systems. When computers were first introduced, very few people had the knowledge and ability to operate them. Computers were extremely expensive, so the most important concern was to optimize use of computer time. Users submitted “batch jobs” in advance which would keep the computer constantly busy; it did not matter if both the turn-around time, and the resulting user productivity, were very poor. The introduction of interactive teletype and command-line interfaces was an improvement, but most people still did not have the expertise to perform tasks using a computer.

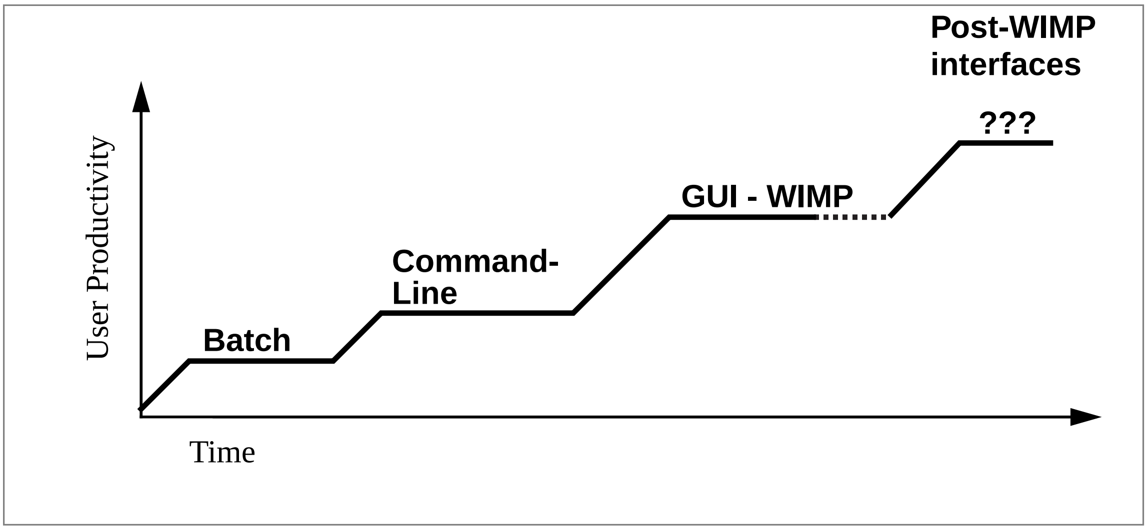


Figure 1.2 Historical plateaus in quality and facility of user interaction.

By the late 1970’s, with computer costs continuing to fall, the cost of time that users spent operating a computer became comparable to the cost of the computer itself. This led to a new philosophy for interface design, first embodied by the Xerox Star personal computer [9]: the Star’s designers realized that it was more important to make effective use of the user’s time than to optimize processor cycles. In the process of attempting to make users more productive, the Star became the first computer to include a user interface based on a bitmapped display and a mouse. The Star inspired the designers of the Apple Lisa [9], and later the Apple Macintosh computer, and led to many of the current conventions for graphical user interfaces. For the first time, these innovations made it possible for a large segment of the population, without extensive knowledge of how computers work, to use personal computers as tools.

The current paradigm for graphical user interfaces (GUI’s) has been dubbed the “WIMP” (Windows, Icons, Menus, and Pointer) interface. Many WIMP graphical interaction techniques were originally designed by Xerox and Apple for low-powered processors with small black-and-white displays. Yet as computing technology becomes ubiquitous and the capabilities of processors, displays, and input devices continue to grow, the limitations of the WIMP interface paradigm become increasingly apparent. To get past this “WIMP plateau,” devising new interface metaphors will not be enough. We need to broaden the input capabilities of computers and improve the sensitivity of our interface designs to the rich set of human capabilities and skills.

There are many devices, displays, and interaction techniques which are candidates for the post-WIMP interface. Candidates identified by Nielsen [123] include virtual realities, sound and speech, pen and gesture recognition, animation and multimedia, limited artificial intelligence (in the form of so-called “interface agents”), and highly portable computers. Weiser [181] has proposed the *ubiquitous computing* paradigm, which suggests that networked computers will increasingly become integrated with ordinary implements and that computers will be embedded everywhere in the user’s environment.

Demonstrations of point designs will not be sufficient to take advantage of advanced interaction techniques. Departing with tradition is expensive and risky, and while a demonstration might be compelling, it cannot tell an interface designer how to generalize a proposed interface technique to a new situation. A demonstration also cannot answer questions about when a new interface technique should be used, or when it may (or may not) have measurable advantages for the user. One must perform careful scientific evaluation of interaction techniques to understand how to use a proposed technique as well as why a proposed technique may result in improved performance for the user. Without such knowledge, and interaction technique essentially must be re-invented every time a designer attempts to use it.

Perhaps the best-known archetype for how one should approach formal user interface evaluations is provided by Card’s experimental comparison of the mouse and other input devices [39]. Card formed mathematical models of each device and tested these models against observed performance data. Card found that the mouse was effectively modelled by Fitts's Law, which predicts movement time based on the amplitude of a movement and width of the target area. Furthermore, Card showed that the same model, with roughly the same parameters, modelled both movement with the mouse and movement with the hand alone. This suggests that designing an input device which allows one to point at targets more quickly than one can point using the mouse would be difficult to achieve.

Summarizing the philosophy of his approach, Card has written:

> [User technology]... must include a technical understanding of the user himself and of the nature of human-computer interaction. This latter part, the scientific base of user technology, is necessary in order to understand why interaction techniques are (or are not) successful, to help us invent new techniques, and to pave the way for machines that aid humans in performing significant intellectual tasks. [38]

My dissertation seeks to follow Card's general approach: by formulating hypotheses and subjecting hypotheses to experimental tests, I demonstrate some fundamental mechanisms of human behavior and general design principles which suggest new design possibilities for the post-WIMP interface.

1.2 Virtual manipulation

Three-dimensional (3D) interfaces and interaction techniques form one candidate for the post-WIMP interface, especially in application areas such as computer-aided design (CAD), architectural design, scientific visualization, and medicine. A central problem for three-dimensional interaction is that of virtual manipulation, which concerns the general problem of grasping and manipulating computer-generated 3D virtual objects. Virtual manipulation includes tasks such as specifying a viewpoint, planning a navigation path, cross-sectioning an object, or selecting an object for further operations.

Virtual manipulation poses a difficult dilemma: one wants virtual objects to violate reality so that one can do things that are not possible to do in the real world, yet one also wants virtual objects to adhere to reality so that the human operator can understand what to do and how to do it. The interface design challenge is find ways that real and virtual objects and behaviors can be mixed to produce something better than either alone can achieve [156]; and part of this challenge is to discover interaction techniques that do not necessarily behave like the real world, yet nonetheless seem natural. This leads to a key point: to design interaction techniques which meet these criteria, short of taking wild guesses in the dark, *the interface designer needs to understand the human*.

1.3 Passive haptic issues

This thesis mainly deals with what a psychologist would consider proprioception (the reception of stimuli produced within the body) and kinesthesis (the sense of bodily movements and tensions). I focus not on the feel of objects, but rather on how users move their hands to explore and manipulate virtual objects, and how users are sensitive to changes in the relative distances between the hands and the body. But, as psychologist J. J. Gibson has pointed out [63], even the terms proprioception and kinesthesis are quite imprecise. The terms include diverse sensations such as sense of movement of the body, joint angles, upright posture and body equilibrium, forces exerted by or on the body, orientation relative to gravity, as well as sense of linear and angular accelerations [63]. Nonetheless, the hypotheses raised by this thesis do not depend upon whether or not these diverse sensations comprise separate sensory systems or one super-ordinate sense.

Haptic comes from the Greek word *haptesthai*, which means to touch. In human-computer interaction, and for the purposes of this thesis, the term haptic encompasses a broad set of related sensations including proprioception and kinesthesis. By describing all of these sensations as “haptic” my intent is not to suggest that these are all facets of a unitary sense, but rather to suggest that together these sensations constitute the *feel* of an interface and present a set of issues which graphical user interfaces and prior efforts in virtual manipulation have often neglected [31][32]. Haptic issues for virtual manipulation involve issues of manual input and how they influence the feel of the human-computer dialog, including issues such as the muscle groups used, muscle tension, the shape and mechanical behavior of input devices, coordination of the hands and other limbs, and human sensitivity to the articulation of the body and the position of its members relative to one another.

Haptic issues include the related topic of active haptic feedback using force-returning armatures or exoskeletons. These devices allow one to feel the shape and texture of a computer-generated object. The Phantom [147], for example, is an armature with a stylus at its distal end. The user can feel a virtual object through the stylus. When the user holds and moves the stylus, the Phantom knows the position of the tip of the stylus and can generate forces which resist the user’s motion when a virtual surface is encountered. Although there are some promising applications for active haptic feedback [21], the technology is still relatively primitive and expensive.

I primarily focus on the information which the user’s own manual movements provide for feedback during manipulation-- what one might call *passive haptic feedback* because the computer can’t directly control or alter these sensations.[1] But the design of the user interface can and should take passive haptic issues into account to provide a human-computer dialogue which is natural and which takes advantage these innate human capabilities and sensations. In essence, my approach focuses not on a particular technology, but rather on the capabilities of the human participant.

---

[1] An ecological psychologist (one who studies perception in the natural environment, emphasizing the moving, actively engaged observer) would think of these as “active” feedback sensations, because the user seeks out and explores the stimuli rather than passively receiving them [75]. From the standpoint of feedback which is under direct control of the computer, however, the information is static or passive, even if the user is active in exploring the passive stimulus.

1.4 Humans have two hands

The cooperative action of the two hands and the proprioceptive information which they provide to the user comprise aspects of passive haptic feedback which have been particularly neglected in virtual manipulation. Humans not only have two hands, but they also have highly developed manual skill with *both* hands. Then why doesn't contemporary computer interface design reflect this? When the nonpreferred hand is used at all, it is usually banished to the occasional keyboard button press. In the traditional "WIMP" interface, even with the preferred hand, all of the human's manual skills are boiled down to moving a single speck on the screen with a mouse. We can do better than this.

The above motivation takes an admittedly informal approach, but Guiard's research on the cooperative action of the two hands working together [67], which Guiard terms *bimanual action*, provides a scientific basis for understanding how humans use two hands. For the present discussion, I will primarily rely on intuitive examples to illustrate Guiard's work; chapter 2, "Related Work," treats Guiard's theoretical contributions in greater depth.

A fundamental observation provided by Guiard is that, in the set of human manipulative tasks, purely unimanual acts are by far a minority, while bimanual acts are commonplace. Manipulative motions such as dealing cards, playing a stringed musical instrument, threading a needle, sweeping, shovelling, striking a match, using scissors, unscrewing a jar, and swinging a golf club all involve both hands. Even writing on a piece of paper with a pen, which has sometimes been mistakenly classified as a unimanual behavior [67], is demonstrably two-handed: Guiard has shown that the handwriting speed of adults is reduced by about 20% when the nonpreferred hand cannot help to manipulate the page [67].

Some of these tasks can be performed unimanually if necessary due to injury or disability, but the normal human behavior is to use both hands. Threading a needle is an interesting example: logic tells us that it should be easier to thread a needle if it is held still in a clamp, and one just has to hold the thread in the preferred hand to guide it through the eye of the needle. Holding the needle in the “unsteady” and “weak” nonpreferred hand introduces a second moving thing and should make the task more difficult. But using both hands instead makes the task easier: when faced with a stationary needle, the first instinct is to grab it with the nonpreferred hand. The nonpreferred hand acts as a dynamic and mobile clamp which can *skillfully* coordinate its action with the requirements of the preferred hand.

Virtual manipulation is a particularly promising application area for two handed interaction. There is not yet an established “standard interface” such as the WIMP interface’s mouse-and-keyboard paradigm which dominates the marketplace. Furthermore people naturally use both hands to indicate spatial relationships and to talk about manipulations in space [76], and as Guiard has argued, the vast majority of real-world manipulative tasks involve both hands [67]. Finally, virtual manipulation presents tasks with many degrees-of-freedom; using both hands can potentially allow users to control these many degrees-of-freedom in a way that seems natural and takes advantage of existing motor skills.

1.5 Thesis statement

Behavioral principles (as proposed by Guiard) suggest that humans combine the action of the hands through an asymmetric hierarchical division of labor. I assert that a system must respect the subtleties of the application of these principles; the mapping of hand motion to virtual object motion is not obvious.

1.6 Contributions and overview

1.6.1 Interdisciplinary approach

Human-computer interaction is inherently interdisciplinary; some would argue that computer science itself is inherently interdisciplinary, with (as Fred Brooks has put it) the computer scientist taking on the role of a toolsmith [20] who collaborates with others to develop useful tools. In my research, the collaboration of neurosurgeons, computer scientists, and psychologists has been essential. The input of neurosurgery domain experts validates the computer science research: rather than addressing a “toy problem,” the work has meaning and application for something real. Working with neurosurgeons has also forced me to address problems I otherwise might have ignored, such as making the physical input devices look professional and “polished,” or carefully optimizing the program code to allow real-time, interactive visualization of volumetric cross-sectional information. The collaboration with psychologists has forced me to think carefully about the underlying behavioral issues, and has allowed me to pursue careful scientific evaluations of these issues. And of course, without a computer science systems-building approach, there would be no tool for neurosurgeons to use, and no design issues to validate with experiments. In short, an interdisciplinary approach has enabled a decisive contribution to the virtual manipulation and two-handed interaction research fields. Without any one part of this collaboration, the work would not be as convincing. Taken as a whole, my thesis contributes a case study for interdisciplinary research methodology in human-computer interaction.

1.6.2 Revisiting haptic issues

The sense of touch is crucial to direct skilled manipulative action. With all the emphasis on *graphical* user interfaces, it is easy to forget that physical contact with the input device(s) is the basis of manipulative action. While much research has focused on the visual aspects of 3D computer graphics for virtual reality, many of the haptic issues of virtual manipulation have gone unrecognized. The phrase *Look and Feel* is frequently used to capture the style of both the visual appearance (the output) and the physical interaction (the input) afforded by an interface-- but as Buxton has argued, to reflect the relative investment of effort on the output side versus the input side, the phrase should perhaps instead be written [31]:

Look and Feel.

At the broadest level, in the process of revisiting haptic issues, I have demonstrated that facile virtual manipulation requires studying the feel of the interface as well as the look of the interface.

1.6.3 Application to neurosurgical visualization

I have worked with the Department of Neurological Surgery to develop a 3D volume visualization interface for neurosurgical planning and visualization (*fig. 1.1*). The 3D user interface is based on the two-handed physical manipulation of hand-held tools, or *props*, in free space. These user interface props facilitate transfer of the user's skills for manipulating tools with two hands to the operation of a user interface for visualizing 3D medical images, without need for training.

From the user's perspective, the interface is analogous to holding a miniature head in one hand which can be "sliced open" or "pointed to" using a cross-sectioning plane or a stylus tool, respectively, held in the other hand. Cross-sectioning a 3D volume, for example, simply requires the surgeon to hold a plastic plate (held in the preferred hand) up to the miniature head (held in the nonpreferred hand) to demonstrate the desired cross-section.

1.6.4 Two-handed virtual manipulation techniques

The interface demonstrates interaction techniques which use the nonpreferred hand as a dynamic frame of reference. Unlike a clamp (whether physical or virtual), the nonpreferred hand provides *mobile* stabilization. An important design principle is for the interface to preserve the mobile, dynamic role of the nonpreferred hand as a reference. The nonpreferred hand adjusts to and cooperates with the action of the preferred hand, allowing users to restrict the necessary hand motion to a small working volume. Informal usability evaluations with hundreds of test users confirms that two-handed virtual manipulation can be effective and easy to learn when designed appropriately.

1.6.5 Basic knowledge about two hands

I contribute formal experimental data which confirms Guiard’s suggestion that the hands have specialized, asymmetric roles. In particular, the thesis includes the first experimental data which suggests that the preferred hand operates relative to the frame-of-reference of the nonpreferred hand. I also show that the advantages of hand specialization are most significant for “high degree of difficulty” tasks.

1.6.6 Two hands are not just faster than one hand

Using two hands provides more than just a time savings over one-handed manipulation. Two hands together provide the user with information which one hand alone cannot. Furthermore, a simultaneous two-handed task is not the same thing as the serial combination of the sub-tasks controlled by each hand. The simultaneous task allows for hierarchical specialization of the hands. It also provides the potential to impact performance at the cognitive level: it can change how users *think* about a task. Since the user can potentially integrate subtasks controlled by each hand without an explicit cost to switch between subtasks, this encourages exploration of the task solution space.

1.7 Organization

This thesis is organized as follows:

Chapter two: *Related Work* describes related work in virtual manipulation, multimodal input, and theory and experiments for two-handed interaction.

Chapter three: *System Description* discusses interaction techniques, implementation issues, and notes on user acceptance for a three-dimensional interface which I designed for neurosurgical visualization.

Chapter four: *Design Issues in Spatial Input* presents a synthesis of design issues drawn from the literature on virtual manipulation, tempered by my interface design experience with the three-dimensional interface for neurosurgical visualization.

Chapter five: *Research Methodology* discusses the issues raised by evaluation of user interfaces and working with domain experts. This chapter also discusses the rationale and the process for the formal experimentation approach used in my thesis work.

Chapter six: *Usability Analysis of 3D Rotation Techniques* presents a formal analysis of techniques for three-dimensional rotation of virtual objects.

Chapter seven: *Issues in Bimanual Coordination* describes experiments which analyze manipulation of physical objects with both hands.

Chapter eight: *The Bimanual Frame-of-Reference* presents an experiment which demonstrates how two hands together provide the user with information about the environment which one hand alone cannot, and also demonstrates that physical articulation of a task can influence how users think about that task.

Chapter nine: *Conclusions* summarizes the main results of my thesis work and proposes some potential areas for future research.

"Experience keeps a dear school, but fools will learn in no other."

Benjamin Franklin, Poor Richard's Almanac, 1757

Chapter 2

Related Work

2.1 Introduction

This chapter provides an illustrated tour of systems and techniques for manipulation of virtual objects and navigation in virtual environments. Although the overall emphasis is on two-handed, three-dimensional approaches, I have attempted to provide a sampling of 2D and 3D approaches, as well as one-handed and two-handed techniques, to provide some overall context for my own work. Since experimental work is an emphasis of this dissertation, this review also includes experiments analyzing issues in two-handed interaction.

2.2 Two-dimensional approaches for 3D manipulation

A number of interaction techniques have been proposed which allow 2D cursor motion to be mapped to three-dimensional translation or rotation. These techniques usually consist of a combination of mappings of mouse motion to virtual object motion and a mechanism for selection of modes to alter which degrees-of-freedom the 2D device is currently controlling.

2.2.1 3D Widgets

3D Widgets are graphical elements directly associated with objects in the scene. For example, an object might be augmented with handles to allow rotation of the object (*fig. 2.1, left*), or with shadows to allow constrained translation of the object (*fig. 2.1, right*). The 3D Widgets disguise separate manipulation modes as spatial modes; that is, moving the mouse and dragging a new handle invokes a new constraint mode. Of course, the widgets must be added to the scene as extra components, which may clutter the scene or obscure the object(s) being manipulated.

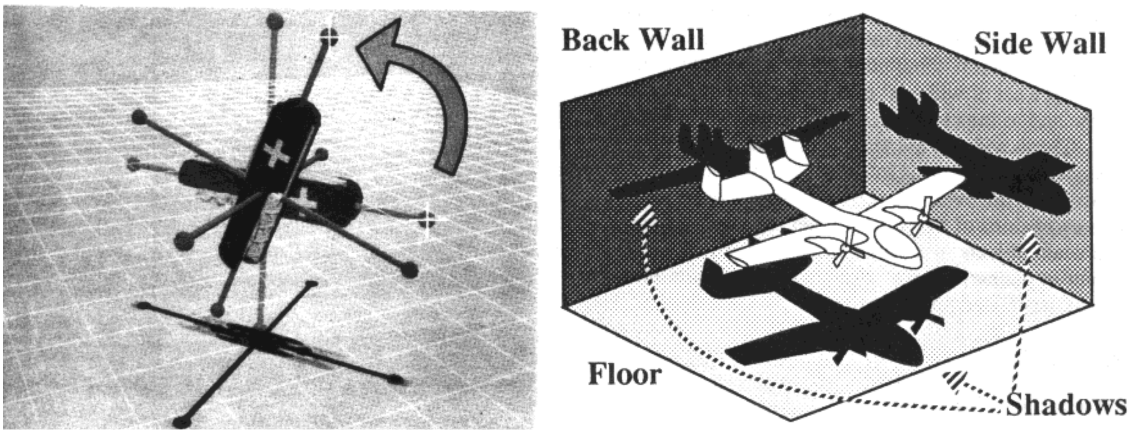


Figure 2.1 Examples of "3D Widgets" (taken from [43] and [77])[1].

Other, more abstract, varieties of 3D widgets include the Arcball [153], Chen's Virtual Sphere [40], and the Evans & Tanner [51] rotation controllers. The Arcball and Virtual Sphere both use the metaphor of rotating an imaginary sphere; the Evans & Tanner technique, while not explicitly using this metaphor, was a predecessor of these techniques which exhibits performance similar to the Virtual Sphere [40]. These techniques will be discussed in further detail in the context of Chapter 6, "Usability Analysis of 3D Rotation

---

[1] Please note that *every* figure in this chapter has been reproduced from another source. Hereafter, a reference number in each figure caption will indicated the source from which the figure has been reproduced or adapted.

Techniques.” As an additional example, Nielson and Olsen’s *triad mouse* technique allows specification of 3D points by mapping mouse motion onto the projected principle axes of an object (*fig. 2.2*). The resulting points can be mapped to translation, rotation, and scaling in three dimensions. This technique works well at three-quarter perspective views (as seen in figure 2.2), but does not work as well when the projection of two axes fall close to one another.

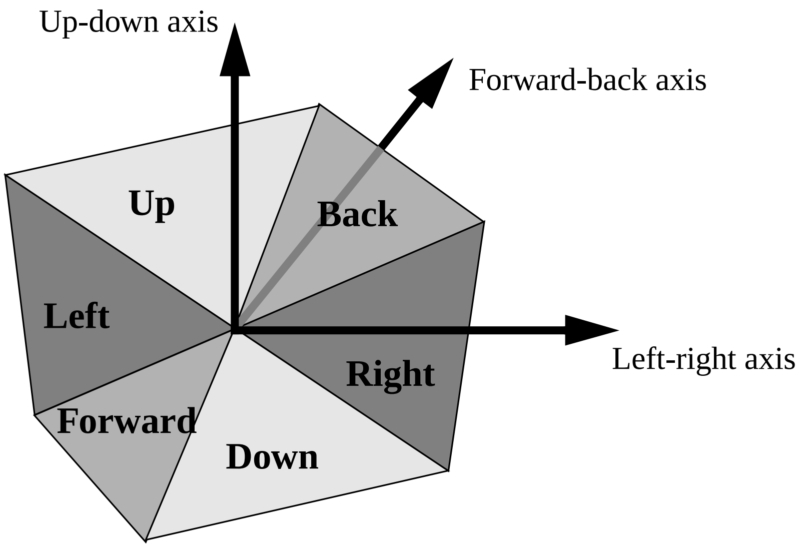


Figure 2.2 Cursor movement zones for the triad mouse technique [125].

2.2.2 SKETCH

The SKETCH system [186] provides a gesture-based 2D interface for generating approximate 3D CAD models. SKETCH’s gesture-based interface consists of a small vocabulary of axis-aligned strokes, mouse clicks, and modifier keys which can be composed following some simple rules. SKETCH also makes clever use of heuristics, such as making a stroke down from an existing object or sketching a shadow below an object (*fig. 2.3*), to infer the 3D location of objects from a 2D sketch. In this sense, SKETCH makes use of the dynamic process of sketching itself to infer information, rather than relying only on static 2D images, as is usually the case in vision research.

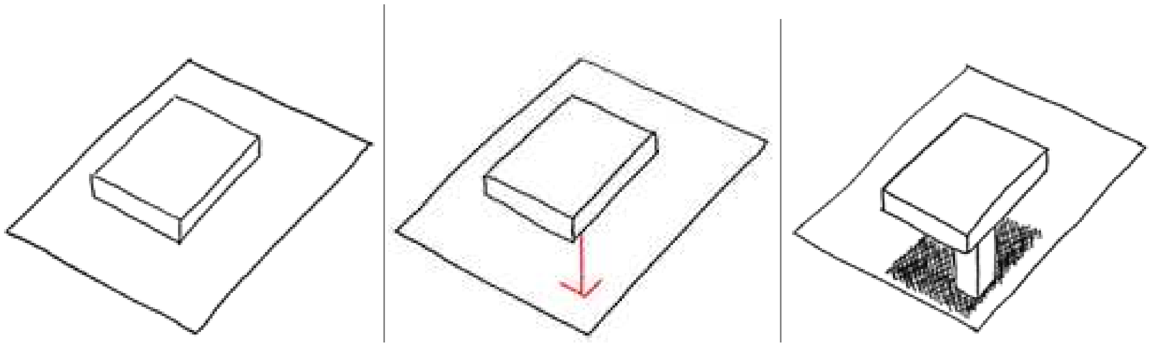


Figure 2.3 SKETCH uses heuristics to infer 3D placement [186].

2.3 Using the hand itself for input

2.3.1 Gloves and Gesture for virtual manipulation

Sturman [165] and Zimmerman [189] describe instrumented gloves which can detect the flex angles of the hand and finger joints (*fig. 2.4*). These gloves are typically used in conjunction with a six degree-of-freedom magnetic sensor attached to the wrist, providing a combined input stream of hand position and finger flexion. The development of techniques to reliably detect gestures, such as making a fist, an open palm, or pointing, from multiple streams of input data, is still an active area of research (for example, see Wexelblat [183]). Other research, such as Baudel's Charade system [10], has focused on techniques for designing gesture command languages which are easy for users to learn.

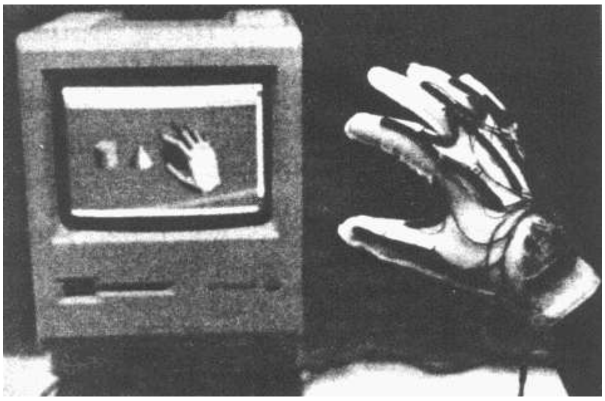


Figure 2.4 Glove used by Zimmerman [189].

Recently introduced gloves, called *chord gloves* or *pinch gloves* [52], detect contact between the fingertips, such as when pinching the thumb and forefinger together. Detecting contact is much more reliable than detecting gestures through finger flexion or hand movement. The gloves cost significantly less without the flexion sensors, making it more practical to use two gloves in one application, as demonstrated by the Polyshop system [1] (*fig. 2.5*), as well as Multigen's SmartScene application [121]. But most importantly, touching one's fingers together is much easier for users to learn than a gesture language, so fingertip contact is an effective means to provide frequently used commands.

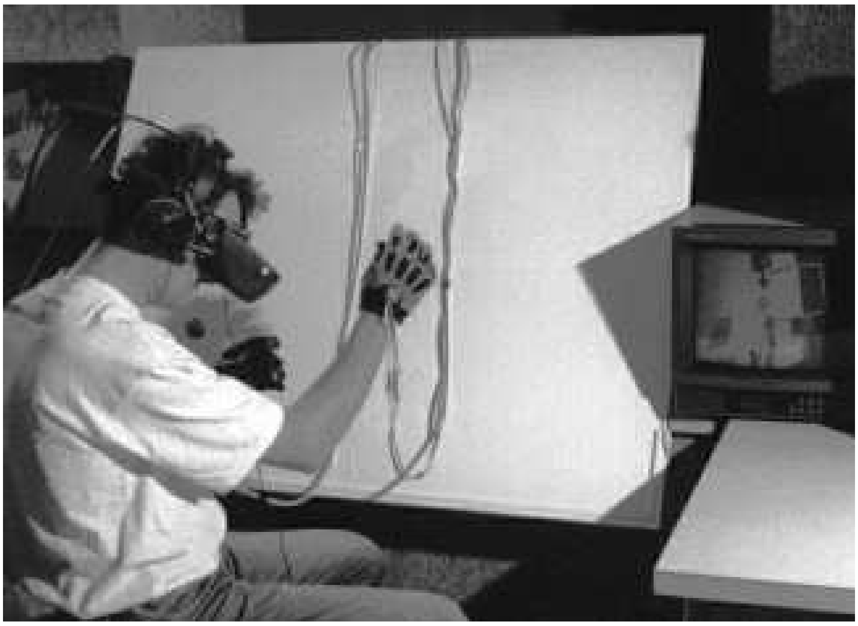


Figure 2.5 Polyshop employs two-handed interaction with chord gloves [1].

2.3.1.1 The Virtual Wind Tunnel

Bryson has implemented a virtual wind tunnel which allows scientists to visualize properties of airflow and pressure in computational fluid dynamics volumetric data sets. These enormous data sets yield many computational challenges [25]. Interaction with the data is achieved with a glove for pointing and issuing gestural commands plus a boom display for viewing the environment (*fig. 2.6*). The boom display is used instead of a head-mounted display to avoid problems with encumbrance. The boom also makes it easy for a pair of scientists to share views of the data.

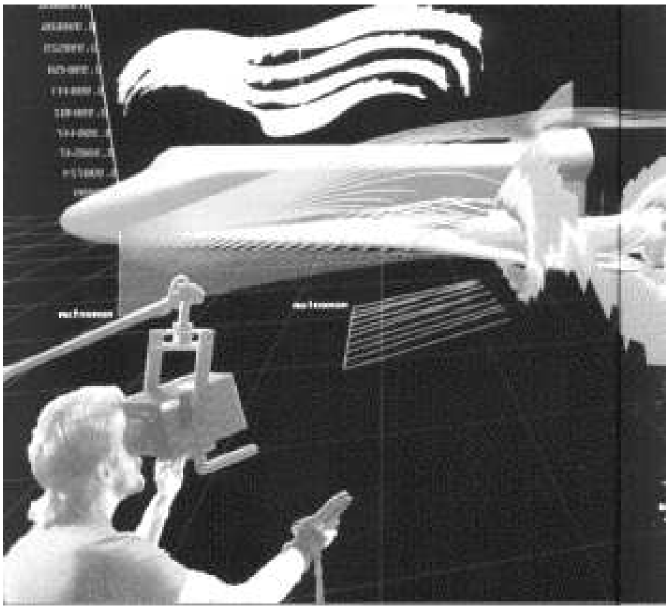


Figure 2.6 Using a boom display and glove with the Virtual Wind Tunnel [25].

2.3.2 Electric field sensing

Zimmerman [190] describes a technique for detecting hand or body position without any encumbering input device at all. The technique works by sensing the capacitance associated with the human body. The body acts as a ground for a low frequency energy field, displacing a current which is functionally related to position within the field. The technique can also be used in the opposite manner, using the body as a short-range electric field transmitter: this has applications in mobile computing, for example. Since the technique relies only on electrical components, it has the potential to become a ubiquitous, cheap technology available as a PC plug-in board.

2.3.3 VIDEODESK and VIDEOPLACE

Kreuger’s VIDEODESK system [104] uses video cameras and image processing to track 2D hand position and to detect image features such as hand, finger, and thumb orientation. This approach reads to a rich vocabulary of simple, self-revealing gestures such as pointing, dragging, or pinching which do not require any explicit input device [103]. For example, the index finger and thumb of both hands can be used to simultaneously manipulate four control points along a spline curve (*fig. 2.7, left*) or the finger tips can be used to sweep out a rectangle (*fig. 2.7, right*).

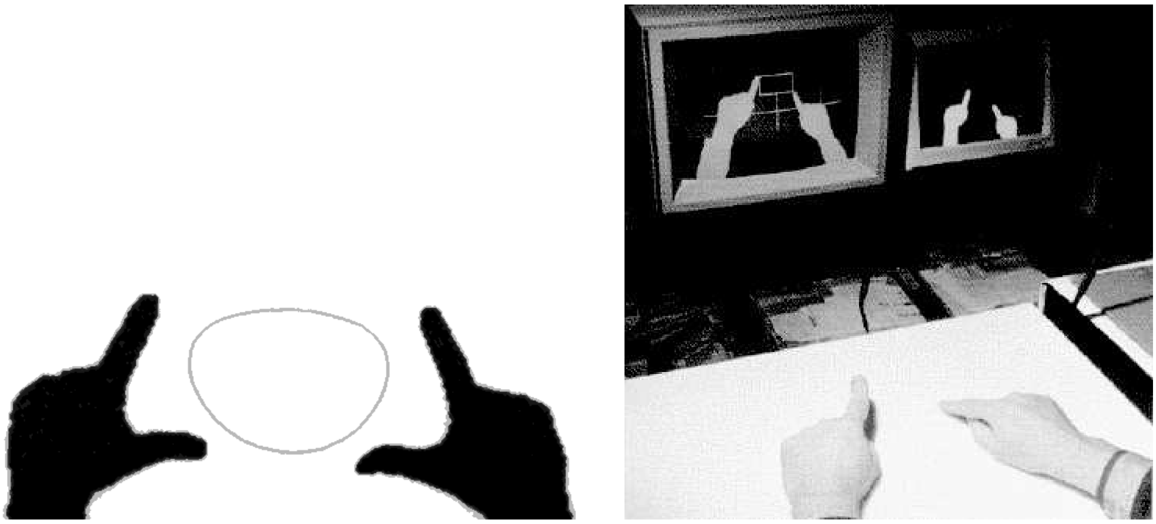


Figure 2.7 Example VIDEODESK applications using two hands [104].

Kreuger has also explored two-dimensional camera-based tracking of the entire body in the VIDEOPLACE environment [104][105], and he has linked this environment with the VIDEODESK. For example, the VIDEODESK participant can pick up the image of the VIDEOPLACE participant's body, and the VIDEOPLACE participant can then perform gymnastics on the other person's fingers [104]. Although these examples are somewhat whimsical, they demonstrate that the medium has many design possibilities.

2.3.4 Hands and voice: multimodal input

2.3.4.1 Hauptmann's behavioral study

Hauptmann [76] reports a "Wizard-of-Oz" study of voice and gesture techniques for 3D manipulations such as translation, rotation, and scaling of a single object. In a Wizard-of-Oz study, instead of implementing an entire system, a “man behind the curtain” plays the role of the computer. Such studies are often used to classify the types of responses that users will naturally generate, so that designers will have a better sense of the issues which an actual system implementation must handle. Hauptmann uses the Wizard-of-Oz strategy to explore the possibilities for voice and gesture, together and in isolation. Hauptmann found that the vocabulary of speech and gestural commands was surprisingly compact. Hauptmann’s test subjects strongly preferred to use simultaneous voice and gesture. Many subjects also spontaneously used both hands, particularly when indicating scaling operations.

2.3.4.2 Put-That-There

Bolt’s “Put-That-There” system [16] provides a compelling demonstration of multimodal voice input plus gesture. Put-That-There allowed users to manipulate objects on a large-screen graphics display showing (for example) a map. The user could point at objects on the map, and speak commands to modify them.

The key insight of the Put-That-There system is that either modality alone is quite limited. Voice recognizers (even today) are nowhere near 100% reliable, especially when the syntax of possible commands is complex. So describing everything verbally is both tedious and error-prone. Gesture (pointing) alone allows selection and dragging, but not much else. But together, the modalities provide a rich medium. Instead of saying “Put the orange square to the right of the blue triangle,” the user can speak commands such as “Put that... to the right of that.” The recognizer needs to know only one object name: “that.” The pronoun is disambiguated using the pointing gesture. This simplifies the syntax of voice commands, and also frees the user from having to remember the names of objects.

2.3.4.3 Two hands and voice

Bolt and Herranz [17] report an implementation of some of the techniques explored by Hauptmann's study [76], among others, using voice input, two-handed gesture, and eye gaze information. The eye gaze information is used to select objects. Voice commands help to disambiguate the eye gaze and gesture information. Two-handed techniques include dual-hand rotation about principle axes of an object as well as two-handed relative placement of objects, where one hand acts as a reference, and the other hand indicates the relationship of a second object to the first hand.

Weimer and Ganapathy [180] discuss voice and gesture input for a 3D free-form surface modeling system. Weimer and Ganapathy initially based the interface purely on gestural information, but added voice for three reasons: (1) people tend to use gestures to augment speech; (2) spoken vocabulary has a more standard interpretation than gestures; and (3) hand gesturing and speech complement one another. Voice is used for navigating through commands, while hand gestures provide shape information. Weimer and Ganathapy report that there was "a dramatic improvement in the interface after speech recognition was added" [180].

2.4 One-handed spatial interaction techniques

Hand-held input devices afford different interaction techniques than systems based on glove and gesture-based approaches. This section focuses on techniques in free space that utilize three-dimensional hand-held input devices.

2.4.1 Schmandt's stereoscopic workspace

Schmandt [142] describes a stereoscopic workspace which allows the user to see his or her hand in the same apparent volume as computer-generated 3D graphics, providing direct correspondence between the 3D input and the 3D output volumes (*fig. 2.8*). A hand-held wand is seen through a half-silvered mirror, upon which the computer graphics are projected. A white spot on the wand itself can be seen through the half-silvered mirror and acts as a real-world cursor. Despite the corresponding input and output volumes, the semi-transparent mirror used in the workspace cannot provide correct occlusion cues. Occlusion is widely regarded as the most important depth cue [19], so despite the corresponding input and output volumes, users sometimes found the workspace difficult to use.

Figure 2.8 Schmandt's stereoscopic workspace [142].

2.4.2 Ware's investigations of the "bat"

Ware has investigated interaction techniques for a six degree-of-freedom magnetic tracker, which he refers to as the "bat" [176][179]. Ware reports that it is difficult to achieve precise 3D positioning using a 1:1 control ratio when the arm or hand is unsupported. He finds that rotations of the bat produce inadvertent translations; interaction techniques which require the user to precisely control both position and orientation at the same time are difficult to master. But when the requirement for precision is relaxed, test users can make effective use of all six degrees of freedom: simultaneous positioning and orientation yields faster task completion times than separate positioning and orientation. Ware uses the bat as a relative device. A button on the bat acts as a clutch allowing or disallowing movement, enabling users to employ a "ratcheting" technique to perform large translations or rotations.

2.4.3 High Resolution Virtual Reality

Deering [46] describes a system which uses a hand-held 3D input device in conjunction with stereoscopic projection (using a standard desktop monitor and shuttered glasses). The effect is similar to that originally explored by Schmandt (*fig. 2.8*), as the stereoscopic 3D volume and the 3D input volume correspond. In one demonstration of the system, the user holds his or her hand up to a (stereoscopically projected) miniature virtual lathe. Deering particularly emphasizes achieving highly accurate registration between the 3D input device and the projected graphics, resulting in a fairly convincing illusion of “touching” the high-resolution 3D graphics with the input device.

2.4.4 JDCAD

Liang’s JDCAD [111] is a solid modeling system which is operated using a magnetic tracker similar to Ware’s “bat” [176]. By switching modes, the bat can be used to interactively rotate and move the model under construction, to select objects for subsequent operations or to orient and align individual pieces of the model.

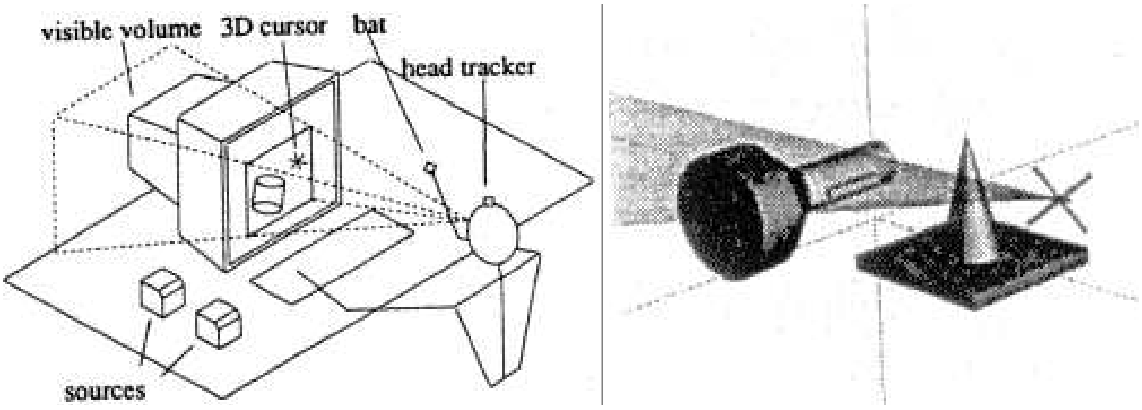


Figure 2.9 JDCAD configuration and cone selection technique [111].

The system set-up is similar to Deering’s system [46], with head tracking and a single hand-held input device used in front of a standard monitor (*fig. 2.9, left)*. JDCAD, however, uses an object selection mechanism based on ray or cone casting. The bat is used to shoot a ray or cone into the scene, allowing the user to hold the input device in a comfortable position and rotate it to change the cone direction. The objects intersected by the cone then become candidates for selection (*fig. 2.9*, *right*).

Another innovation introduced by Liang is the *ring menu*. This is a menu selection technique that is designed specifically for a hand-held 3D input device, and thus eliminates the need to switch between the bat and the mouse for menu selection. The items available for selection appear in a semi-transparent belt; a gap in this belt always faces the user and indicates the selected item. Figure 2.10 shows a user selecting a cylinder geometric primitive using the ring menu. Rotation of the bat about the axis of the user's wrist causes a new item to rotate in to the gap and become selected. This technique is useful, but it does not scale well to a large number of items, or to selection of menu items which cannot be easily represented by a 3D graphics icon.

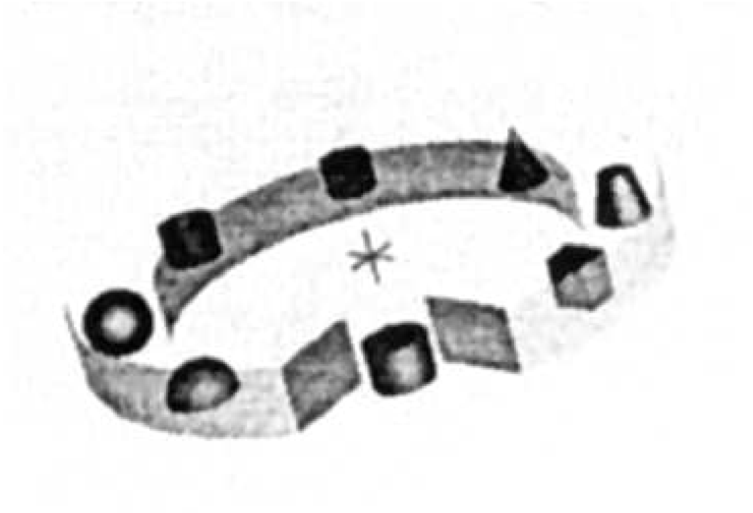


Figure 2.10 The ring menu technique for 3D menu selection [111].

Liang has evaluated JDCAD by comparing it in usability tests with commercially available 3D modelling packages [112]. Liang has not, however, performed formal experimental analysis of the factors which contributed to JDCAD's design.

2.4.5 Butterworth’s 3DM (Three-Dimensional Modeler)

Butterworth [26] describes a 3D CAD (Computer Aided Design) system for use in head-mounted displays. One hand is used for all interactions, including selecting commands from a floating menu, selecting objects, scaling and rotating objects, or grabbing vertices to distort the surface of an object. Modes for gravity, plane snapping, and grid snapping aid precise placement. The interface has support for multiple navigation modes, including *growing* and *shrinking* the user to allow work at multiple levels of detail; *walking* a short distance within tracker range; and *grabbing the world* to drag and rotate it. Butterworth reports that “since the user can become disoriented by all of these methods of movement, there is a command that immediately returns the user to the initial viewpoint in the middle of the modeling space”[26]. This is evidence of an underlying usability problem; the command to return the user to the initial viewpoint helps, but the reported user disorientation indicates that the underlying problems of navigation and working at multiple scales still need additional research.

2.5 Two-handed spatial interaction techniques

2.5.1 3Draw

Sach’s 3-Draw system is a computer-aided design tool which facilitates the sketching of 3D curves [141]. In 3-Draw, the user holds a stylus in one hand and a tablet in the other. These tools serve to draw and view a 3D virtual object which is seen on a desktop monitor. The palette is used to view the object, while motion of the stylus relative to the palette is used to draw and edit the curves making up the object (*fig. 2.11*). Sachs notes that “users require far less concentration to manipulate objects relative to each other than if one object were fixed absolutely in space while a single input sensor controlled the other”.

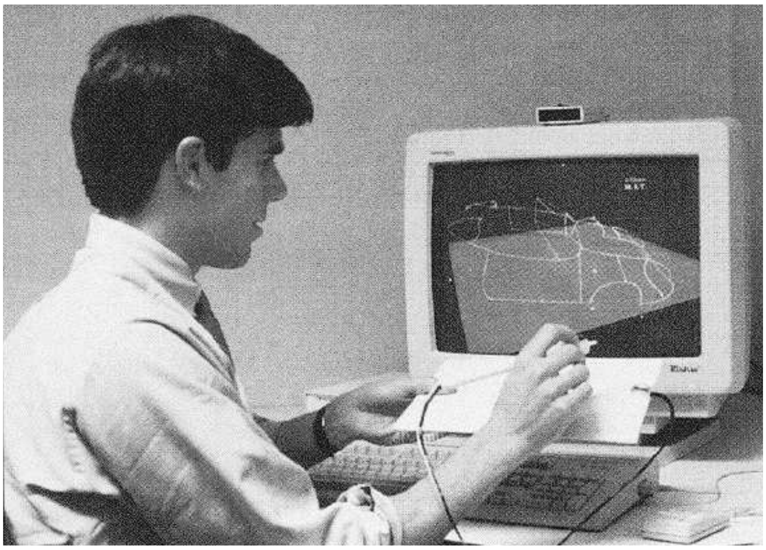


Figure 2.11 The 3Draw computer-aided design tool [141].

2.5.2 Interactive Worlds-in-Miniature

Stoakley's Worlds-in-Miniature (WIM) interface metaphor [164] provides the virtual reality user with a hand-held miniature representation of an immersive life-size world (*fig. 2.12, left*). For example, the user can design a furniture layout for the room in which he or she is standing. Users interact with the WIM using both hands. The user's non-preferred hand holds the WIM on a clipboard while the preferred hand holds a "button ball," a ball-shaped 3D input device instrumented with some buttons (*fig. 2.12, right*). Moving a miniature object on the WIM with the ball moves the corresponding life-size representation of that object.

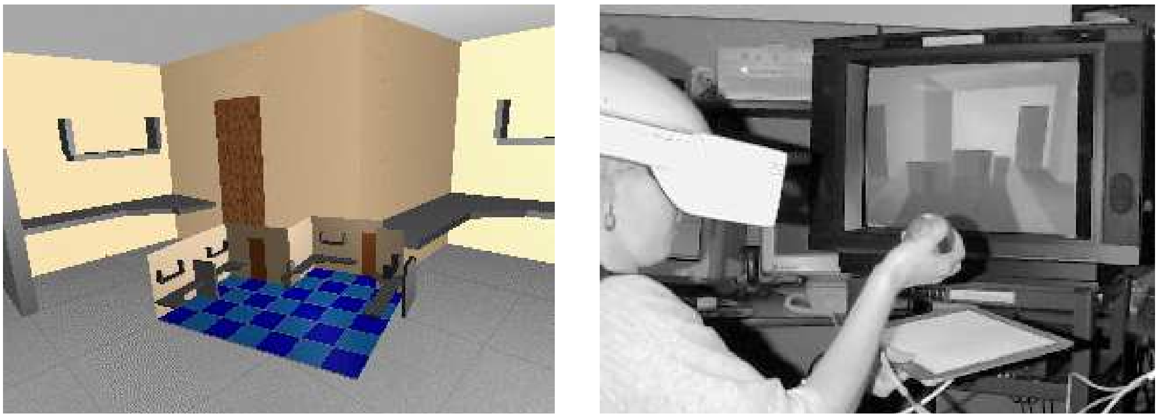


Figure 2.12 The Worlds-In-Miniature (WIM) metaphor [164].

The WIM effectively integrates metaphors for viewing at 1:1 scale, manipulating the point-of-view, manipulation of objects which are out of physical reach or occluded from view, and navigation [133]. I will discuss some further issues raised by two-handed interaction with the WIM in the context of Chapter 4, “Design Issues in Spatial Input,” as well as Chapter 8, “The Bimanual Frame-of-Reference.”

2.5.3 The Virtual Workbench

Poston and Serra [138] have implemented the Virtual Workbench (*fig. 2.13*), a mirrored display similar to the system implemented by Schmandt (*fig. 2.8*), but the mirror which Poston and Serra use is opaque. The means that all images are completely computer generated, allowing correct occlusion cues to be maintained. The Virtual Workbench employs a physical tool handle which can have different virtual effectors attached to it, depending on the current mode. In some modes it is possible to use both hands to assist manipulation (*fig. 2.13*). Poston and Serra [148] report that physical layout of the workspace is important: the user’s head has to stay in front of mirror, so users with differing body sizes may need to make adjustments to their working environment to avoid uncomfortable postures.

The Virtual Workbench has been developed with medical applications in mind. It includes support for tasks of interest in neurosurgical visualization, including cross-sectioning a volume. Poston and Serra have not, however, performed any formal experimental evaluation of the design issues raised by the Virtual Workbench.

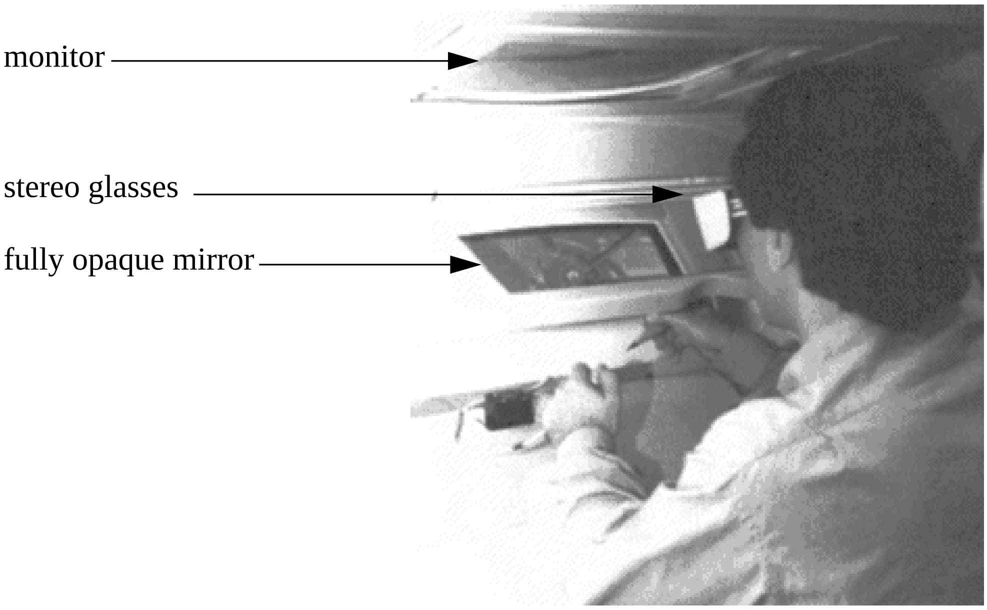


Figure 2.13 The Virtual Workbench [138].

2.5.4 The Reactive Workbench and ImmersaDesk

Another recent trend in interactive 3D graphics is to use stereoscopic projection on a tilted large-screen projection display. Examples of such displays include the Reactive Workbench [52] and the ImmersaDesk [49]. The user typically wears stereo glasses, causing stereoscopically projected objects to appear to occupy the space in front of or behind the display. The large display fills a wide visual angle and gives a better sense of presence than an ordinary desk top display. These displays are useful when multiple persons wish to collaborate on a single problem; much like drafting board invites others to look over one’s shoulder at ongoing work, the large display surface conveys a subtle social message which tells others that the information is meant for sharing. Multiple participants, however, cannot simultaneously view accurate stereoscopic data, because the geometry of the stereoscopic projection can only be correct for one viewer.

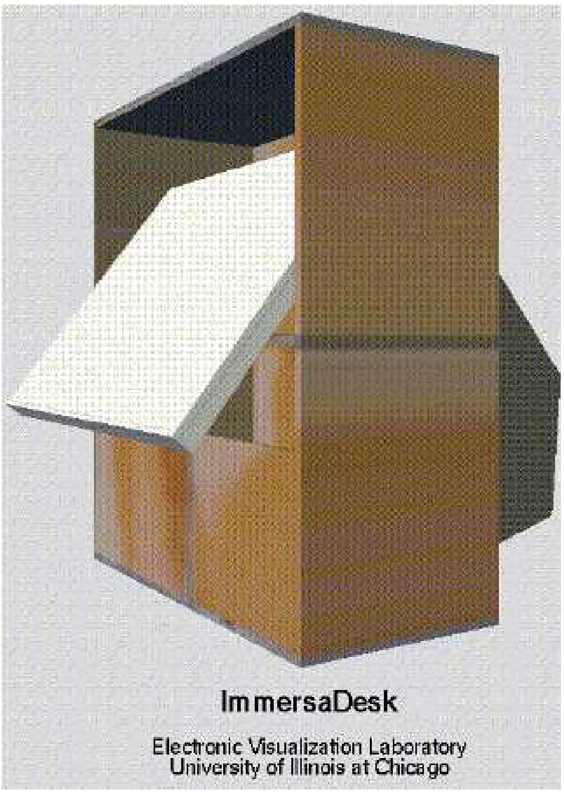


Figure 2.14 The ImmersaDesk [49].

2.5.5 Other notable systems

Several other systems demonstrate two-handed techniques for virtual object manipulation. Shaw's THRED system for polygonal surface design [149] employs both hands. The nonpreferred hand indicates constraint axes, performs menu selections, and orients the scene, while the preferred hand performs all detailed manipulation. Shaw has performed users tests on THRED [150], but he has not reported any formal behavioral studies. Leblanc [107] describes a sculpting application which uses a Spaceball (a force-sensing six degree-of-freedom input device that rests on the desk surface) in the nonpreferred hand to orient a sculpture, while the preferred hand uses a mouse to select and deform the vertices making up the sculpture. Mine [119] has recently added two-handed manipulation to an immersive architectural design application. Mine uses the separation of the hands to indicate the magnitude of object displacements, and to select options from a “mode bar” extending between the hands.

2.6 Theory and experiments for two hands

2.6.1 Guiard’s Kinematic Chain Model

Guiard’s analysis of human skilled bimanual action [67] provides an insightful theoretical framework for classifying and understanding the roles of the hands. The vast majority of human manual acts involve two hands acting in complementary roles. Guiard classifies these as the *bimanual asymmetric* class of manual actions.

Guiard has proposed the Kinematic Chain as a general model of skilled asymmetric bimanual action, where a *kinematic chain* is a serial linkage of abstract motors. For example, the shoulder, elbow, wrist, and fingers form a kinematic chain representing the arm. For each link (e.g. the forearm), there is a *proximal* element (the elbow) and a *distal* element (the wrist). The distal wrist must organize its movement relative to the output of the proximal elbow, since the two are physically attached.

The Kinematic Chain model hypothesizes that the preferred and nonpreferred hands make up a *functional* kinematic chain: for right-handers, the distal right hand moves relative to the output of the proximal left hand. Based on this theory and observations of people performing manual tasks, Guiard proposes three high-order principles governing the asymmetry of human bimanual gestures, which can be summarized as follows:

(1) *Motion of the preferred hand typically finds its spatial references in the results of motion of the nonpreferred hand.* The preferred hand articulates its motion relative to the reference frame defined by the nonpreferred hand. For example, when writing, the

nonpreferred hand controls the position and orientation of the page, while the preferred hand performs the actual writing by moving the pen relative to the nonpreferred hand (*fig. 2.15*). This means that the hands do not work independently and in parallel, as has often been assumed by the interface design community, but rather that the hands specify a hierarchy of reference frames, with the preferred hand moving relative to the nonpreferred hand.

(2) *The preferred and nonpreferred hands are involved in asymmetric temporal-spatial scales of motion*. The movements of the preferred hand are more frequent and more precise than those of the nonpreferred hand. During handwriting, for example, the movements of the nonpreferred hand adjusting the page are infrequent and coarse in comparison to the high-frequency, detailed work done by the preferred hand.

(3) *The contribution of the nonpreferred hand starts earlier than that of the preferred*. The nonpreferred hand precedes the preferred hand: the nonpreferred hand first positions the paper, then the preferred hand begins to write. This is obvious for handwriting, but also applies to tasks such as swinging a golf club [67].

Figure 2.15 Guiard's handwriting experiment [67].

A handwriting experiment by Guiard illustrates these principles in action (*fig. 2.15*). The left half of the image shows an entire sheet of paper as filled out by the subject on dictation. The right half of the image shows the impression left on a blotter which was on a desk underneath the sheet of paper. The experiment demonstrates that movement of the dominant hand occurred not with respect to the sheet of paper itself, but rather with respect to the postures defined by the non-dominant hand moving and holding the sheet of paper over time.

In a related experiment, Athenes [7], working with Guiard, had subjects repeatedly write a memorized one-line phrase at several heights on individual sheets of paper, with the nonpreferred hand excluded (no contact permitted with the sheet of paper) during half the trials. Athenes's study included 48 subjects, including a group of 16 right-handers and two groups of left-handers,[1] each again with 16 subjects. Athenes's results show that when the nonpreferred hand was excluded, subjects wrote between 15% and 27% slower, depending on the height of the line on the page, with an overall deficit of approximately 20%. This result clearly shows that handwriting is a two-handed behavior.

Although for consistency I have used handwriting as an example throughout this section, note that Guiard's original analysis is rich with many examples of bimanual acts, such as dealing cards, playing a guitar, swinging a golf club, hammering a nail, or unscrewing a jar cap. Guiard carefully describes how the principles which he proposes apply to a wide range of such examples.

Guiard also provides an overall taxonomy of bimanual actions. The three principles outlined above apply to the *bimanual asymmetric* class of manipulative actions only. The above principles do not apply to *bimanual symmetric* motions, such as weight lifting,

---

[1] Athenes used two groups of left-handers because there are at least two distinct left-handed handwriting postures; these postures will be discussed further in section 7.2.1 of Chapter 7, "Issues in Bimanual Coordination."

climbing a rope, rowing, jogging, or swimming. In particular, during locomotion, symmetry seems to be a virtue, since asymmetric motions would result in postural instability (loss of balance). Symmetric motion is also useful to specify scale or extent, as demonstrated by Kreuger's two-handed technique for sweeping out a rectangle (*fig. 2.7, right*) [104]. As a final note, Guiard's principles also do not apply to the communicative gestures which accompany speech (for example, see the work of Kimura [101][102]), or those that are used for sign language. Such gestures are most likely a completely separate phenomenon from the manipulative gestures studied by Guiard.

Looking beyond the hands, one might also apply the Kinematic Chain model to reason about multiple effector systems ranging from the hands and voice (playing a piano and singing [68]), the hands and feet (operating a car's clutch and stick shift), or the multiple fingers of the hand (grasping a pen). For example, for the task of playing a flute, Guiard argues that the preferred hand organizes its action relative to the nonpreferred hand as usual, but that the mouth (breathing and tongue motion) performs the highest frequency actions relative to the hands. Similarly, one's voice can also be thought of as a link in this hierarchy. Although multimodal two handed input plus voice input will not be a subject of this dissertation, Guiard's model could offer a new perspective for studying how and when the hands and voice can work well together, and how and when they cannot. In existing studies that I am aware of, the hands and voice have generally been thought of as independent communication channels rather than as mutually dependent or hierarchically organized effectors [17][76][180]. For example, a pianist can easily sing the right-handed part of a piece of music, while continuing to perform the left-handed part, but singing the left-handed part without error is difficult or impossible to achieve, even for simple compositions [68].

2.6.2 Formal experiments

There are few formal experiments which analyze two-handed interaction, and I am not aware of any previous formal experimental work which has studied two hands working in three dimensions. There is an extensive literature of formal analyses of bimanual tasks in the psychology and motor behavior fields; I will briefly review this literature in the context of Chapter 7, “Issues in Bimanual Coordination.”

2.6.2.1 Buxton and Myers experiments

A classic study by Buxton and Myers [27] (*fig. 2.16*) demonstrated that two-handed input can yield significant performance gains for two compound tasks that were studied: a select / position task and a navigate / select task. Their results demonstrated that two-handed techniques were not only easily learned by novices, but also that the two-handed techniques improved the performance of both novice and expert users.

Figure 2.16 Configuration for Buxton and Myers experiments [27].

In the first experiment reported by Buxton and Myers, subjects positioned a graphical object with one hand and scaled it with the other. The task allowed subjects to adopt either a strictly serial strategy (i.e., position first, then scale) or a parallel strategy (position and scale at the same time). Buxton and Myers found that novices adopted the parallel task strategy without prompting and that task performance time was strongly correlated to the degree of parallelism employed.

In a second experiment, subjects scrolled through a document and selected a piece of text. Buxton and Myers found that both experts and novices exhibited improved performance using two hands to perform this task versus using one hand, and they also found that novices using two hands performed at the same level as experts using just one hand.

2.6.2.2 Kabbash experiments

Kabbash, MacKenzie, and Buxton [94] compared pointing and dragging tasks using the preferred hand versus the non-preferred hand. For small targets and small distances, the preferred hand exhibited superior performance, but for larger targets and larger distances, there was no significant difference in performance. Contrary to the traditional view that humans have one “good” hand and one “bad” hand, the authors concluded that “the hands are complementary, each having its own strength and weakness” [94].

A second experiment by Kabbash, Buxton, and Sellen [95] studied compound drawing and color selection in a “connect the dots” task (*fig. 2.17*). The experiment evaluated the two-handed ToolGlass technique [15]. ToolGlass consists of a semi-transparent menu which can be superimposed on a target using a trackball in the nonpreferred hand. The preferred hand can then move the mouse cursor to the target and *click through* the menu to apply an operation to the target. Note that this integrates the task of selecting a command (or mode) from the menu and the task of applying that command to objects being edited. In Kabbash’s experiment, the ToolGlass was used to select one of four possible colors for each dot.

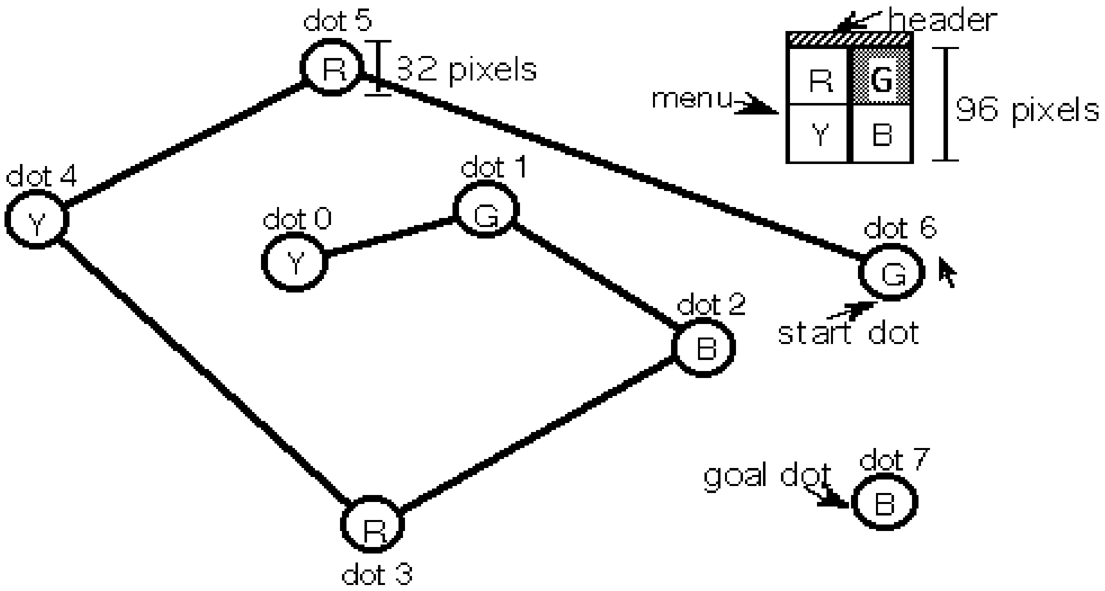


Figure 2.17 Second experimental task used by Kabbash [95].

Kabbash proposes that two-handed input techniques which mimic everyday tasks conforming to Guiard’s bimanual asymmetric class of gestures can produce superior overall performance. Kabbash’s results suggest that everyday two-handed skills can readily transfer to the operation of a computer, even in a short interval of time, and can result in superior performance to traditional one-handed techniques. This result holds true *despite* the benefit of years of practice subjects had with the one-handed techniques versus only a few minutes of practice with the two-handed techniques.

Kabbash also demonstrates that, if designed incorrectly, two-handed input techniques can yield *worse* performance than one-handed techniques [95]. In particular, the authors argue that techniques which require each hand to execute an independent sub-task can result in increased cognitive load, and hypothesize that consistency with Guiard’s principles is a good initial measure of the “naturalness” of a proposed two-handed interaction.

2.6.2.3 Leganchuk's area selection experiment

Leganchuk [108] has adapted Kreuger's camera-based techniques (*fig. 2.7*) [104] to tablet-based input using a pair of wireless devices (*fig. 2.18*). Leganchuk's experiment studied an "area sweeping" task in which subjects selected an area encompassing a target. This is similar to sweeping out a rectangle to select a set of targets in a graphics editing application. Using both hands allowed subjects to complete the task significantly faster than using just one hand.

For the one handed technique, Leganchuk also timed how long it took users to switch between two separate control points used to sweep out the rectangle. Even when this time was removed from the one-handed technique, the two-handed technique was still faster, especially for sweeping out larger, more difficult areas. Thus, Leganchuk argues that the difference in times cannot be attributed to the increased time-motion efficiency alone, and interprets this as evidence that the bimanual technique "reduces cognitive load."

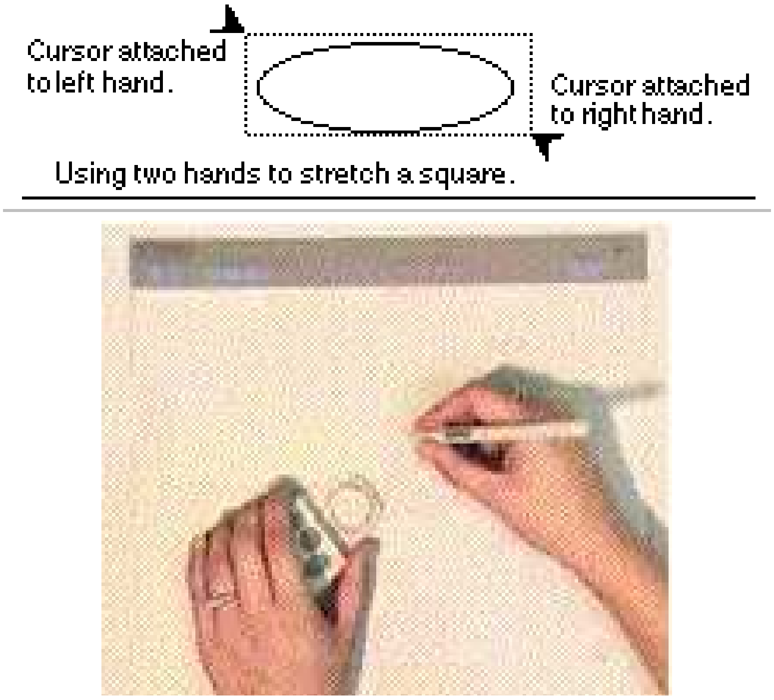


Figure 2.18 Experimental task and apparatus used by Leganchuk [108].

2.7 Summary

One thing that stands out in a review of the user interface literature is how few systems have supported two-handed interaction. There are several reasons for this, including cost, lack of support from windowing systems and user interface toolkits, and the inexperience of designers in developing two-handed interfaces. And I am only aware of a handful of formal experiments which have studied the nature of two-handed interaction with computers, most of which have been performed in the context of Buxton's Input Research Group at the University of Toronto. To my knowledge there has not yet been any experimental work which has studied two hands working in three dimensions with multiple degree-of-freedom input devices[1].

My work is different and new along several dimensions. First, I have applied virtual manipulation to a develop a novel interface for volume visualization; the application to neurosurgical visualization is new. Second, much of the previous work has been driven by the technology rather than the real-world needs of an actual user. By working closely with domain experts, I have been able to focus on a real problem and I have been able to contribute not only a description of the design, but also the results of usability tests with real users. Many other efforts have been aimed at general applications without a specific focus or a specific user base.

Finally, most of the prior research has taken either a purely systems-building approach, which boils down to an account of "here is something we built," or it has taken a purely experimental approach, performing psychology experiments which may or may not be relevant to the problems faced by designers. By combining these approaches, my work offers an integrated discussion which is both timely and relevant to design.

---

1. Chris Shaw has performed usability studies on the THRED two-handed polygonal surface design system [150], but he has not performed any experimental analyses of the underlying behavioral issues.

> *Hitching our research to someone else's driving problems, and solving those problems on the owners' terms, leads us to richer computer science research.*
>
> *Fred Brooks [20]*

Chapter 3

System Description

3.1 Overview

I describe a 3D neurosurgical visualization system which incorporates a 3D user interface based on the two-handed physical manipulation of hand-held tools in free space. These *user interface props* facilitate transfer of the neurosurgeon's skills for manipulating tools with two hands to the operation of a user interface for visualizing 3D medical images, without need for training.

From the surgeon's perspective, the interface is analogous to holding a miniature head in one hand which can be "sliced open" or "pointed to" using a cutting-plane tool or a stylus tool, respectively, held in the other hand. The interface also includes a touchscreen which allows facile integration of 2D and 3D input techniques. Informal evaluations of over fifty neurosurgeons, as well as hundreds of non-neurosurgeons, have shown that with a cursory introduction, users can understand and use the interface within about one minute of touching the props.

3.2 The application domain: neurosurgery and neurosurgeons

Neurosurgeons are driven by a single goal: deliver improved patient care at a lower cost. While improving quality of care and reducing costs might seem to be at odds, in practice one can achieve both ends by reducing the time required to perform surgery. Operating room time itself is of course very expensive. But more importantly, the longer a patient’s brain is exposed during a procedure, the greater the chance for expensive and life-threatening complications.

The key to reducing operating room time is superior surgical planning. Neurosurgery is inherently a three-dimensional activity; it deals with complex structures in the brain and spine which overlap and interact in complicated ways. To formulate the most effective surgical plan, the neurosurgeon must be able to visualize these structures and understand the consequences of a proposed surgical intervention, both to the intended surgical targets and to surrounding, viable tissues.

A certain class of procedures known as *stereotaxy,* which use a metal *stereotactic frame* bolted to the patient’s skull, are particularly amenable to computer-assisted surgical planning because the frame provides a known and fixed reference coordinate system. The frame also serves to guide surgical instruments to the target with millimeter accuracy. For procedures which cannot use a frame, such as aneurysm or blood vessel surgeries, the surgeon typically views the surgical theatre through a microscope while operating directly on the brain with hand-held tools.

3.2.1 Traditional practice

Traditionally, neurosurgeons have planned surgery based on 2D slices, acquired through scanning techniques such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). This restriction to 2D slices is not necessarily by preference. MRI is acquired as 3D volumetric data, and its presentation as a set of 2D slices is an artifact of limited computer technology. Even CT is usually acquired as multiple slices that are closely spaced, and thus often can be treated as volumetric data.

The 2D slice paradigm has imposed a further restriction. The images are typically restricted to appear in planes orthogonal to canonical axes through the patient’s head. These orthogonal planes are known as the sagittal, coronal and axial planes and are shown in figure 3.1, reading clockwise from the top. The orthogonal planes form the frame of reference in which medical students learn their anatomy, and as such they are the planes in which physicians can best understand and reason about the anatomy. But many structures within the brain, and many surgical paths to these structures that are clinically useful, are oblique to these canonical views. For example, to reach a deep-seated structure, the neurosurgeon might follow a fold of the brain to reduce the amount of transected brain tissue. Such an approach is often not possible using the canonical planes.

Even for experienced neurosurgeons, the development of a three-dimensional mental model from 2D slices of the anatomy remains a challenging task. Traditionally, following an oblique trajectory to a target has been risky since it has been difficult or impossible to produce appropriate visualizations. This is why visualization of oblique slices is so important: it is difficult to understand the anatomy at oblique angles, so the surgeon wants to be able to see these views and relate them back to the more familiar canonical views.

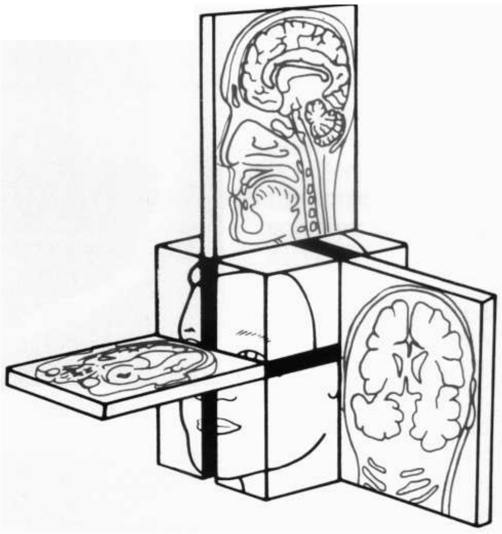


Figure 3.1 The canonical sagittal, coronal, and axial orthogonal planes.

3.2.2 Computer-assisted surgery

Neurosurgeons have recently become increasingly interested in computer-based surgical planning systems which allow them to quantify and visualize the three-dimensional information available from medical imaging studies. By making better use of this three-dimensional information, and by allowing the surgeon to quickly and intuitively access it, computer-based visualization and planning systems can positively impact both cost-of-care and patient outcome.

A typical computer-assisted surgery consists of the following elements:

Stereotactic frame placement: The stereotactic frame is bolted to the patient’s skull prior to imaging, to ensure that it stays rigidly fixed in place during imaging and surgery. At the University of Virginia, we use the Leksell stereotactic frame [109]. The frame serves two main purposes: it provides a coordinate system for the patient’s head, and it carries mechanical guides that serve as a rigid mounting platform for surgical instruments. During medical image acquisition, the frame is fitted with a system of fiducial markers. These markers form a distinctive pattern in the digital patient images which can be used to calculate a transformation between the digital image coordinates and the stereotactic frame coordinates, allowing transfer of the computer-based surgical plan to the operating room.

Medical image acquisition: Our neurosurgeons principally depend upon MRI (*fig. 3.2*, *left*) for anatomic information due to its high soft tissue contrast and its capability for true 3D data acquisition. When necessary, the surgeon may also request digital subtraction angiography images (*fig. 3.2, center*), which detail the vascular structures (veins or arteries) within the brain, or CT (*fig. 3.2, right*), which is sometimes needed for bony detail or visualization of calcifications. In general, however, acquiring multiple image modalities is expensive, time-consuming, and stressful for the patient, so the surgeon will only order the minimum necessary set of image modalities.

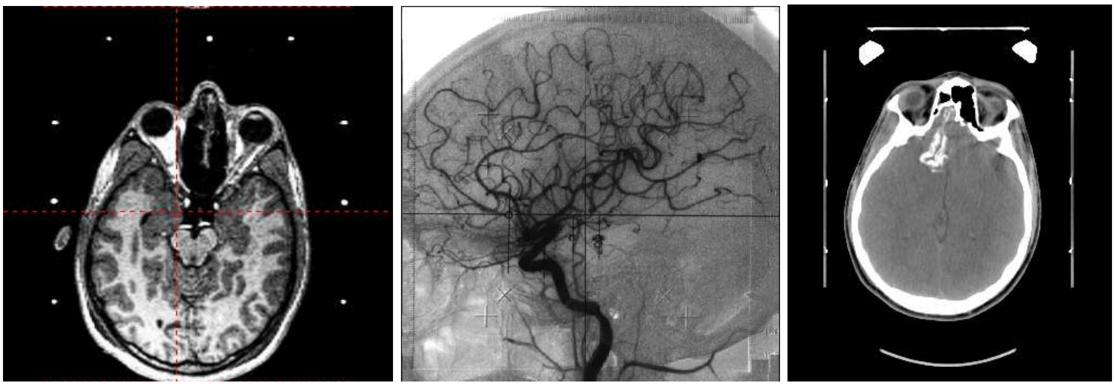


Figure 3.2 Example MR, angiography, and CT images.

Image segmentation and classification: We employ computer algorithms to identify the surface and major structures of the brain, to delineate pathology, and to identify major blood vessels [92][97][157][158][159]. Robust identification of these structures is often helpful during visualization and is necessary for quantitative evaluations (such as tracking the volume of a tumor over time). Manual approaches are possible, but are time-consuming and error prone, making them impractical for surgical planning.

Pre-surgical planning: The user interface discussed in this dissertation focuses on the pre-surgical planning phase, which usually takes place on the morning of surgery. Planning is typically done in a separate surgical planning room rather than in the operating room itself. To develop a plan, the surgeon uses visualization, measurement, and other planning tools to select the surgical targets and to select a path to those targets that produces the least possible damage to viable tissue. To plan the best possible trajectory, the surgeon needs to understand the spatial relationships between the targets, surrounding structures, functional areas which must not be damaged, and feasible entry points. Feasible entry points must avoid cortical vessels or skin surface areas which are inaccessible due to mechanical limitations of the stereotactic frame. Visualizations of the cortical surface, the stereotactic frame, proposed surgical trajectories, and volume cross-sections at both orthogonal and oblique angles can all help the neurosurgeon to make informed decisions. These visualizations also serve as visual checks that there has not been human or software-generated errors during imaging, frame registration, or image segmentation. For example, if the visualization of the stereotactic frame does not line up with the registration markers in the image, then the registration is clearly incorrect.

Surgery: Finally, the surgery itself is ready to proceed. At this point, the surgeon has carefully studied the intended target(s) and the surgical trajectory, so the main difficulty is accurate navigation along the selected trajectory. Intra-operative guidance for surgical navigation is a current research area [48][106], but not a topic of this dissertation. In some rare cases, the surgeon may elect to modify the plan intra-operatively based on information which was not visible in the medical images.

3.2.3 Some system requirements

All the activity leading up to the surgery itself must occur during a span of approximately 3-4 hours on the morning of surgery[1]. To be clinically useful, a computer-based surgical planning system must be able to produce all of its results within this time window. Since the principal neurosurgeon is extremely busy and may be providing care for several other patients, the actual time available for planning may be as little as fifteen minutes for the more straightforward cases.

Thus the user interface for a neurosurgical planning and visualization system must permit the surgeon to work quickly: the morning of surgery is perhaps the least optimal time for a surgeon to be fussing with a maze of slider bars and command prompts. Also, the surgeon must cope with frequent distractions, and therefore must be able to quickly detach from the user interface, both physically and cognitively. Thus, the interface must not encumber the surgeon with devices such as gloves or head-mounted displays that will be difficult to remove, and it must not have explicit modes that are easily forgotten during a phone call or a discussion with a colleague.

Software usability is crucial to get neurosurgeons to actually use advanced visualization software in the clinical routine. I have sought to design interaction techniques which facilitate use of the software by surgeons, without need for technical assistance. The manipulative capabilities of input devices such as mice and keyboards are poorly matched to the volumetric manipulation and visualization tasks of interest to the neurosurgeon. Rather than typing in commands or moving sliders with a mouse, the neurosurgeon thinks in terms of real objects in real space; a three-dimensional user interface should allow the

---

1. It is important to minimize the time window between imaging and actual surgery so that the digital images do not become a stale representation of the brain (the brain can change over time, especially in a sick or injured patient), and of course because the patient (who must be conscious for some procedures) may experience considerable pain while wearing the stereotactic frame.

neurosurgeon to work and think in these same terms. As one surgeon put it, “I want a skull I can hold in my hand.”

Our laboratory has worked closely with the neurosurgeons at the University of Virginia throughout the design of our 3D neurosurgical planning system and the associated 3D user interface. My work on the user interface, in particular, has necessarily been heavily collaborative, relying on the advice and opinions of neurosurgeons to provide goals and specifications throughout the design process.

Note that the neurosurgeon’s existing visualization and planning tools are almost exclusively two-dimensional. This is an artifact of historical technological limitations rather than preference; computer-assisted three-dimensional surgical planning can allow neurosurgeons to view and explore the individual patient’s anatomy in ways that previously have not been possible. Our laboratory’s initial clinical experience suggests that computer-assisted three-dimensional surgical planning can allow surgeons to approach old problems in new and more efficient ways, and to treat borderline cases that may have been considered too risky to treat with traditional techniques.

3.3 System design philosophy

In our everyday lives, we are constantly confronted with tasks that involve physical manipulation of real objects. We typically perform these tasks with little cognitive effort, with both hands [27], and with total confidence in our movements. For many applications, a three-dimensional user interface should offer equally facile interaction.

I propose a 3D interface which permits the user to manipulate familiar objects in free space. These *passive interface props* act as tools which help users reason about their tasks. With six degree-of-freedom magnetic trackers [137] unobtrusively embedded within the props, the computer can observe the user’s gestures. This results in a human-computer dialog where the system *watches* the user [132], in contrast to the traditional approach where the user generates input tokens in a contrived dialog.

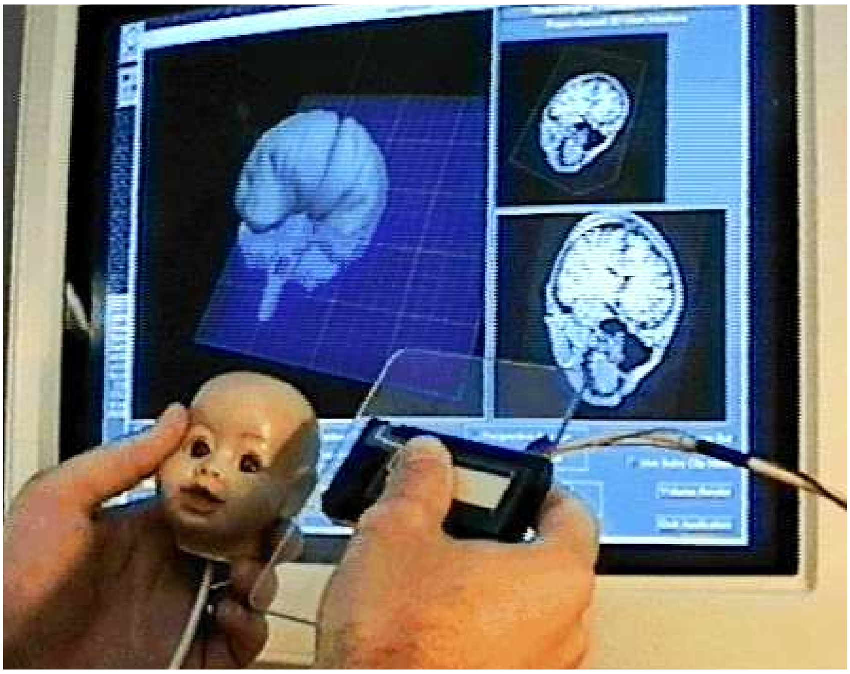


Figure 3.3 User selecting a cutting-plane with the props.

An interface which requires the neurosurgeon to wear an instrumented glove and make grabbing gestures to manipulate imaginary objects would not offer this style of interaction. No matter how realistic the on-screen graphics are, the user does not experience the visceral kinesthetic and tactile feedback which comes from grasping a real-world object. When the user holds a physical tool, he or she has passive haptic feedback to guide the motion of the hand, allowing all the degrees-of-freedom of the fingers, thumb, and palm to participate in the manipulation of the tool.

Compared to “3D widgets” [43][77][78] (as shown in figure 2.1 on page 15) a props-based interface offers several advantages. With props, there is no need to make a widget’s behavior explicit or to make the user realize the widget is an active interface component. The appearance of the props indicates their use and their *palpability* makes users immediately and continuously aware they exist. Drawing a widget without cluttering the scene becomes trivial, since there is no widget. Also, for casual users such as surgeons, manipulating a real tool is familiar and natural, whereas an abstract widget, no matter how well designed, is not. Instead of having to expend cognitive effort on the basic acts of manipulating objects, my approach allows the user to employ the normal perceptual and motor systems to manipulate objects, so that the user can properly focus on the intellectually challenging portion of the task.

In the domain of neurosurgical planning and visualization, the props-based interface has proven successful and has elicited enthusiastic comments from users. Approximately 50 neurosurgeons have tried the interface, representing about 1% of the total user population of 4,000 neurosurgeons in the United States[1]. With a cursory introduction, neurosurgeons who have never before seen the interface can understand and use it without training.

3.4 Real-time interaction

Without the ability to render and manipulate images of the brain in real time, my approach to the interface would be infeasible. The system software has been designed to achieve high performance: typical interactive update rates are approximately 15-18 frames per second[2]. A frame rate of 10 frames per second is usually considered to be the minimum update rate at which humans can still fuse separate frames into apparent motion [38].

During each frame, the system renders a simplified brain surface representation consisting of approximately 9,000 polygons and displays a volumetric cross-section from data which typically consists of 256 x 256 x 128 voxels (volume elements), each 2 bytes

---

[1] The estimate of 4,000 practicing neurosurgeons in the United States was provided by Dr. Neal Kassell of the Neurosurgery department.

[2] The software runs on a Hewlett Packard J210 workstation with the “Visualize” hardware polygonal and texture mapping acceleration.

wide, for a total of 16 megabytes of volume data. To support selection of points on the brain surface, during each frame the system must also compute the intersection of a ray with the 9,000 polygon brain surface. And of course, the system must also communicate with the actual input devices, calculate matrix transformations, and compute geometrical relationships as a precursor to rendering the actual graphics for a frame.

3.5 Props for neurosurgical visualization

3.5.1 Viewing patient data with a head prop

The surgeon uses a *head prop* to manipulate the individual patient's head data. The prop is a small doll's head which can be held comfortably in one hand. I have tried head props of various sizes: if the input device is too small, it is easy to drop it by mistake, and the connecting cable significantly impedes rotating the device. If the input device is too large, it becomes difficult to tumble with the fingers of one hand. The right size seems to be roughly 2.0 to 2.5 inches in diameter.

I have also tried using a small rubber ball, but users prefer the doll's head because it is much richer in tactile orientation cues. The orientational cues help users to better understand what the input device does, and suggests appropriate behavior for three-dimensional manipulation: people's first instinct is to roll a ball on the desk, but they will pick up the head. The doll's head itself also provides a certain amount of "marketing appeal" and serves as a memorable icon for the interface.

The doll's head prop is an absolute rotation controller: rotating the doll's head always causes a polygonal model of the patient's brain to rotate correspondingly on the screen. The user can control the image zoom factor by moving the prop towards or away from his or her body. Note, however, that the software does not in fact know where the user is sitting, so the zoom factor is actually based on the distance between the doll's head and the front of the screen. Also, since the angle subtended by the virtual object on the screen grows geometrically with linear translation towards the user, the virtual distance moved is actually a *log transform* of the physical front-back translation. Without this log transform, a front-back translation near the screen produces almost no effect on the zoom factor, whereas a translation near the user's body suddenly shoots the zoom factor beyond a useful range.

The doll's head provides only four degrees-of-freedom: three degrees-of-freedom for rotation plus one degree-of-freedom for the zoom factor. In the context of surgical visualization, moving the object left-right or up-down is typically not useful, so it is helpful to constrain the polygonal brain to appear at the center of the screen. This simplifies the task and users find it natural.

The original interface design envisioned a realistic skull-shaped prop, but retreated from this approach for the following reasons:

- Although many non-neurosurgeons have suggested using a more realistic head prop, *not even one* neurosurgeon has done so *after* operating the interface. In fact, when I suggest the idea, neurosurgeons flatly resist it. This includes the surgeon who originally said he wanted a *skull* he could hold in his hand.

- The ergonomics of a small device (the doll's head or a ball) are superior because the device can comfortably be held in one hand and at any orientation.

- Using a realistic head prop leads directly to false user expectations. For example, some "off-the-street" test users will sometimes hold the cutting plane up to the doll's eyes, expecting to see a cut directly through the orbits. But since neuroanatomy varies greatly between individuals, the morphology of the real-world prop and the virtual head do not precisely correspond. As a result, the cut does not go exactly through the orbits. Similarly, a neurosurgeon using a realistic skull might assume a precise registration between the real skull and the virtual

> head which simply does not exist. When using a more abstract form such as the doll’s head, surgeons do not expect the prop to precisely match the brain model, so this usability problem does not arise.

Also note that the scale of the doll’s head does not match the scale of the actual patient data. The doll’s head acts only as an orientational reference and its outer surface conveys no information about scale: touching the outer surface of the doll’s head does not correspond to touching the outer surface of the virtual patient data. Scaling down the virtual patient data to roughly match the size of the doll’s head would result in a substantial loss of precision: one millimeter of real-world motion would then correspond to several millimeters of virtual motion.

3.5.2 Slicing the patient data with a cutting-plane prop

The surgeon can also employ a *cutting-plane prop* to specify the position and orientation of an arbitrary slice through the patient’s anatomy. The prop itself is a rectangular plate with a housing for the tracker (*fig. 3.4, left*). Users can spread their fingers across the plate to get a direct haptic sense of how it is oriented in space. The appearance of the cutting-plane prop differentiates it from the head prop and makes its purpose immediately obvious.

Note that the cutting-plane prop is used in concert with the head prop rather than as a separate tool. The user holds the cutting-plane against the head to indicate a slice through the brain data. The computer shows a corresponding virtual tool intersecting the virtual head, along with a cross-section of the volumetric head data (*fig. 3.4, right*). The reader can easily approximate this interface. Seat yourself in a chair with armrests. Grasp a ball in one hand and a small book in the other. While supporting your elbows with the armrests, hold the book up to the ball, and orient each as deemed necessary. This is all that the interface requires for 3D manipulation.

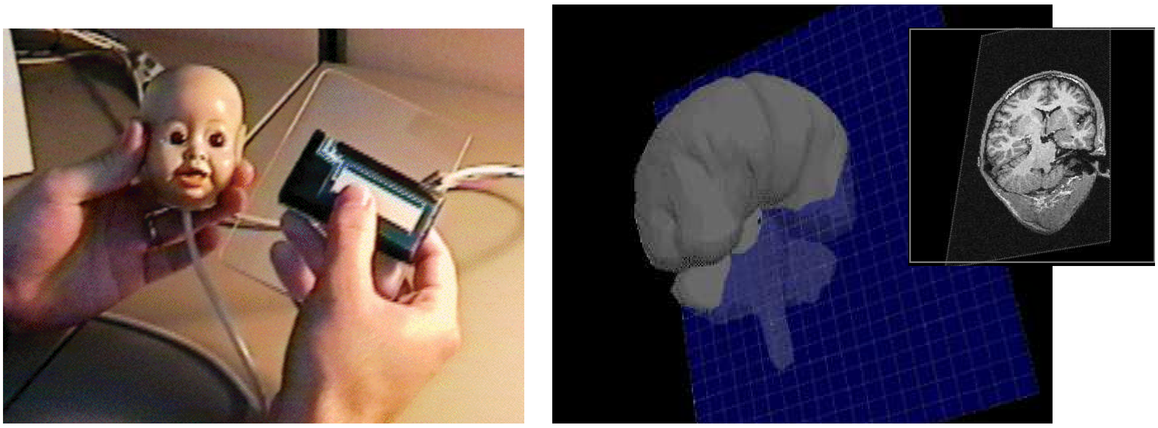


Figure 3.4 User indicating a cross-section.

There are three distinct clinical uses for the cutting-plane prop as I have implemented it:

Volume Exploration: The user can interactively sweep the cutting plane through the volume. Because of the interactive update rate, users can quickly develop a sense of the objects embedded in the volume by sweeping the plane back and forth. For example, structures which are difficult to visualize when viewing orthogonal slices can now be easily found and inspected: figure 3.5 shows a user moving the cutting-plane prop, over a period of a few seconds, to expose the optic nerves.

Volume Dissection: Once the plane is selected, a portion of the volume can be permanently cut away. I have implemented a version of the system with texture-mapping hardware which allows the polygonal object to be “capped” with a texture map of the exposed data. The texture map affords selection of a further cross-sectioning plane which passes through structures revealed by the initial cut, as further discussed further in section 3.7.2 of this chapter.

Measuring Distances: A grid pattern on the computer rendering of the plane can be used as a ruler. I had not anticipated that the cutting-plane prop could be used as a ruler, but much to my surprise some neurosurgeon test users started employing it for this purpose.

When the user manipulates real objects in real space, new or unusual ideas can readily be expressed; the user is not artificially bound by an abstraction or a metaphor.

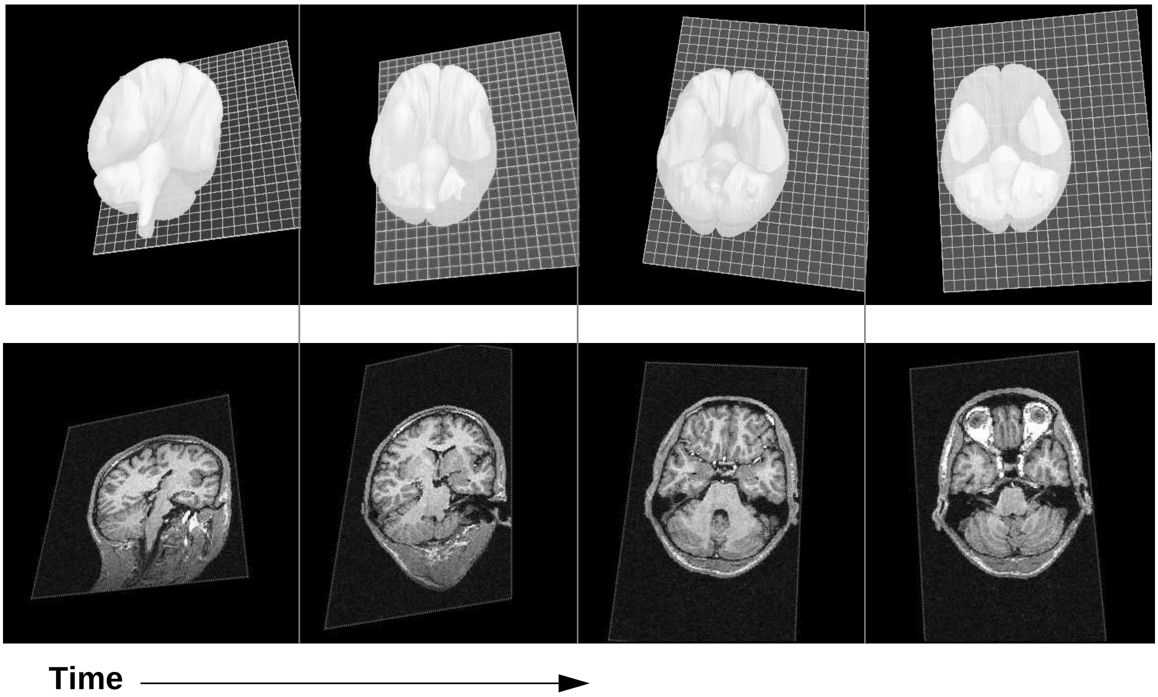


Figure 3.5 User positioning the cutting-plane prop along an oblique plane.

To provide visual correspondence, the virtual representation of the cutting-plane prop mirrors all six degrees-of-freedom of the physical tool. But several of these degrees-of-freedom do not affect the cross-section of the object, because (mathematically) the resulting plane has only four degrees of freedom. For example, rotation about the axis normal to the cutting-plane does not affect the cross section. Similarly, if the tool is moved left-to-right or front-to-back in the current plane, this does not affect the resulting plane equation. In this regard, the cutting-plane prop acts a bit like a flashlight: the virtual plane is much larger than the physical cutting-plane prop, so when one holds the input device to the side of the doll’s head, on the screen the plane still virtually intersects the brain, even though the two input devices don’t physically intersect.

The question of exactly how many degrees-of-freedom the user is manipulating at once is somewhat ill-defined. It can be as many as 12 or as few as 8, depending on how one likes to count. As mentioned above, the visual representation of the plane moves in 6 degrees of freedom, but the resulting cross-section has only 4 degrees of freedom. The doll's head has at least 4 degrees of freedom (rotation plus the zoom factor), but in fact all 6 degrees of freedom of the doll's head are used to compute the mapping of input device motion to virtual object motion (as described further in section 3.6.2). Thus, all 12 degrees of freedom from both input devices influence the display, but mathematically only 8 degrees of freedom are important.

The virtual representation of the cutting-plane prop is a semi-transparent rectangle. The transparency helps users to acquire a desired target: it provides a simple occlusion cue while maintaining the context of what is in front of or behind the plane [188]. This task is much more difficult if the plane is opaque.

3.5.3 Indicating surgical paths with a trajectory prop

The *trajectory selection prop* is a stylus-shaped tool (*fig. 3.6*) that allows the surgeon to specify 3D vectors and points. Moving the trajectory prop relative to the head prop specifies the position and orientation of a cylindrical virtual probe relative to the polygonal brain model. In previous work, Chung has implemented an interface for a similar task (radiotherapy treatment planning) using a head-mounted display, but his results were inconclusive: using a head-mounted display to select the trajectory of the radiotherapy beam did not have clear task performance advantages over hand-guided rotation, which is more similar to my approach.

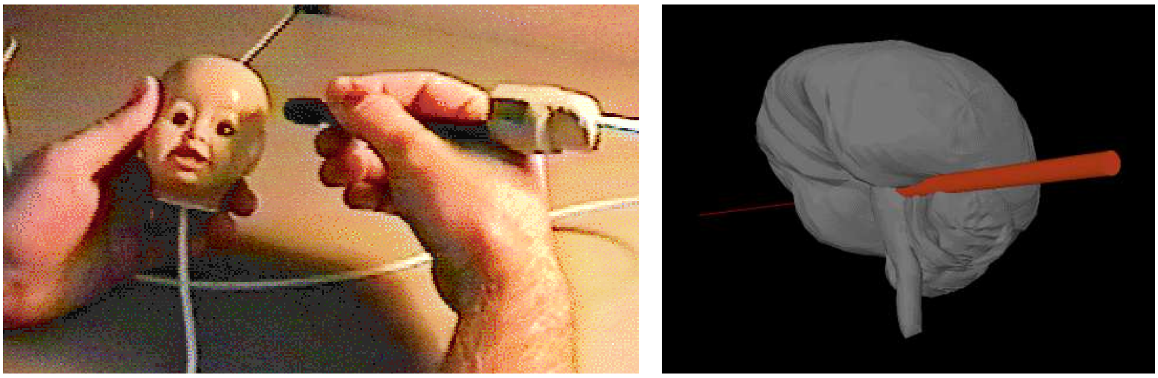


Figure 3.6 User selecting a trajectory.

In neurosurgery, a trajectory is defined as a three-dimensional path from the exterior of the head to a surgical target inside the brain. A *linear* trajectory is adequate for most cases, but occasionally a *nonlinear* trajectory is required to avoid vasculature or healthy brain tissue. Our laboratory's surgical planning software does not currently support nonlinear trajectories, nor does the props-based interface.

A linear trajectory consists of a target point inside the brain and a vector to that point. The trajectory prop indicates the vector by its orientation relative to the head prop. The target of the trajectory is indicated by the intersection of a ray cast from the virtual probe and the brain model's surface. Points which lie on the interior of the brain model can be selected by first bisecting the volume with the cutting plane to expose the contents of the volume, and then selecting a point on the exposed surface. Note that in this case the plane not only exposes the interior of the data, but it also expresses constraint of the point indicated by the trajectory prop to a plane, without requiring an explicit mode to do so.

3.6 Two-handed interaction

Guiard proposed that humans use the preferred and nonpreferred hands to control frames of reference which are organized in a hierarchy. For right handers, the left hand specifies a *base frame of reference* relative to which the right hand expresses a second *active frame of reference*. The props interface assigns the base frame of reference to the doll's head and the active frame of reference to the cutting plane. Since the neurosurgeon's task is to specify a cutting plane relative to a particular desired view of the brain, the interface's frames-of-reference assignment matches the surgeon's mental model of the task, resulting in an easily understood two-handed interface.

During the early stages of the interface design, I felt some concern that users might not be able to effectively control the four degrees-of-freedom provided by the doll's head using only their "weak" hand. In practice, however, informal evaluations have confirmed that the non-dominant hand is well suited to this task. The nonpreferred hand is not merely a poor approximation of the preferred hand, but can bring skilled manipulative capabilities to a task [94], especially when it acts in concert with the preferred hand.

Traditionally, two-handed input has been viewed as a technique which allows the user to save time by performing two sub-tasks in parallel [27]. For 3D input, however, two-handed interaction may be of even greater importance. Most everyday manipulative tasks, such as peeling an apple or cutting a piece of paper with scissors, involve both hands [67]. Previous work [76] has also shown that people often naturally express spatial manipulations using two-handed gestures. Based on my user observations and design experience, I can suggest some additional potential advantages for using two hands in 3D:

- Users can effortlessly move their hands relative to one another or relative to a real object, but it requires a conscious effort to move a single hand relative to an abstract 3D space. Note that the designers of 3-Draw [141] have made a similar observation.

- Use of two hands provides physical support. One-handed three-dimensional input in empty space can be fatiguing, but if the hands can rest against one another or against a real object, fatigue can be greatly reduced.

- Using two hands can help users to transfer existing skills for manipulating objects with two hands to the operation of a computer interface.

- The user can express complex spatial relations in a single two-handed action. Users can manipulate the interface props with two hands to specify a cut relative to a particular brain orientation in a single gesture. Not only does this make the interaction parallel (as opposed to being sequentially moded), but it also results in an interface which more directly matches the user's task. This issue is discussed further in the following section.

3.6.1 Two-handed input and the task hierarchy

One might argue that using two hands to operate the interface only adds complexity and makes an interface harder, not easier, to use-- after all, it is difficult to "rub your head and pat your stomach at the same time." Rubbing your head and patting your stomach are independent subtasks which bear no relation to one another. There are many compound tasks, however, such as navigation and selection in a text document or positioning and scaling a rectangle, which users perceive as *integral attributes* [91] that are aspects of a single *cognitive chunk* [28]. When designed appropriately, a two handed interface for *integral* compound tasks does not necessarily impose a cognitive burden, and can help users to reason about their tasks.

Figure 3.7 illustrates how the props-based interface simplifies the compound task of selecting a cutting-plane relative to a specific view of the polygonal brain. Cutting relative to a view consists of two sub-tasks: viewing and cutting. Viewing can further be subdivided into orienting the brain and specifying a zoom factor, and so forth. At the lowest level, there are ten separate control parameters (yaw, pitch, roll, and zoom for the view; x, y, z, yaw, pitch, and roll for the cutting tool) that can be specified. In a sliders or knob-box implementation of this interface, the user would have to perform ten separate one-dimensional tasks to position the cutting plane relative to a view, resulting in a user interface which is nearly impossible for a surgeon to use. Using the props with both hands, however, reduces this entire hierarchy into a single transaction (cognitive chunk) which directly corresponds to the task that the user has in mind. As a result, the user perceives the interface as being much easier to use.

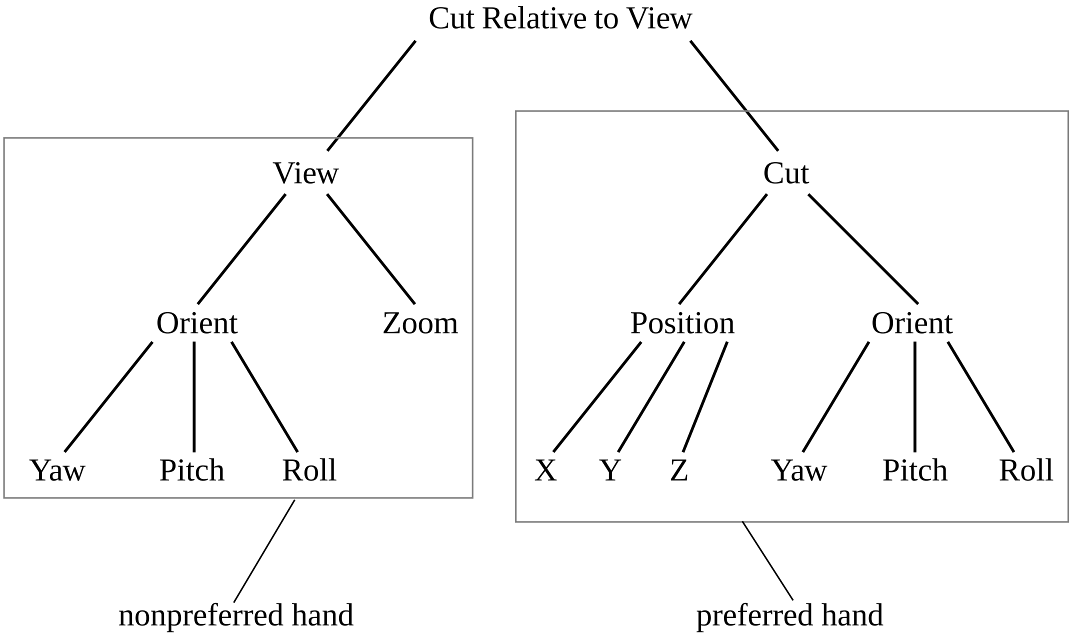


Figure 3.7 Task hierarchy for selecting a cut relative to a specific view.

This framework, suggested by Buxton's work on chunking and phrasing [28], is useful for reasoning about the differences between one and two-handed interfaces. With a unimanual interface, *View* and *Cut* would always have to be performed as purely sequential subtasks. There is also the need to switch back and forth between viewing and cutting, so this implies a third sub-task, that of changing modes. Changing modes might involve acquiring another input device, speaking a voice command, or moving the mouse to another region of the screen -- the exact interface is irrelevant to this discussion -- but all of these mode switching techniques take a non-zero amount of time. This process can be modelled as a simple state diagram (*fig. 3.8*).

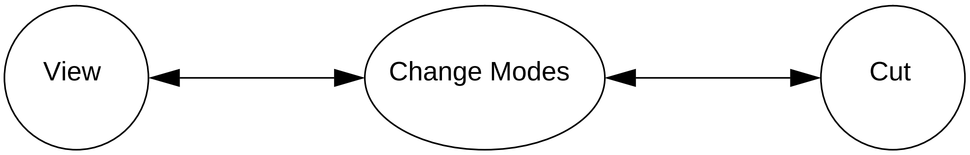


Figure 3.8 State diagram for unimanual subtasks.

A two-handed interface changes the syntax for this task. Under bimanual control, a new meta-task with a single *Cut Relative to View* state becomes possible. The simultaneous *Cut Relative to View* task is not the same thing as the serial combination of the sub-tasks. The simultaneous task allows for hierarchical specialization of the hands, and there is no cost (or need) to switch between *View* and *Cut* subtasks. Thus, there is the potential for bimanual control to impact performance at the cognitive level: it can change how users think about the task. Since the *View* and *Cut* subtasks can be integrated without cost, this encourages exploration of the task solution space. And since the user never has to engage in a *Change Modes* sub-task, there is no possibility for this extraneous sub-task to interfere with the user's primary goal of viewing and cutting. Chapter 8, "The Bimanual Frame-of-Reference," further explores the hypothesis that bimanual control can impact performance at the cognitive level.

3.6.2 The natural central object

As mentioned earlier, the polygonal brain is constrained to appear at the center of the screen. The brain is the natural central object of the manipulation and exploration supported by the interface. This design decision interacts with two-handed control and leads to an interaction technique which does not strictly copy physical reality, yet nonetheless seems quite natural. The key design principle is *not* to maintain a direct 1:1

correspondence between physical and virtual object motion, but rather it is to maintain the nonpreferred hand as a dynamic frame-of-reference, as described below.

Users do expect the real-world relationship between the props to be mirrored by their on-screen graphical representations. Simplifying control of the virtual brain by centering it on the screen, however, requires a software mapping of its real-world position to its centered position by constraining the x and y translations (note that no such mapping is required for the *orientation* of the prop). Define the position of the head prop in the real world as (H_{Rx}, H_{Ry}, H_{Rz}). If the center point of the screen is defined as (C_x, C_y), then the *virtual* constrained head position is given by (H_{Vx}, H_{Vy}, H_{Vz}) = (C_x, C_y, H_{Rz})[1].

When the user moves the cutting plane prop relative to the doll's head, the user expects to see this relative motion mirrored on the screen. This implies that the virtual representation of the cutting plane prop is drawn relative to the virtual position of the head prop. That is, the virtual position of the plane is equal to the virtual position of the *head* plus the real-world offset between the head prop and the cutting plane prop. Define the position of the cutting plane prop in the real world as (P_{Rx}, P_{Ry}, P_{Rz}). The offset is:

$$(\Delta_x, \Delta_y, \Delta_z) = (H_{R_x}, H_{R_y}, H_{R_z}) - (P_{R_x}, P_{R_y}, P_{R_z})$$

The virtual position of the plane is then given by:

$$(P_{V_x}, P_{V_y}, P_{V_z}) = (H_{V_x}, H_{V_y}, H_{V_z}) + (\Delta_x, \Delta_y, \Delta_z)$$

This mapping results in the following non-correspondence artifact: if the user holds the cutting-plane prop still and translates only the head prop, the polygonal brain will

---

[1] The system actually performs a log transformation on the zoom factor. Thus H_{Vz} is mapped to a function of log(H_{Rz}). As long as a corresponding log transform is also applied to the Z coordinate of the cutting plane, P_{Rz}, this log transform does not affect the mapping described in this section.

remain centered and the virtual plane will move in the *opposite* direction. This violates the generally accepted design principle that an interface should always maintain a direct 1:1 correspondence between physical and virtual object motion. But it adheres to the design principle that the object in the nonpreferred hand (the doll’s head) should form the base frame of reference. In hundreds of informal user trials, I have found that users almost never discover this artifact, because they typically hold and orient the head prop in a relatively stable location while moving the cutting plane prop relative to it. The net effect is that the interaction behaves as users expect it would; the mapping is the software embodiment of Guiard’s principle that the nonpreferred hand sets the frame of reference while the preferred hand articulates its motion *relative to* the nonpreferred hand.

Centering the reference object also has some other subtle effects on the interaction techniques and resulting user behavior. Since the nonpreferred hand now defines a dynamic frame of reference relative to which all manipulation occurs, this means that the user is not forced to work relative to the screen itself or relative to some center point within the environment, as is required by unimanual desk-top 3D interfaces [46][111]. Users are free to shift their body posture, to hold their hands on the desk surface, or to hold them in their laps. There is also no need for a “homing” or re-centering command to move the center point, since the nonpreferred hand automatically and continuously performs this function just by holding the doll’s head.

In a scenario where the user is working with multiple objects, and not a single natural central object as is the case for neurosurgical visualization, I speculate that the natural central object technique described above could still be used if the notion of a *current central object* were introduced. For example, the user might hold a clipboard in the nonpreferred hand that represents the current object of interest. Whenever the user selects a new object for manipulation from the environment (using an object selection technique suited to the application), the selected object might fade in and appear on the clipboard.

When the user finishes manipulating an object, the object could be detached from the clipboard, saved for later use, or perhaps replaced by the next object for manipulation.

3.7 Interactive volume cross-sectioning

There are several possibilities to consider for presenting an interactive update of the volume cross-section. The data from the volume cross-section could be superimposed on the polygonal view of the brain model, allowing the slice data to be seen directly in the context of the 3D object. The slice data could also be shown in a separate window off to the side of the polygonal graphics display area, which removes the slice from the context of the 3D object, but which thereby allows for a “map view” or “bird’s eye view” of the slice from directly above. The selection of an appropriate technique depends on the specific task as well as the practical constraints of what can be implemented with real-time interaction.

3.7.1 Frames-of-reference for the cross section display

In the original concept of the interface, the surgeons wanted to interactively clip away a portion of the object and paste the volumetric cross-section data on to the resulting capping polygon, so that the cross-section would always be seen directly in the context of the polygonal brain model. Without texture mapping hardware, the implementations I attempted could only render the volume slice in context at a maximum of about 2-3 frames per second, compared to about 15-18 frames per second for a display in a separate window. Thus I pursued a separate, out-of-context display approach.

The separate display can take the slice data that would have been superimposed on the polygonal model and draw it in a separate window, giving a *perspective view* of the slice data, or alternatively the previously mentioned *map view* from directly above could be used. Figure 3.9 compares these approaches by showing how the separate cross-section display behaves for the perspective view and map view techniques as the user changes the view of the polygonal brain over time.

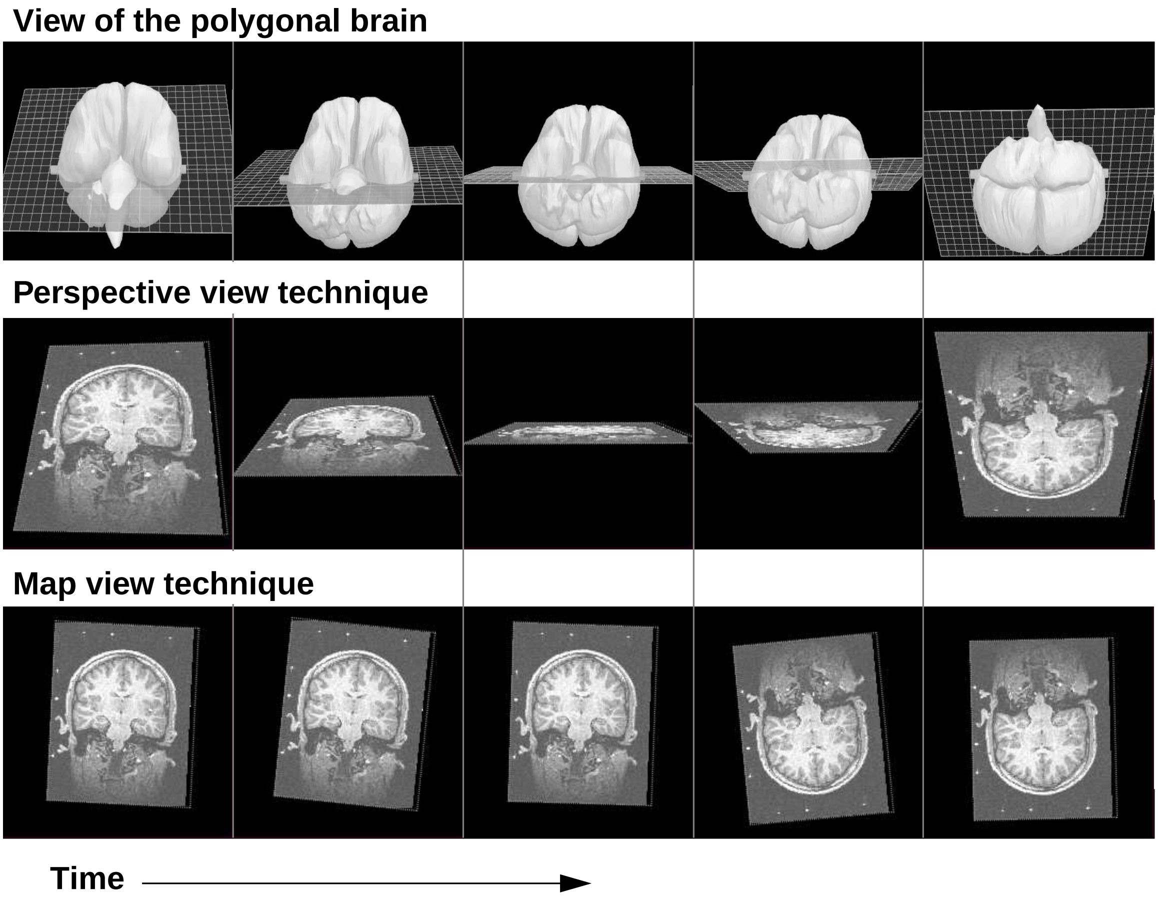


Figure 3.9 Comparison of perspective and map views of the cross-section data.

As seen in the figure, when the user tilts or moves the plane, the perspective view changes accordingly. Thus, if the user is primarily visually attending to the cross-section display (and not to the 3D view of the polygonal objects), the perspective view technique provides useful visual feedback of the motion of the plane. However, if the user holds the plane so that it is seen roughly edge-on in the 3D view of the polygonal objects, the perspective view conveys essentially no information. This imposes certain restrictions on how one can hold the cutting plane prop: it has to be held roughly vertical so that the cross-section data can be seen during interactive manipulation. Some users initially find this requirement to be confusing, though it is easy to adapt to.

The map view technique (*fig. 3.9, bottom row*) does not impose any restrictions on the orientation of the cutting plane prop relative to the view-- even when the user holds the plane edge-on, he or she can still see the data resulting from the cross-section. But the map view technique lacks the motion cues provided by the perspective view; as seen in figure 3.9, the map view essentially does not change at all over the first three frames of the time sequence, despite a large angle of rotation. Thus, each technique can be considered to have its own strengths and weaknesses. For the dynamic view of the cross-section, the current version of the props interface defaults to the perspective view and has the map view available as an option. When the user selects a plane (by releasing the clutch button on the cutting-plane prop), a separate static view of the cross-section is saved for subsequent 2D manipulation, and this static view is always displayed using the map view technique.

3.7.2 Texture mapping hardware

Texture mapping hardware allows for the possibility of in-context viewing of the cross-section data, as seen in the prototype implementation of figure 3.10. Standard texture mapping hardware does not help to calculate the cross section itself; this still must be calculated in software to compute the initial texture map. Unfortunately, selecting an arbitrary plane through a volume is probably a worst case for traditional texture mapping. Loading a texture map to texture memory is an expensive operation[1]. Texture mapping hardware typically assumes that most textures are known when an application begins, and that they will be used repeatedly, so the system performs pre-computations and creates data structures when a texture is first loaded. But when users move the cutting plane prop to interactively select a volume cross-section, the texture map changes every frame and the exact same plane is not usually accessed more than once.

---

[1] This is changing on high-end graphics workstations. The new SGI architecture is designed to provide high bandwidth to texture memory. Hewlett Packard is planning to make several improvements to their texture-mapping facilities based on feedback from my project.

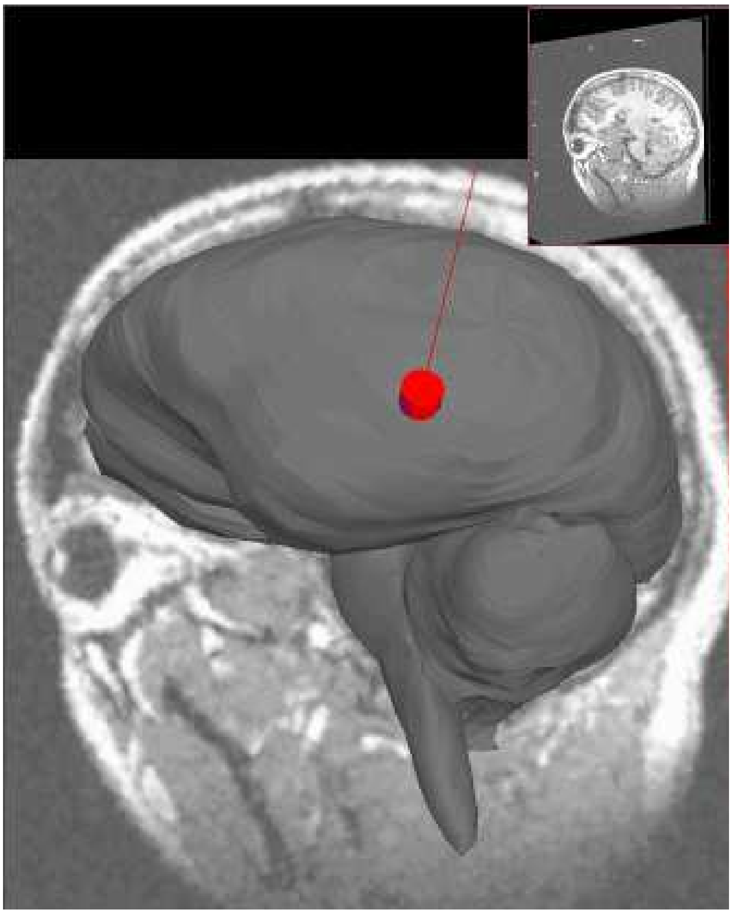


Figure 3.10 Texture mapping shows a slice in the context of the 3D brain model.

Thus, commonly available texture mapping hardware still does not allow an in-context presentation of the cross-section while the cross-section itself is changing[1]. However, once the user selects a cross-section, this cross-section can be turned into a texture map and integrated with the view of the polygonal objects. With the cross-section available in context, users find it much easier to select a subsequent plane which passes through a target revealed by a previous cut. For example, it would be easy to select a cut which passes through the lens of the eye already revealed in figure 3.10.

---

[1] Some high-end SGI graphics supercomputers do allow definition of “3D texture maps” or “voxel maps” which can perform this function in hardware.

3.7.3 Disappearing object problem

When clipping away the polygonal model is enabled, it is possible to position the cutting plane such that it slices away the entire object (*fig. 3.11, top*), which leads to confusion. One possible solution is to draw a wireframe wherever the object has been cut away (*fig. 3.11, bottom*). This solution works well with a separate cross-section display. With an integrated display using texture mapping hardware, however, this solution is not ideal because the wireframe obscures the cross-section.

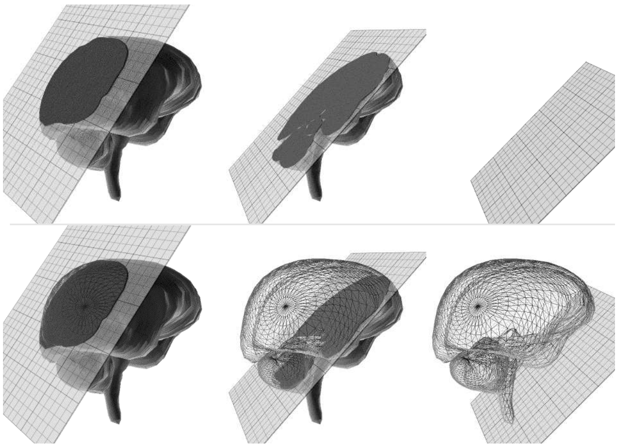


Figure 3.11 The disappearing object problem and wireframe solution

3.8 Clutching mechanisms

A mechanism is often needed to tell the computer to “stop watching” a particular prop. This allows the surgeon to freeze the image in a desired configuration and put down the props. In the original design, a foot pedal was provided to “clutch” the head prop and a thumb button to clutch the cutting-plane prop. The foot pedal behaved like a gas pedal: the user held it down to move. Similarly, the cutting-plane prop only allowed motion while the thumb button was held down. Sellen [145] has shown such tension can reduce mode errors.

I have also experimented with voice control of the clutch. Saying *move {head | plane}* enables motion, while saying *stop {head | plane}* disables motion. Since the user is engaged in a real-time manipulation task, the time to speak and recognize a voice command causes an irritating delay. It is not clear if this problem would persist with a more sophisticated voice recognizer than the low-cost unit [174] which I used on an experimental basis, but I expect it would; the delay introduced by speaking the command might itself prove intolerable. Under some conditions voice input can also interfere with short term memory [96], which poses another possible difficulty.

Further user testing has suggested that the interface is easiest to use when there *is no clutch* for the doll's head. In the current design, the doll's head is always allowed to move. Freezing the polygonal brain in place seems like a useful thing to do, but again the most important design principle is to maintain the nonpreferred hand as a dynamic reference for the action of the preferred hand. If the doll's head is "clutched" so that it cannot move, it is no longer useful as a reference and the preferred hand again must move relative to the environment. I have watched many users clutch the head and then become confused as they subconsciously begin to move their nonpreferred hand to aid the action of the preferred hand, only to have no effect. After gaining some experience with the interface, users generally tend to constantly hold down the footpedal for the doll's head anyway. This questions whether the head clutch serves any real purpose other than to initially confuse users.

The interface does provide the ability to generate a detailed volume rendering of the head; this acts very much like a clutch in the sense that it generates a still image "snapshot," but it does not have the side-effect of interfering with further manipulation. With this capability, there is no apparent need for the head clutch except in unusual circumstances, such as when trying to photograph the interface, or when a user without the ability to use both hands wishes to operate the interface. For such situations, it is still possible to use a footpedal for clutching the doll's head, but this behavior is not enabled by default.

3.9 Touchscreens for hybrid 2D and 3D input

The 3D interface props excel for 3D manipulation, but when 2D tasks such as loading a new patient from an image database arise, there is an awkward pause in the human-computer dialog while the user must put down the props to move the mouse or to use the keyboard. The problem lies in a 3D input versus 2D input dichotomy: some tasks are best done in 3D, others are better suited to 2D, and users need an intuitive and consistent mechanism for switching between the different styles of input. Users are distracted from the focus of their work because they must decide which device to acquire for a given input task.

To address this shortcoming, we added a touchscreen sensor to the monitor used with the interface props. This hybrid interface combines 3D input with more traditional 2D input in the same user interface. Note the ergonomic facility with which a touchscreen can be used: the surgeon can move in 3D using the props; then, without having to put the props down, the surgeon can reach out and touch the screen to perform 2D tasks, since the hand is sufficiently free to extend a finger or knuckle (*fig 3.12*). This provides a consistent input medium for both 2D and 3D tasks, since the user always interacts gesturally with objects in the real environment: one interacts gesturally with the props to perform 3D operations; one interacts gesturally with the touchscreen to perform 2D operations.

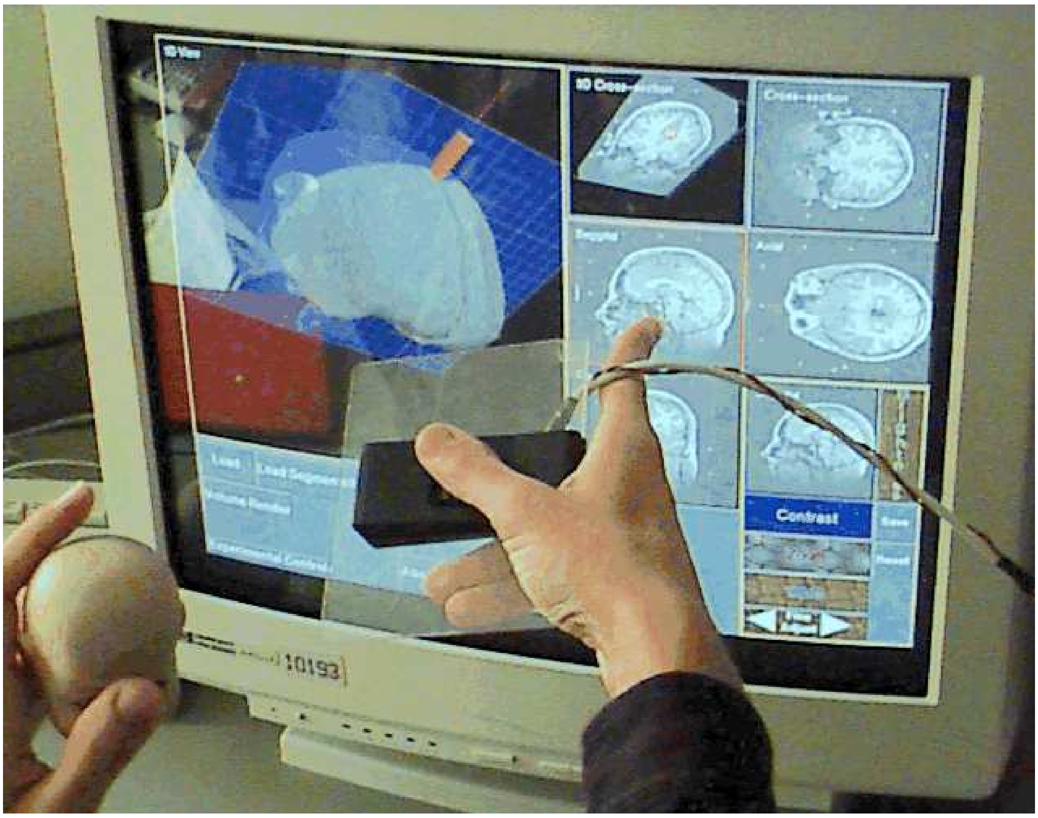


Figure 3.12 User employing the touchscreen in combination with the props.

3.9.1 Previous techniques for hybrid input

Most previous work in 3D interaction has either ignored the problem of adding 2D input or has integrated it in an ad hoc manner, but there are a few notable exceptions. Feiner [55] has constructed a system which combines a 3D augmented reality head-mounted display with a standard 2D desktop display. This allows the user to see 3D virtual objects arranged around his or her head, but does not block out the user's view of the real world, allowing standard mice, keyboards, and CRT displays to be used.

Shaw's approach [149] is to put as many 2D capabilities as possible into the 3D environment. Shaw's system for polygonal surface design and CAD tasks uses two small hand-held "bats" for input devices. Each bat is a six-degree-of-freedom magnetic tracker augmented with three small push-buttons, which can be pushed to indicate various discrete commands. Shaw's interface also incorporates a ring menu [111], which is a pop-up menu which can be accessed in 3D by rotating the bat. The menu appears as a ring surrounding the 3D cursor on the screen; by rotating the bat, the user can select different items which are highlighted in the ring. This works reasonably well when only a few items (no more than about 15-20) are displayed in the ring.

The Polyshop system [1] (*fig. 2.5 on page 18*) uses a drafting board as a 2D constraint surface. Since the user is wearing a head-mounted display, graphics can appear to lie on a virtual representation of the drafting board, allowing it to be used in a manner analogous to a touchscreen. The drafting board does not actually detect finger touches; it must rely on hand position sensing to estimate when the user is trying to work with the 2D surface. Thus, while promising, this approach lacks the reliable feedback of knowing that an action occurs if and only if one's finger physically touches the display surface.

Work at Brown University [43][77][78] has looked at ways of using mouse-controlled "3D Widgets" for 3D interaction (*fig. 2.1 on page 15*). Here the problem of combining the 3D and 2D interfaces is obviated, as the mouse is used for all input tasks. Similarly, Brown's Sketch system [186] (*fig. 2.3 on page 17*) has demonstrated that a 2D interface based on some simple gestures and well-chosen heuristics can be very powerful for sketching 3D objects or scenes.

3.9.2 Description of the touchscreen interface

The surgeon uses a combination of 3D and 2D manipulation facilities when planning a surgery. There is not a distinct 3D interaction phase followed by a distinct 2D interaction phase. Rather there is a continuous dialog which combines 2D and 3D visualization tools to accomplish the surgical plan. Using a mouse for 2D input does not facilitate this style of dialog. In my experience surgeons are hesitant to use mice in the first place; when manipulating the props in addition to the mouse, the surgeon would typically rely on someone else to move the mouse.

The touchscreen graphical user interface (GUI) divides the screen into a set of tiles (*fig 3.13*) which contain different views of the same volumetric data set. These tiles are interchangeable; for example, to increase the screen real estate for the sagittal view, the user can drag it with his finger into the large area on left side of the screen. The region in the lower right hand corner of the screen acts as a general purpose control panel for all tiles. When the user touches a tile, it becomes selected and a miniature copy of the tile appears in the control panel. The control panel widgets can then be used to interactively manipulate the miniature copy, and after a brief pause, the changes are propagated to the original tile. The control panel includes controls for image contrast and brightness, zooming and panning, browsing or precisely stepping through parallel slices, saving the current image, and resetting the default viewing parameters.

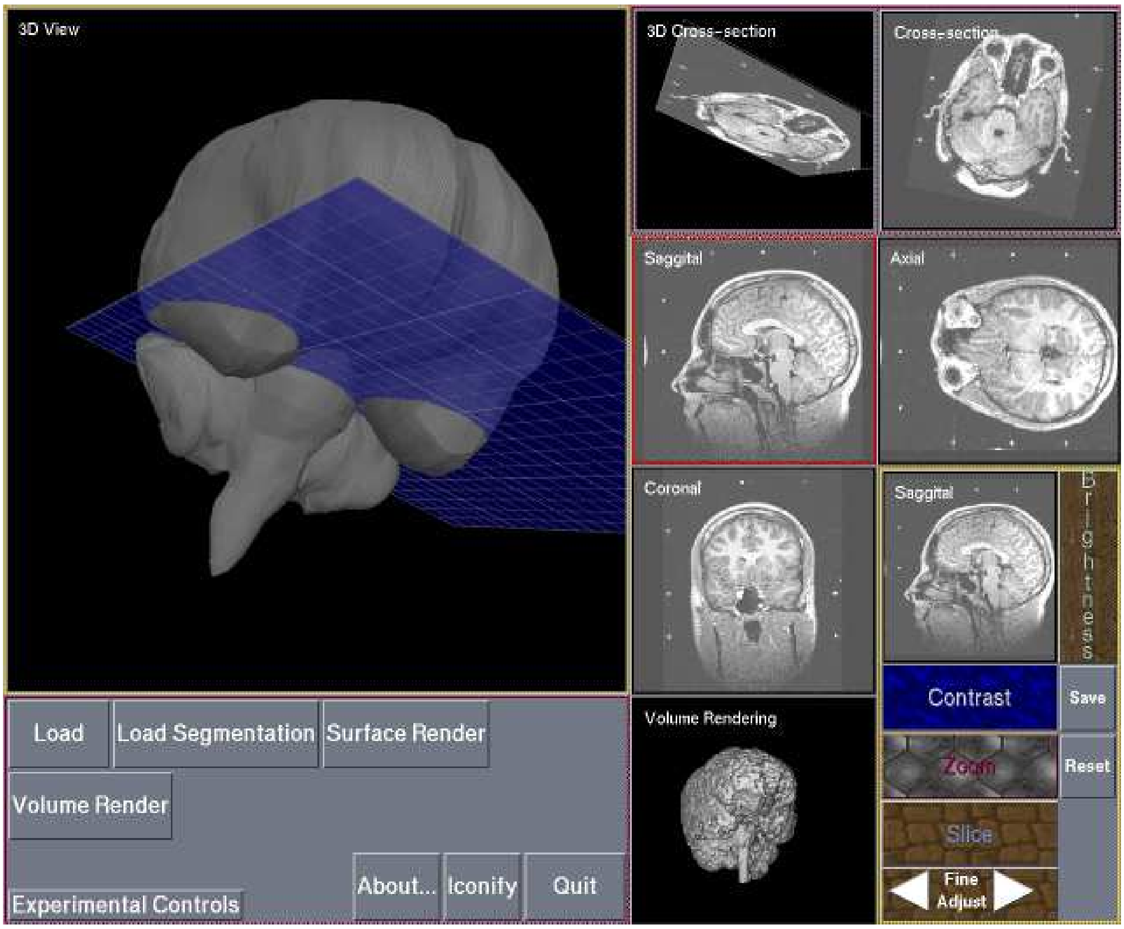


Figure 3.13 Touchscreen graphical user interface for use with the 3D props.

The brightness, contrast, zoom, and slice navigation touchscreen controls in the control panel (*fig. 3.14*) were suggested by the physical thumb-wheels which are found on many 2D medical image viewing consoles. When the user touches and drags the touchscreen thumb-wheels, the background textures slide up and down, giving immediate visual feedback. The touchscreen thumb-wheels were designed to be used without looking directly at them, because the user is typically focusing on the image being modified, and not the widget itself. The moving background textures can be seen even with peripheral vision, and so meet this “eyes-free” requirement effectively. Using a standard scrollbar on a touchscreen completely fails in this regard, since the user’s finger occludes the thumb of the scrollbar. An experimental implementation of the touchscreen thumb-wheels with nonspeech audio feedback has suggested that the technique would be even more effective with appropriate audio feedback.

Figure 3.14 Close-up of touchscreen control panel.

Rather than having constraint modes for the 3D devices, I found that users were more comfortable with expressing constraints using the naturally constrained dialog afforded by the touchscreen. Thus, the tiles which show the standard sagittal, axial, and coronal slice orientations act as subtle constraint modes; all operations on these tiles are naturally constrained to the appropriate axis of the volume. Similarly, once an oblique slice has been selected with the props, this becomes a tile (seen in the upper right of figure 3.13) which expresses constraint along the normal of the currently selected oblique cutting plane.

The touchscreen provides access to a couple of other important facilities. First, at any time, the surgeon can generate a high-quality volume rendering (*fig. 3.15*), which takes approximately five seconds to generate, by touching a button labelled *Volume Render* in the interface. The volume rendering can show the cortical surface at the full resolution of the MR scan (rather than at the approximate resolution of the polygonal model), and as such is essential to the surgeon's end goals. Second, to load a new data set, the surgeon presses a button labelled *Load* which brings up a database browsing dialog (*fig. 3.16*). This provides facilities for browsing the database, picking patients from an index, and selection of a specific study for a patient. The interface maintains a cache of recently accessed patients, as the surgeon may need to frequently switch between a small number of cases.

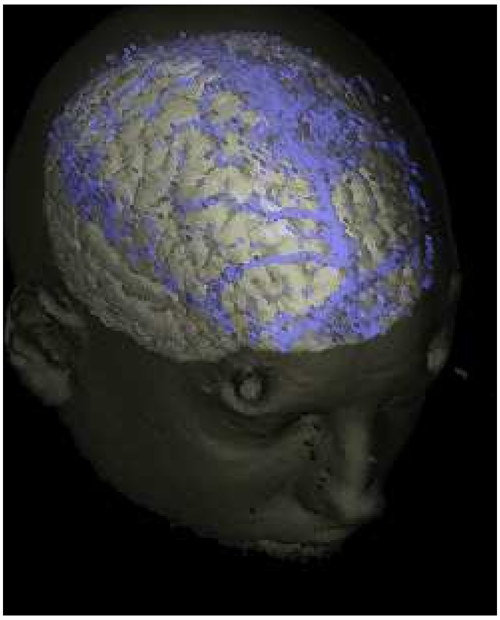
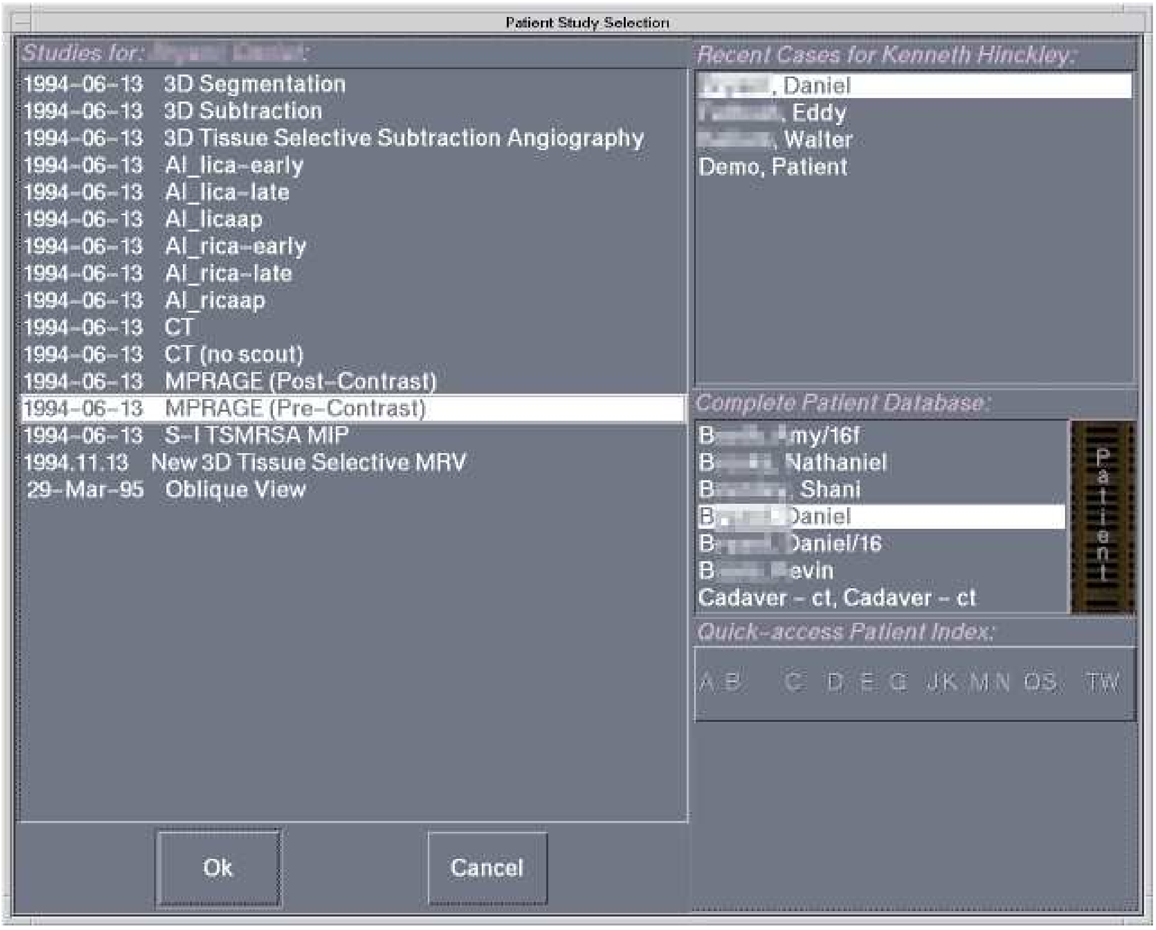


Figure 3.15 Volume rendering showing the brain and skin surface.



Studies for: [redacted]:

1994-06-13 3D Segmentation
1994-06-13 3D Subtraction
1994-06-13 3D Tissue Selective Subtraction Angiography
1994-06-13 AI_lica-early
1994-06-13 AI_lica-late
1994-06-13 AI_licaap
1994-06-13 AI_rica-early
1994-06-13 AI_rica-late
1994-06-13 AI_ricaap
1994-06-13 CT
1994-06-13 CT (no scout)
1994-06-13 MPRAGE (Post-Contrast)
1994-06-13 MPRAGE (Pre-Contrast)
1994-06-13 S-I TSMRSA MIP
1994.11.13 New 3D Tissue Selective MRV
29-Mar-95 Oblique View

Recent Cases for Kenneth Hinckley:

[redacted], Daniel
[redacted], Eddy
[redacted], Walter
Demo, Patient

Complete Patient Database:

B[redacted]my/16f
B[redacted] Nathaniel
B[redacted], Shani
B[redacted]Daniel
B[redacted]Daniel/16
B[redacted]evin
Cadaver - ct, Cadaver - ct

Figure 3.16 Touchscreen interface for selecting patient images from a database[1].

Informal observation of users of the touchscreen and props hybrid interface suggest that touchscreens offer not only intuitive 2D input which is well accepted by physicians, but that touchscreens also offer fast and accurate input which blends well with 3D interaction techniques.

3.9.3 Limitations and proposed enhancements

While promising, the implementation of the hybrid interface concept has several shortcomings. The major usability problems which I encountered were related to the surface acoustic wave (SAW) touchscreen technology which our laboratory chose to use.

---

[1] To maintain patient confidentiality, the actual patient names in this figure have been blurred out.

Disadvantages (for this application) include parallax errors and a limitation on the type of materials which can be used to touch the screen. Parallax errors result from a gap between the touch screen and the actual monitor; these errors make the touchscreen difficult to use if the user is standing or not directly in front of it. Further, since it works by detecting sound waves, the SAW touchscreen can only detect the presence of soft materials such as a finger or an eraser tip. This means that the interface props cannot be used to directly touch the screen, something which users would naturally and repeatedly try to do. I attempted to implement “soft” versions of the props, but it was difficult to find durable materials that were soft enough to do the job, and none of them worked as effectively as a finger.

A resistive membrane touchscreen may provide a more suitable technology for this application. Any type of material can be used to touch resistive membrane touchscreens, and since they consist of a film directly on the monitor glass, parallax errors are reduced. They also respond to a lighter touch than SAW touchscreens, making “drag and drop” style interactions easier to perform. Potential disadvantages include lower transmissivity (they attenuate more of the light coming from the screen) and the inability of currently available commercial models to detect pressure.

Liang [112] has suggested using a graphics tablet, such as those manufactured by Waccom [175], in a similar capacity, though the idea has not been tested with an implementation. Although a tablet would not have the ability to act as a display, this may provide an alternative approach for some applications.

Another promising avenue for exploration is a multimodal combination of voice, touchscreen, and two-handed 3D manipulation. Voice recognition would allow the user to perform discrete commands when both hands are busy manipulating the props, or it could be combined with touch to disambiguate a gesture, as demonstrated by the “Put that There” system [16].

3.10 Informal evaluation: notes on user acceptance

Without the cooperation and close collaboration of real neurosurgeons, this work would not have been possible. The design of the interface has been an iterative process where I slowly learned what interaction techniques neurosurgeons could use to accomplish the goals they had in mind, and the surgeons learned what possibilities and limitations the computer offers as a surgical planning and visualization tool. Watching a surgeon become frustrated with an inadequate interface design that seems “intuitive” to computer scientists is a very strong motivation to discover how to improve the interface.

3.10.1 User observations

The most valuable tool for evaluating and improving the interface design has been informal observation of test users. This section focuses on my user observations of neurosurgeons, but I should point out that I have tested the interface with many other types of physicians who work with volumetric data, such as neurologists, cardiologists, and radiologists. I have also performed many informal user observations during demonstrations to the general public and during debriefing sessions with experimental subjects.

According to my collaborators in neurosurgery, there are currently about 4,000 neurosurgeons practicing in the United States. Over the history of the project about 50 neurosurgeons have tried the props-based interface during informal observation sessions[1]. Surgeons who tried the system typically included attending neurosurgeons or neurosurgical residents from UVA, prestigious visiting neurosurgeons from other institutions, and interviewees for neurosurgical residencies. Some additional neurosurgeons tried the system during demonstrations at academic neurosurgery conferences.

The methodology for testing was simple but effective. On some occasions I set up the system in the hospital or another location convenient for the surgeons, but more often

---

[1] Only about 5 neurosurgeons have tried the touchscreen interface, since it is a fairly recent enhancement, and my research has since shifted away from this avenue.

surgeons would visit our laboratory for an informal demonstration. I almost always began by briefly showing the system to the visiting neurosurgeon. Many surgeons were eager to try it and would jump in themselves without needing an invitation. Then I would stand back and just watch the neurosurgeon operating the interface. If the surgeon seemed to be having some troubles, rather than immediately intervening with advice or suggestions, I would wait for the surgeon to ask a question or make a comment. In this way, I could understand a problem in the terms the surgeon was using to think about it. I found that a good question for stimulating discussion of the interface was: “Can you show me what you did during your most recent surgical procedure?”

Surgeons usually would offer opinions, advice or suggestions for the interface on their own without any need for prompting. This would start a discussion where I would also ask the surgeon about particular problems I had observed, what they saw as capabilities or limitations of the interface, and how the prototype interface might be augmented to become a clinical tool. Another question which helped me to see what the interface looked like from *the surgeon’s perspective* was: “How would you explain the interface you just saw to a colleague?”

Neurosurgeons have been very enthusiastic about the props-based interface. All of the neurosurgeons who have tried to use the interface were able to “get the hang of it” within about one minute of touching the props; many users required considerably less time than this. This clearly demonstrates that with a cursory introduction, neurosurgeons who have never before seen the interface can rapidly apply their existing skills for manipulating physical objects with two hands, and can understand and use the interface without training.

In addition to the neurosurgeons, approximately 50 physicians from other specialities have tried the interface. The overall response has been similar to that of neurosurgeons, particularly from specialists who commonly deal with volumetric MRI data, such as data representing the knees, shoulders, or heart. From talking with these physicians, it is clear that oblique plane visualization problems commonly occur in these fields as well; for example, when visualizing the knee, oblique cuts along the ligaments are clinically valuable to assess the severity of an injury.

One exception has been with radiologists, where the response has been more variable. Radiology deals primarily with diagnosis, not planning surgical interventions, and furthermore radiological training emphasizes the *mental* 3D visualization of structures from 2D information. Thus, the interface does not support tasks which are of interest to most radiologists, and therefore it does not represent a useful tool to most of these individuals. This is not to say that all radiologists found the interface useless-- on the contrary, several radiologists with an interest in the issues posed by digital medical imaging have been quite enthusiastic.

Finally, I estimate that I have given or observed hands-on demonstrations to over 1,000 users from a broad sample of the general public, ranging from small children to elderly university benefactors. These informal demonstrations have been useful for testing some of my ideas on a larger sample of test users than is possible with physicians. Furthermore, these observations strongly suggest that people in general, and not just skilled and dexterous surgeons, can use both hands to perform 3D manipulation tasks.

People don't understand 3D. They experience it.

Ivan Sutherland

Chapter 4

Design Issues in Spatial Input

4.1 Introduction

A broad, scattered set of interface designs and experimental results influenced the design of the props-based interface described in the previous chapter. My experiences from implementing the system and performing user observations suggest some general issues which span this set of research. The present chapter integrates these issues into a survey of free-space 3D input techniques. This survey also suggests some general research areas which interface designers and researchers have only begun to explore and which are in need of further work.

The term *spatial input* refers to interfaces based upon free-space 3D input technologies such as camera-based or magnetic trackers [136][137], as opposed to desktop 2D or 3D devices such as the mouse or the Spaceball [161], respectively. In the literature, a wide variety of interfaces for manipulating three-dimensional objects have been described as "3D interfaces," but here the term *spatial input* distinguishes the class of 3D interfaces based upon free-space interaction.

The design issues presented in this chapter are not scientifically demonstrated principles of design or ready-to-go solutions. Rather they are issues to be aware of and some different approaches to try. I explore some of these design issues, particularly those involving two-handed interaction, in the context of formal experiments described in subsequent chapters of this document, but the other design issues are supported only by possibly unrepresentative user observations. Nonetheless, this chapter serves as a useful guide for the community of designers and researchers who wish to explore spatial input techniques.

4.2 Understanding 3D space vs. experiencing 3D space

Anyone who has tried to build a stone wall knows how difficult it is to look at a pile of available stones and decide which stone will best fit into a gap in the wall. There are some individuals, such as experienced stone masons, who have become proficient with this task, but most people simply have to *try* different stones until one is found that fits reasonably well.

In general, people are good at experiencing 3D and experimenting with spatial relationships between real-world objects, but they possess little innate comprehension of 3D space in the abstract. People do not innately *understand* three dimensional reality, but rather they *experience* it.[1]

From a perceptual standpoint, one could argue that our difficulty in building stone walls, and in performing abstract 3D tasks in general, results from a distinction between creating and manipulating images (or objects) versus mentally conjuring images and mentally transforming them. For example, the Shepard-Metzler mental rotation study [151] suggests that for some classes of objects, people must mentally envision a rigid body

---

[1] Ivan Sutherland suggested this distinction between understanding 3D and experiencing 3D in the Fall of 1993. Also, Fred Brooks included this idea in his 1988 review paper where he observes that “3D understanding is difficult” [19].

transformation on the object to understand how it will look from different viewpoints; that is, humans must perceive the motion to understand the effect of the transformation.

Previous interfaces have demonstrated a number of issues which may facilitate 3D space perception, including the following:

- Spatial references
- Relative gesture vs. absolute gesture
- Two-handed interaction
- Multisensory feedback
- Physical constraints
- Head tracking techniques

A spatial interface does not necessarily have to consider all of these issues to be usable. Rather, a designer should consider these issues as a set of approaches which might be applied to a given design problem.

4.3 Spatial references

Badler [8] describes an interface where a stylus is used to control the position of a virtual camera. One version of the interface allows the user to indicate the desired view of an imaginary object using the stylus. Badler reports that “the lack of spatial feedback [makes] positioning the view a very consciously calculated activity.”

Badler repeated the experiment with a real (as opposed to imaginary) object. He digitized a plastic spaceship and allowed the user to specify the virtual camera view of the corresponding wireframe spaceship by positioning and orienting the wand relative to the real-world plastic spaceship. With this single change, Badler's "consciously calculated activity" suddenly became "natural and effortless" for the operator to control.

In general, to perform a task, the user's perceptual system needs something to refer to, something to *experience*. In 3D, using a spatial reference (such as Badler's plastic spaceship) is one way to provide this perceptual experience. Ostby's system for manipulating surface patches [129] was a second early system to note the importance of spatial references. Ostby reported that "[locating] a desired point or area [is] much easier when a real object is sitting on the Polhemus's digitizing surface."

4.4 Relative gesture vs. absolute gesture

In Galyean's 3D sculpting interface [61], the user deforms a 3D model by positioning a single tracker in an absolute, fixed volume in front of a monitor. This leads to an interface which is not entirely intuitive. Galyean reports that "controlling the tool position is not easy. Even though the Polhemus pointer is held in a well-defined region, it is often difficult to correlate the position of the pointer in space with the position of the tool on the screen."

Compare this to Sachs's 3-Draw computer-aided design tool [141], which allows the user to hold a stylus in one hand and a palette in the other (both objects are tracked by the computer). These tools serve to draw and view a 3D virtual object which is seen on a desktop monitor. The palette is used to view the object, while motion of the stylus relative to the palette is used to draw and edit the curves making up the object.

3-Draw's use of the stylus for editing existing curves and Galyean's use of the "Polhemus pointer" for deforming a sculpture represent nearly identical tasks, yet the authors of 3-Draw do not report the difficulties which Galyean encountered. This difference may result from the palette-relative gesture employed by 3-Draw, as opposed to the abstract, absolute-space gesture required by Galyean's sculpting interface. As Sachs notes, "users require far less concentration to manipulate objects relative to each other than if one object were fixed absolutely in space while a single input sensor controlled the other" [141].

Thus, users may have trouble moving in a fixed, absolute coordinate frame. A spatial interface could instead base its interaction techniques upon relative motion, including motion relative to a spatial reference or the user's own body.

4.5 Two-handed interaction

Enabling the use of both hands can allow users to ground themselves in the interaction space; in essence the user's own body becomes a spatial reference. Regarding two-handed interaction in free space, Sachs observes that "the simultaneous use of two [spatial input] sensors takes advantage of people's innate ability--knowing precisely where their hands are relative to each other" [141]. For example, during informal user observations of a virtual reality interface, our user interface lab noted that users of two-handed interaction are less likely to become disoriented versus users who interact with only one hand [130]. Even when manipulating just a single object in 3D, using two hands can be useful and natural: in a classic wizard-of-oz experiment, Hauptmann [76] observed test subjects spontaneously using two hands for single-object translation, rotation, and scaling tasks.

Based on an analysis of human skilled bimanual action [67] Guiard has proposed an insightful theoretical framework and principals governing two-handed manipulative action, as discussed in section 2.6.1 on page 31. Note, however, that the application of Guiard's principles to bimanual interface design have not been formally demonstrated, and may also represent an incomplete set of conditions for usable two-handed interfaces. For example, Kabbash [95] describes a two-handed interface (the "palette menu") where the user moves an opaque menu using a trackball in the left hand and a selection cursor using a mouse in the right hand. Although this interface apparently conforms to Guiard's principles, Kabbash's results suggest that the palette menu interface may induce a cognitive load. Because the palette is opaque, it occludes the context of the working area, and this apparently causes users to be uncertain of what strategy to use when placing the palette. The transparent menu used by the two-handed ToolGlass [15] technique, which Kabbash also analyzed, does exhibit this problem.

4.5.1 Working volume of the user's hands

Guiard's observations of subjects performing writing tasks [67] as well as my own observations of the props-based interface [80] suggest that people tend to move their hands in a surprisingly small working volume. This volume is not only small, but also tends to move over time as the user changes body posture. For example, Guiard's analysis of handwriting tasks suggests that the writer tends to define an active volume relative to his or her non-dominant hand. Guiard also reports that "the writing speed of adults is reduced by some 20% when instructions prevent the nonpreferred hand from manipulating the page" [67]. This suggests that users of a spatial interface which requires movements relative to a fixed frame-of-reference in their environment may experience reduced task performance due to cognitive load, fatigue, or both.

4.6 Multisensory feedback

A key challenging facing spatial interaction is identifying aspects of the proprioceptive senses that designers can take advantage of when interacting in real space. Interacting with imaginary, computer-generated worlds can easily bewilder users; presumably, providing a wide range of sensory feedback might help the user to more readily perceive their virtual environment. Psychologist J. J. Gibson has long argued that information from a variety of feedback channels is crucial to human understanding of space [62].

Brooks [19] discusses interfaces which employ multisensory feedback techniques, including force feedback [21][90][120], space exclusion (collision detection), and supporting auditory feedback. I add *physical manipulation of tools with mass* to these techniques.

For example, in our user interface lab [173] we have experimented with a virtual reality interface for positioning a virtual flashlight using a glove, which users can use to grab and position the virtual flashlight. However, during public demo sessions, we found that users have inordinate difficulty grasping and manipulating the virtual flashlight using the glove. By replacing the glove with a tracked *physical* flashlight, we found that users could position the virtual flashlight with ease. For this application, physical manipulation of a flashlight worked well, while glove-based manipulation of a virtual flashlight was a disaster.

There are several factors which can contribute to ease-of-use for the physical manipulation paradigm:

- When utilizing glove-based input, the user must rotate the entire hand to indicate the rotation of a virtual object. But as Liang notes, “the hand has certain kinematic constraints. For example, it is far more easy to rotate something held by the fingers than to rotate the whole hand itself” [111].

- The mass of the tool can damp instabilities in the user’s hand motion. For example, surgeons are very particular about the *weight* of their surgical instruments, as the proper heaviness can help decrease the amplitude of small, involuntary hand tremors.

- A physical tool provides kinesthetic feedback, due to the tool’s inertia and the force of gravity.

- The physical properties of a tool suggest its use and constrain how the user can manipulate it. For example, a screwdriver affords rotation about its vertical axis while a wrench affords rotation about a horizontal axis. This type of haptic feedback would not possible if the rotational constraints were purely visual, as is the case with graphical 3D widgets [43].

4.7 Physical constraints and affordances

Physical constraints and affordances are widely used in industrial design (Norman [126] provides many examples). Software constraints are often useful, but they do have limitations: the user must understand the constraints and their feedback and then react appropriately. Using physical constraints can remove this problem and also lends support: users can *try* configurations of objects by moving their hands until they hit something.

For example, Schmandt describes an interface for entering multiple layers of VLSI circuit design data in a 3D stereoscopic work space [142]. The user enters the data by pressing a stylus on a stationary 2D tablet; the user can adjust the depth of the image so that the desired plane-of-depth lines up with the 2D tablet. Versions of the interface which constrained the 3D stylus position to lie on grid points via software mapping were less successful; the physical support of the tablet proved essential. Other useful 2D constraining surfaces include the physical surface of the user's desk, the glass surface of the user's monitor, or even a hand-held palette or clipboard.

4.8 Head tracking techniques

In a non-immersive spatial interface, desktop-based head tracking can allow the interface to "give back" some of the information lost by displaying 3D objects on a flat display, via head motion parallax depth cues. Previous research [118][46][178][111] discusses the advantages of head tracking and the implementation issues. An additional user study [131] shows performance improvement for a generic search task using an immersive head-tracked, head-mounted display vs. a non-head-tracked display.

4.9 Related versus independent input dimensions

The Jacob and Sibert study [91] compares user performance for two tasks: the first asks the user to match (x, y, size) parameters of two squares, while the second task requires matching (x, y, greyscale) parameters of two squares. Both tasks require the control of three input dimensions, but Jacob reports that user task performance time for the (x, y, size) task is best with a 3D position tracker, while performance for the (x, y, greyscale) task is best with a mouse (using an explicit mode to change just the greyscale).

Jacob argues that the 3D tracker works best for the (x, y, size) task since the user thinks of these as related quantities (“integral attributes”), whereas the mouse is best for the (x, y, greyscale) task because the user perceives (x, y) and (greyscale) as independent quantities (“separable attributes”). The underlying design principle, in Jacob’s terminology, is that “the structure of the perceptual space of an interaction task should mirror that of the control space of its input device” [91].

This result points away from the standard notion of logical input devices. It may not be enough for the designer to know that a logical task requires the control of three input parameters (u, v, w). The designer should also know if the intended users perceive u, v, and w as related or independent quantities. In general it may not be obvious or easy to determine exactly how the user perceives a given set of input dimensions.

4.10 Extraneous degrees of freedom

Most spatial input devices sense six dimensions of input data, but this does not mean that all six dimensions should be used at all times. If, for example, the user’s task consists only of orienting an object, it makes little sense to allow simultaneous translation, since this only makes the user’s task more difficult: the user must simultaneously orient the object and keep it from moving beyond their field of view. Extraneous input dimensions should be constrained to some meaningful value; designing an interface that is useful does not necessarily require realistically imitating the behavior of objects in the physical world.

In general, it makes good common sense to exploit task-specific needs to reduce dimensionality. For example, the mouse-based interactive shadows technique [77] allows constrained movement in 2D planes within a 3D scene. If the user's task consists only of such constrained 2D movements, this may result in a better interface than free-space 3D positioning. The same general strategy can be applied to the use of spatial input devices.

4.11 Coarse versus precise positioning tasks

In two dimensions, the direct manipulation paradigm allows rapid, imprecise object placement. But to perform useful work in the context of a complex application such as a document editor, direct manipulation often needs to be constrained by techniques such as gridding or snap-dragging [13][14]. Corresponding three-dimensional constraint techniques and feedback mechanisms need to be developed.

Users may have difficulty controlling an interface which requires simultaneous, precise control of an object's position and orientation. The biomechanical constraints of the hands and arms prevent translations from being independent of rotations, so rotation will be accompanied by inadvertent translation, and vice versa. Even in the real world, people typically break down many six degree-of-freedom tasks, such as docking, into two subtasks: translating to the location and then matching orientations [21].

The design hurdle is to provide an interface which effectively integrates rapid, imprecise, multiple degree-of-freedom object placement with slower, but more precise object placement, while providing feedback that makes it all comprehensible. As Stu Card has commented, a major challenge of the post-WIMP interface is to find and characterize appropriate mappings from high degree-of-freedom input devices to high degree-of-freedom input tasks.

Applications such as 3-Draw [141] and abstractions such as Gleicher’s snap-together math [64] make good initial progress toward providing constrained input in 3D, but the general “spatial input constraint problem,” and the issue of providing appropriate feedback in particular, is still a challenging area for future research.

4.12 Control metaphors

Ware [177] identifies three basic control metaphors for 3D interaction:

Eyeball-in-hand metaphor (camera metaphor): The view the user sees is controlled by direct (hand-guided) manipulation of a virtual camera. Brooks has found this metaphor to be useful when used in conjunction with an overview map of the scene [18][19].

Scene-in-hand metaphor: The user has an external view of an object, and manipulates the object directly via hand motion. Ware suggests this metaphor is good for manipulating closed objects, but not for moving through the interior of an object [177].

Flying vehicle control (flying metaphor): The user flies a vehicle to navigate through the scene. Ware found flying to be good for navigating through an interior, but poor for moving around a closed object [177]. Special cases of flying include the “car driving metaphor,” as well as the “locomotion metaphor,” where the user walks through the scene [18].

A fourth metaphor can be appended based on subsequent work:

Ray casting metaphor: The user indicates a target by casting a ray or cone into the 3D scene. The metaphor can be used for object selection [111] as well as navigation [115]. It is not yet clear under which specific circumstances ray casting may prove useful.

The selection of an appropriate control metaphor is very important: the user's ability to perform 3D tasks intuitively, or to perform certain 3D tasks at all, can depend heavily on the types of manipulation which the control metaphor affords. Brooks addresses this issue under the heading "metaphor matters" [19].

4.13 Issues in dynamic target acquisition

The term *dynamic target acquisition* refers to target selection tasks such as 3D point selection, object translation, object selection, and docking. There are several issues related to dynamic target acquisition tasks:

- Use of transparency to facilitate target acquisition
- Ray casting vs. direct positioning in 3D
- Cone casting vs. ray casting

The first two issues suggest general strategies, while the second two issues address 3D point selection and 3D object selection, respectively.

4.13.1 Use of transparency to facilitate target acquisition

Transparency is a good general technique to aid in dynamic target acquisition tasks for two reasons:

- *Occlusion cues*: Placing a semi-transparent surface in a 3D scene provides occlusion cues. The user can easily perceive which objects are in front of, behind, or intersected by a transparent surface.
- *Context*: Since the surface is semi-transparent, objects behind it are not completely obscured from view. This allows the user to maintain context as the transparent surface is manipulated.

Zhai [188] describes the use of a semi-transparent volume, known as the “silk cursor,” for dynamic target acquisition. Zhai’s experimental results suggest that for the 3D dynamic target acquisition task which he studies, transparency alone leads to greater performance improvements than stereopsis alone. Zhai’s work is the first I know of to generalize the benefits of transparent volumes for target acquisition tasks.

Other example uses of transparency to aid target acquisition include use of a 3D cone for object selection [111], use of a semi-transparent tool sheet in the Toolglass interface [15], or the use of the semi-transparent cutting plane in the props interface [80].

4.13.2 Ray casting versus direct positioning in 3D

Perhaps the most obvious way to implement point selection is to base it on the (x, y, z) position of the tracker, but in many circumstances 3D ray casting may be a superior strategy for selecting 3D points. Instead of directly specifying the 3D point, the spatial input device is used to shoot a ray into the scene, allowing the user to hold the input device in a comfortable position and rotate it to change the ray direction [111].

The 3D points selectable by casting a ray are constrained to lie on the surface of virtual objects in the scene. In many circumstances this is exactly what is desired. If it is necessary to select points on objects which are inside of or behind other objects in the scene, the ray casting can be augmented with a mechanism for cycling through the set of all ray-object intersection points. For disconnected 3D points, 3D snap-dragging techniques [14] can be used if the disconnected points are related to existing objects in the scene. If the disconnected points are on the interior of objects, ray casting can be combined with a “cutting plane” operator, which is used to expose the interior of the objects [80][111].

Digitizing points on the surface of a real object is an instance where ray casting may not be helpful. In this case, the real object provides a spatial reference for the user as well as physical support of the hand; as a result, direct 3D point selection works well [129].

4.13.3 Cone casting versus ray casting

For gross object selection, ray casting may become less appropriate, especially if the object may be distant. One could alternatively use a translucent 3D cone to indicate a region of interest; distance metrics can be used to choose the closest object within the cone. Note that “spotlighting” visual effects afforded by many graphics workstations can provide real-time feedback for this task. An implementation of this strategy is reported by Liang [111].

4.14 Clutching mechanisms

Most spatial interfaces incorporate some type of *clutching mechanism*, that is, a software mode which allows the spatial input device to be moved without affecting the 3D cursor. In my experience, some of the most confounding (for the user) and hard-to-fix (for the implementor) usability problems and ergonomic difficulties can arise due to poor clutch design. Section 3.8 of the previous chapter discusses some of the issues I have encountered in designing clutching mechanisms.

4.14.1 Recalibration mechanisms

At a low level, all spatial input devices provide the software with an absolute position in a global coordinate frame. The user interface should provide a *recalibration mechanism* for mapping this absolute position to a new logical position, which allows the user to specify a comfortable resting position in the real world as a center point for the interaction space. There are at least three basic recalibration strategies:

Command-based: The user explicitly triggers a recalibration command, sometimes referred to as a *centering command* or a *homing command*. JDCAD, for example, uses this strategy [111] to bring the 3D cursor to the center of the visible volume.

Ratcheting: Many spatial interfaces (e.g. [41], [176]) utilize the notion of *ratcheting*, which allows the user to perform movements in a series of grab-release cycles. The user presses a clutch button, moves the input device, releases the clutch button, returns his or her hand to a comfortable position, and repeats the process.

Continuous: In some cases recalibration can be made invisible to the user. For example, in a virtual reality system, when the user moves his body or head, the local coordinate system is automatically updated to keep their motions body-centric. Another example is provided by my props-based interface, where the nonpreferred hand is used to define a dynamic frame-of-reference relative to which other tools may be moved with the dominant hand, as discussed in section 3.6.2 (“The natural central object”) of the previous chapter.

These strategies can be composed. In a virtual reality application, for instance, the position of the hands will be continuously recalibrated to the current position of the head, but an object in the virtual environment might be moved about via ratcheting, or brought to the center of the user’s field of view by a homing command.

4.15 Importance of ergonomic details in spatial interfaces

Manipulating input devices in free space can easily fatigue the user. The designer of a spatial interface must take special pains to avoid or reduce fatigue wherever possible. A poor design risks degraded user performance, user dissatisfaction, and possibly even injury to the user. Some issues to consider include the following:

- Users should be able to move around and shift their body posture. The interface should not require the spatial input devices to be held within a fixed volume that cannot easily be adjusted by the user. Use of recalibration mechanisms is one way to address this problem.

- For desk top configurations, an adjustable height chair with arm rests can help to provide proper arm support. Also, using a **C** or **L** shaped desk can provide additional surface area to rest the arms.

- If the interface is designed well, fatigue should only be associated with prolonged, uninterrupted use. It may be useful to build time-outs into the system which remind the user to take an occasional break, and of course the system must be designed so that it is easy to pause the work flow, and resume it at a later time.

Based on my user observations, the posture of users' hands while manipulating spatial input devices is *not* the same as the hand posture required during typing. The palms face each other (instead of facing downward) and users typically either rest the sides of their palms on the desk top, or they support their forearms at the elbows using chair arm rests, and hold their hands in the air above the desk top. This suggests that the ergonomics requirements for spatial manipulation may be different than those for typing.

4.16 Discussion

This chapter extracts design issues from a large body of work. I have identified common themes in what has worked well for spatial input, and what has not. The issues presented in this chapter are not formally proven principles of design. They are meant to suggest ideas and serve as a guide to designers working with spatial input techniques, and should not be expected to serve as a substitute for testing an interface design with real users. Situations where these issues and strategies are useful, or where they are not, need to be better defined and characterized, and ultimately subjected to formal study.

“Knowledge and timber shouldn’t be much used till they are seasoned.”

Oliver Wendell Holmes, The Autocrat of the Breakfast

Chapter 5

Research Methodology

5.1 Purpose

After implementing the props-based interface, I felt encouraged by the enthusiastic informal evaluations offered by neurosurgeons and others. The system was also well received as a point design by the human-computer interaction community, offering additional encouragement, and my survey of techniques and issues across a broad selection of systems and experiments had suggested some areas worthy of further exploration. However, to make some salient general points about interface design, I felt that my research needed to move beyond the informal approaches used so far, and perform some formal evaluations under experimentally controlled conditions.

My goal was to move beyond point design and to introduce some careful scientific measurement of relevant behavioral principles. Even given that one wants to “evaluate” a system, there are many possible strategies for evaluation that one might choose. The purpose of this chapter is to outline some of the possibilities, to provide a rationale for the research methodology which was ultimately chosen, and to discuss the process for applying that methodology.

5.2 Evaluation with experts versus non-experts

My work has focused on three-dimensional interfaces for neurosurgeons, who are clearly a form of expert user. But what exactly is meant by an expert user? What are some of the issues raised by working with expert users, and evaluating the results of that work? There are at least three different types of experts, plus the category of non-experts, to consider:

Domain experts: These are experts, such as my neurosurgeon collaborators, who have thorough expertise and experience in a particular field or activity. Performing evaluations with domain experts is appropriate when the goal is to develop tools for the domain application and to demonstrate that those tools can improve current practice.

Interface experts: These are expert users who are proficient with a particular computer interface or set of tools. For example, Card [38] reports experiments with word processing experts. Card's goal is to develop "an applied cognitive science of the user" [38] which proposes some models of the human as a cognitive processor. In the context of Card's studies, interface experts are ideal candidates because they exhibit less variable behavior, which is consistent with cognitive skill rather than the searching behavior of novices performing problem solving.

Manipulation experts: These are individuals with great dexterity for skilled manual behaviors, such as painting, sculpting, or playing a violin. Studies of these people and the incredible things they can do with their hands might be helpful for answering questions of how one can build tools that help people to develop comparable skills.[1]

---

[1] Bill Verplank of Interval Research suggested the concept of "manipulation experts" during the ACM CHI'96 Workshop on Virtual Manipulation in April 1996.

Non-experts: These are users who may not share a common domain knowledge, may not have any experience with a task of interest, nor will they necessarily have a clear goal in mind with regards to an interface or technology being evaluated. Evaluation with non-experts is appropriate when an artificial goal or task can be introduced and the intent is to see if people can “walk up and use” an interface to accomplish that task. Non-experts are also appropriate for experimental testing of behavioral hypotheses about humans in general.

Neurosurgeons clearly are domain experts, and many neurosurgeons might also be considered manipulation experts because of the fine manual skill required during delicate surgery. Using neurosurgeons for evaluation imposes some constraints on what type of evaluation can be done. Neurosurgeons have heavily constrained schedules, and the available user community of neurosurgeons is quite limited.

5.3 Approaches for evaluation

Given the above constraints, during the planning stages for this research, I considered three general evaluation strategies which might be used[1]:

Informal usability testing: Demonstrate the interface to domain experts and solicit comments (verbally or through questionnaires) to assess how well the interface meets the task needs of the domain expert. This form of evaluation is essential to develop a useful tool for the domain expert, provides rapid feedback which is well suited to an iterative design process, and is helpful when forming initial hypotheses about factors which can influence the design. Informal usability testing cannot answer general questions as to why an interface might be better than alternative techniques, nor can it address specific experimental hypotheses.

---

[1] I would like to acknowledge personal communication with Rob Jacob which discussed these approaches.

Use by domain experts for real work: The ultimate proof of any tool is for a group of domain experts to use it to achieve goals in the process of their real work. If the domain experts say it is useful, then the tool is declared a success. This approach has been advocated by Fred Brooks [20]. For neurosurgery, the ideal test would be to deploy a tool in the clinical routine and to plan surgical interventions on real patients. This requires development and support of a commercial-quality tool which has been carefully tested for robustness and safety.

Formal Experimentation: Formal experimentation allows careful study of specific hypotheses with non-expert subjects under controlled conditions. Formal experimentation requires introduction of abstract tasks that non-experts can be trained to do quickly and which are suited to the experimental hypotheses, but which may or may not be directly analogous to actual tasks carried out by domain experts.

I decided that my primary goal for this dissertation was to make some general points about interface design and human behavior, so that some of the lessons I had learned in the neurosurgery application could be applied to other interface designs. The formal experimentation strategy best meets the requirements to achieve this goal: ample non-expert subjects are available for experimental testing of hypotheses about human behavior.

Even though formal experimentation is my primary approach, this work as a whole includes elements of all three strategies outlined above. I have performed extensive informal testing with domain experts to drive the interface design itself. Furthermore, although the interface is a research tool and not a clinical tool, it has been tested in the context of actual surgical procedures with real patients, in conjunction with our laboratory's surgical planning software [65][66][160], and Multimedia Medical Systems [122] is currently working to develop a commercial version of the interface for clinical use.

5.4 Principled experimental comparisons

The formal experimentation strategy can only make general points about interface design and human behavior when a principled approach is taken. A careless experimental design is subject to many pitfalls. A pitfall of particular concern when attempting to evaluate and compare user interfaces is known as the A *vs. B comparison* pitfall. In such evaluations, the purpose is typically to demonstrate that interface A is "superior to" interface B. But unilateral, unqualified statements of this form are almost always meaningless. Interface or input device comparisons should be made in the context of a specific task or set of tasks, and in the context of a specific class of intended users.

Buxton [31] presents the example of two drawing toys: an Etch-a-Sketch and a Skedoodle. The Etch-a-Sketch has two separate one-degree-of-freedom knobs to control the motion of the stylus, while the Skedoodle has a joystick which allows one to manipulate both stylus degrees-of-freedom simultaneously. The "research question" is this: Which toy has the better interface for drawing? For drawing one's name in cursive script, the Skedoodle excels. But for drawing rectangles, the Etch-a-Sketch is superior. The point is that neither toy is unilaterally "better for drawing," but rather that each style of interaction has its own strengths and weaknesses.

Another related fault of A vs. B comparisons is that they typically offer no insight as to why one interface differs from another. An unprincipled comparison of competing interfaces can easily confound independent experimental factors, making results difficult to interpret or generalize. For example, concluding that "touchscreens are easiest to use" from a comparison of a touchscreen and a mouse confounds the independent factors of absolute versus relative control, direct versus indirect input, device acquisition time, and the required accuracy of selection, among others. One must carefully formulate specific experimental hypotheses or predictions. Actually *testing* the hypotheses might still involve a carefully controlled comparison of alternative interfaces, but with the goal of testing specific hypotheses in mind, one can design principled evaluations which demonstrate the fundamental mechanisms or human capabilities at work, and thereby suggest new possibilities for design.

5.5 The process of experimental evaluation

There is a general pattern which should be followed in experimental design. It is vital to begin with principled hypotheses upon which to base an experimental comparison. Without some theory to guide the experimental design and the interpretation of experimental data, it is difficult to draw any firm conclusions.

During the initial stages of experimental design, pilot studies are conducted on a small number of subjects. Pilot studies for experiments are much like throw-away prototypes for software systems. The goal is to rapidly discover major surprises or flaws in the concept for the experiment before investing large amounts of time in a more formal study. If one is not getting the expected results, why not? Is the experiment fundamentally flawed, or are there minor problems with the experimental design? Pilot studies also allow one to work out the details of the experiment, such as the specific instructions to give to subjects or the amount of time to allow for each trial. The pilot studies drive modifications and improvements to the experimental design, in an iterative process which may go through several cycles.

The final formal study requires collecting the data for a meaningful sample size. The data from pilot studies can be used to calculate an effect size in terms of standard deviation units (the difference between means divided by the standard deviation). The effect size, in turn, can be used to estimate a sample size which will yield sufficient statistical power. Statistical power is the probability that an effect of a given magnitude can be found with a sample of fixed size, assuming that the effect really exists [93]. If the sample size is too small, statistical power is low, and the probability of detecting an effect, even if that effect actually exists, will go down. Thus, the pilot studies also serve to ensure that the final study will not fail to find an effect because of lack of statistical power.

5.6 Data analysis

The formal study provides the data, but data analysis is still necessary to demonstrate the experimental hypotheses. Data analysis requires a careful and thorough exploration of the data. The data analyses in this dissertation use standard *analysis of variance* (ANOVA) techniques, which perform a linear least squares fit of a model (derived from the experimental hypotheses) to the data [93].

Linear analysis of variance makes several assumptions about the data, each of which must be checked [12][93]. For example, the errors (differences between the predicted model and the observed data) should be normally distributed, and the errors should not be correlated with the predicted values. Also, the data must be checked for outliers (such as trials where a computer glitch may have occurred, or the subject might have sneezed) which could unduly bias the analysis. These details of analysis are not addressed in the context of the individual experiments presented here, but the statistical confidences reported are based on a thorough analysis.

All of the experiments described in this thesis use *within-subjects* designs counterbalanced for order of presentation. This means that each subject performs all experimental conditions, but that the order in which subjects perform the conditions is systematically varied. This is sometimes referred to as a *latin squares* design. A latin squares design helps to ensure that the results will not be biased by order of presentation effects, because order is an explicit between-subjects factor that can be analyzed separately.

For example, imagine that we are designing an experiment that compares condition A with condition B. In a within-subjects design, each subject will perform both condition A and condition B, yielding two groups of subjects: the A-condition-first subjects and the

B-condition-first subjects. These groups will be balanced so that half of the subjects try condition B before condition A, while the other half try condition A before condition B. This can help to ensure that any detected difference between condition A and condition B is not entirely due to the order in which subjects performed the conditions: the effects of order are controlled and can be explicitly analyzed.

Order	First Condition	Second Condition
A-condition-first subjects	A	B
B-condition-first subjects	B	A

Figure 5.1 Example Latin Square for two experimental conditions.

Within-subjects designs typically use a *repeated measures* technique for analysis of variance. Each subject performs multiple conditions, and therefore contributes multiple data observations. But these are not truly separate observations because the observations from a single subject will tend to be correlated: for example, if a subject performs faster than average during one condition, it is more likely that the subject will perform faster for other conditions as well. Repeated measures analysis of variance takes this repetition of observations into account when computing statistical confidence levels.

The statistical results given in my formal experiments make frequent mention of the *F statistic* and p, the *statistical confidence*. Both of these statistics are related to testing a model to see how well it predicts a given set of observations. The observed data values Y are the sum of the value Y_p predicted by the model plus unexplained error E:

$$Y = Y_p + E$$

Adding a variable to a model either helps to explain some additional error that was unexplained before, or it has no predictive power, contributing nothing to the model. For statistical analysis, the relevant question is whether or not adding a parameter to a model accounts for more error than one would expect by random chance alone. The F statistic is a ratio which quantifies this. F is the sum of squared error accounted for by adding a new variable to a model divided by the sum of squared error one would expect to account for by adding a random variable to a model. The F statistic can be used to look up p, the statistical confidence for the model. The statistical confidence p is the probability that random selection of N observations from a population resulted in a F statistic of a given magnitude, given the number of parameters (or degrees-of-freedom) of the model. If p is low enough (usually $p < .05$ is considered sufficient), then the statistical analysis suggests that the F statistic did not result from a deviation caused only by random selection, and therefore one concludes that adding the new variable to the model significantly improves prediction of the observed data, with a confidence level of p.

5.7 Conclusion

This chapter has provided a discussion of the issues of working with expert domain users and has demonstrated the rationale and process for the experimental work described in subsequent chapters. Each evaluation technique has its own advantages and disadvantages. For the purpose of extracting general design principles and knowledge about human behavior, evaluation with non-experts is a suitable strategy. This thesis offers some elements from each of the informal usability testing, use by domain experts for real work, and formal experimentation strategies, providing a synergy of real-world application and careful scientific measurement. The remaining chapters will now focus on the formal experiments.

Chapter 6

Usability Analysis of 3D Rotation Techniques

6.1 Introduction

One of the main points I would like to make in this dissertation is that passive haptic issues are important: facile virtual manipulation requires studying the *feel* of the interface, and not just the *look* of the interface. Although I believe my main contributions are related to the two-handed issue, I certainly had to consider other issues as I designed the interface for the neurosurgery application. One of the first major design decisions I faced was whether or not an approach based on 3D input was appropriate for the neurosurgery application at all. Visualizing and manipulating volumetric data sets with dexterity is important for this application, but the literature often seems to suggest that 3D input may have some shortcomings in this regard. My design experience suggested that, with appropriate design, this is not necessarily true: 3D input devices can be useful for both dextrous and fast virtual manipulation. The main contribution of the experiment presented in this chapter is to formalize this suggestion for the task of orienting virtual objects.

A more general motivation may also be appropriate here. With the rapid migration of fast, cheap 3D graphics to the PC platform [170], and with 3D graphics beginning to make their presence felt on the web through standards such as VRML [23], applications which incorporate 3D manipulation and 3D object viewing will become increasingly prevalent. In particular, orienting a virtual object to a desired view is a fundamental task, since viewing or inspecting an object is often a precursor to further manipulation.

Since a 1988 study by Chen which introduced the Virtual Sphere [40], I am not aware of any quantitative data comparing performance of currently proposed techniques for specifying 3D orientation. For example, there is no formal user study to compare the Virtual Sphere and the Arcball proposed by Shoemake [153][154], nor is it known what advantages (if any) direct orientation input using a 3D input device might offer for 3D orientation tasks. Chen's study does not include detailed observations of user expectations or the common difficulties encountered. Given that computer users typically have no experience with interactive computer-based 3D rotation tasks, a description of the usability problems that novice users may encounter when first exposed to these techniques forms a useful contribution.

The high cost of 3D input devices has traditionally limited their use to research or niche market applications such as head-mounted virtual reality systems, high-end animation software, or medical visualization. Free-space 3D input devices are still expensive compared to the mouse, but with the recent introduction of PC-based devices [137][5] priced near $1000, these devices are more affordable and practical now than they have ever been, and a growing number of interface designers will have the opportunity to explore the possibilities of free-space 3D input.

The intent of this study is not to argue that any one device or technique is “best.” Each interface device or input technique will excel for some tasks and languish for others [31]; the most appropriate device for an application depends on the context of tasks to be supported and the intended users. The goal here is to collect some solid performance data for an experimental rotation matching task, so that informed design decisions can be made, and to collect some qualitative observations that will help to illustrate some strengths and weaknesses of each technique.

6.2 Overview

This formal user study evaluates interactive 3D rotation techniques including the mouse-driven Virtual Sphere and Arcball techniques (which I will shortly explain in more detail), as well as direct 3D orientation input techniques based on magnetic orientation sensors. The data suggest that when performing an orientation matching task, users can take advantage of the integrated degrees of freedom of 3D orientation input to complete the task up to 36% faster, without necessarily sacrificing any statistically detectable degree of accuracy.

I report detailed observations of user expectations and common usability problems when first encountering the techniques. My qualitative observations also suggest some design issues for 3D input devices. For example, the physical form-factors of the 3D input device had a marked effect on the subjective user acceptance (but not on quantitative task performance) for otherwise identical input sensors. The device should afford some tactile cues, so the user can feel its orientation without looking at it. In the absence of such cues, some test users were not able to form a clear concept of how to use the device.

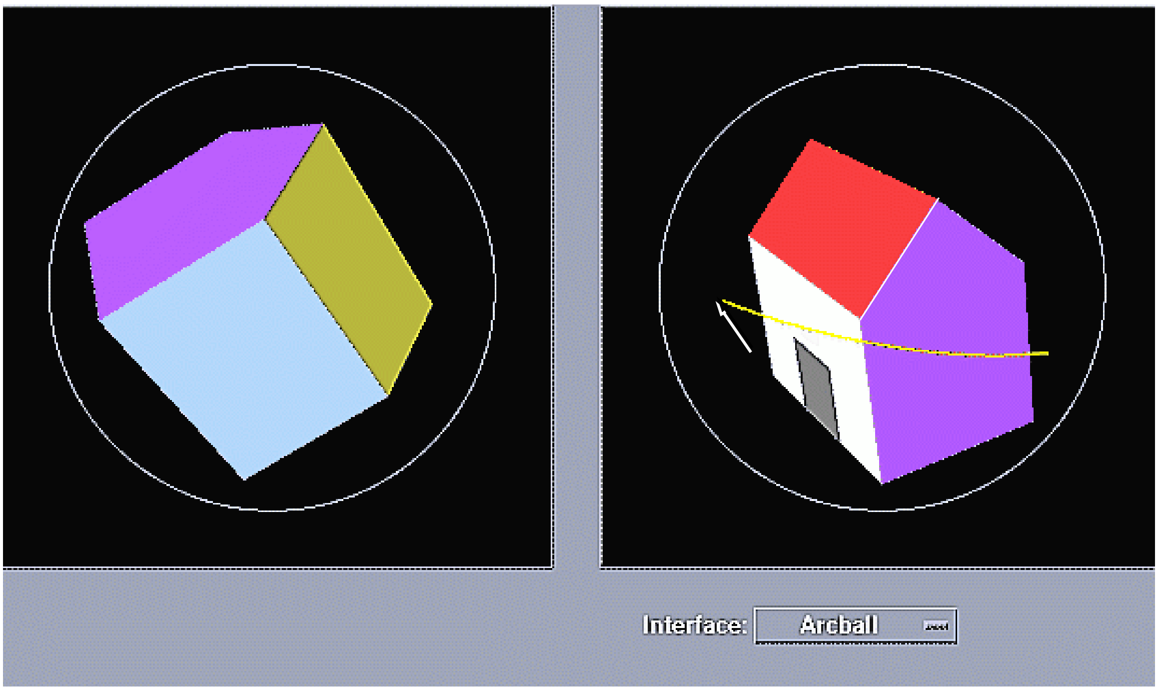
6.3 Interaction techniques

My analysis includes the following interactive 3D rotation techniques:

Virtual Sphere: This mouse-driven 2D interface simulates a physical trackball. The virtual object is shown on the screen, and when the user clicks and drags on the virtual object, the computer interprets these drags as tugging on the simulated trackball. The virtual object rotates correspondingly. To provide the third rotational degree of freedom, a circle is drawn around the object (*figure 6.1*), and when the user clicks and drags in the area *outside* of the circle, rotation is constrained to be about the axis perpendicular to the computer screen. Hereafter, this outside area of the circle will be referred to simply as "the outside."

Arcball: This interface is similar to the Virtual Sphere, but it is based upon a more mathematically rigorous quaternion [152][155] implementation. It does not suffer from problems with gimbal lock or noisy data, and its implementation affords easy addition of constraint modes. Some designers consider the Arcball to be the best known 2D technique for 3D rotation. Shoemake has performed an informal comparison [153], but no quantitative data currently exist which compare the Arcball and the Virtual Sphere.

The Virtual Sphere and Arcball both require the user to achieve some orientations by composing multiple rotations, since only two of the three possible rotational degrees of freedom can be accessed at any one time.



Interface: Arcball

Figure 6.1 Screen snapshot of the experiment software.

3D Ball: The user rotates a two-inch diameter plastic sphere (*fig. 6.2*) instrumented with a magnetic tracker to manipulate the virtual object. The magnetic tracker simultaneously provides all three rotational degrees of freedom, so in principle the user never has to mentally compose rotations with this interface. However, it is not clear if users can employ coupled rotation axes effectively [40], nor is it clear if the multiple degrees of freedom result in faster, but possibly less accurate, input of orientation data.

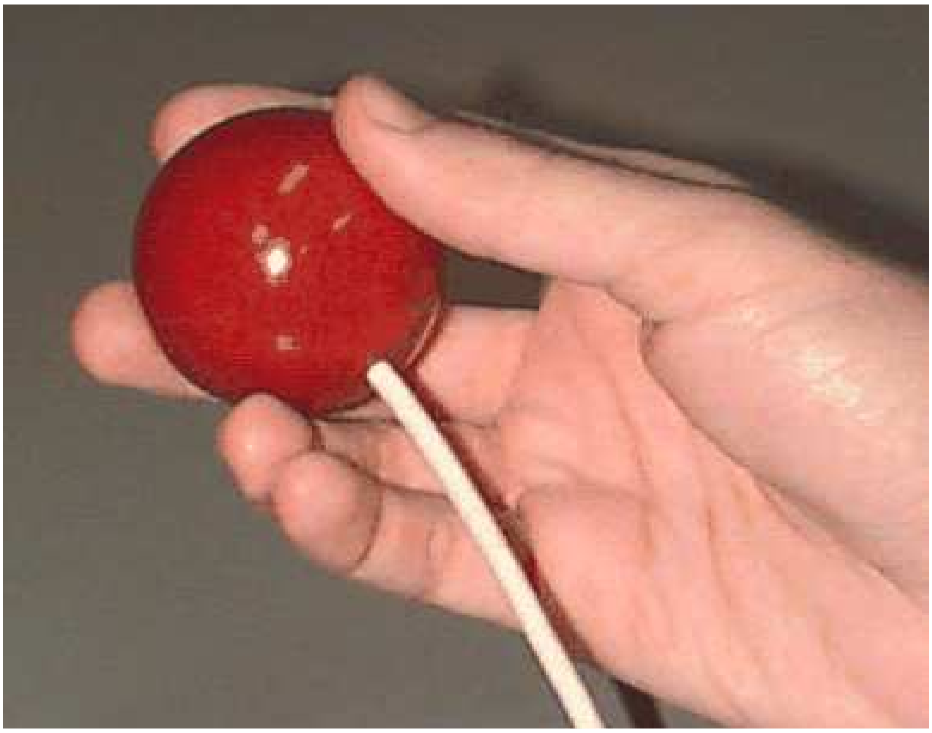


Figure 6.2 The 3D Ball input device.

The 3D Ball always acted as an absolute rotation controller: the orientation of the object being manipulated always matched the orientation of the 3D Ball. With the addition of a clutching mechanism (for engaging and disengaging the ball from a virtual object), it would be possible to use the 3D Ball as a relative rotation controller, by performing “ratcheting” movements [179]. Some other ball-shaped 3D input devices with integrated clutch buttons have used this technique [41][137][164]. The 3D Ball used here did not include any integrated buttons; I will return to the issue of integrated control buttons in section 6.9 of this chapter.

For the purposes of this experiment, the design of the device was kept as simple as possible, to avoid introducing secondary variables (such as absolute versus relative rotation). The basic task of rotating the object with integrated degrees-of-freedom is the main issue for this study.

Tracker: The Tracker interface, which also uses a magnetic orientation sensor, is identical to the 3D Ball in all regards except the physical packaging (*fig. 6.3*). This is the default form for the input device as shipped by the manufacturer [137], and as such represents the only way to use the device without designing or purchasing an alternative housing [47].

Figure 6.3 The Tracker 3D input device.

The Tracker has an unusual and unfamiliar shape. In our virtual reality lab [131], we have noted that with practice, experts can become quite proficient with the Tracker despite its awkward shape. It is not clear how well novice users will be able to adapt to its design.

6.4 Hypotheses

This study investigates the following specific hypotheses:

H1: Users can effectively use coupled rotation axes, and integrated control of all three degrees-of-freedom for rotation will provide significantly faster input of orientation data.

A study by Jacob [91] suggests that multiple degree-of-freedom input will be most appropriate when users think of a task's control parameters as integral attributes; I propose that 3D rotation matching is one such task. Most people are not good at mentally composing rotations, so when attempting to perform complex rotations, the separated 2D+1D control required by the Arcball and Virtual Sphere techniques should reflect this.

H2: Three-dimensional input is often assumed to be fast but inaccurate. I hypothesize that, at least for a 3D orientation matching task, 3D input provides fast orientation input without necessarily sacrificing any accuracy.

H3: The physical shape (or affordances) of the 3D input device can be an important design consideration in itself.

This hypothesis arose from my previous work on the props interface. The props take advantage of natural affordances (as discussed by Norman [126]), which can help users to know what to do just by inspecting or grasping an object or input device. The 3D Ball and Tracker used in this experiment are more general-purpose 3D input devices, yet nonetheless each communicates natural affordances which will implicitly channel user behavior; I intend to explore these issues in the analysis.

H4: The Arcball includes several apparent improvements over the Virtual sphere. As such, the Arcball should outperform the Virtual sphere in terms of task performance, user acceptance, or both.

6.5 The Experiment

6.5.1 Task

Test users performed an orientation matching task based on the task employed by Chen [40]. The goal here is not to reproduce Chen's results, but rather to extend the set of interactive 3D rotation techniques that have been formally evaluated with a common task.

A static view of a solid-rendered 3D model of a house, at a randomly generated orientation [4], was shown on the left side of the screen (*figure 6.1*). Test users attempted to manipulate a second view of the house on the right-hand side of the screen to match the random orientation. Each side of the house was colored uniquely to facilitate the matching. A circle was always drawn around both images of the house to assist matching, even though the circle only was strictly necessary for the Arcball and Virtual Sphere techniques.

When test users felt that the orientations matched, they clicked a footpedal to end the trial. After each trial, performance was rated as “Excellent!” (shortest-arc rotation less than 5.7 degrees), “Good Match!” (less than 7.6 degrees), or “Not good enough, try harder next time.” I chose a footpedal to end the trials, rather than the spacebar used by Chen [40]. This kept the desk surface open for manipulation and it also allowed test users to use both hands (if desired) to manipulate the 3D input devices. The keyboard was removed from the desk during all experimental conditions.

Participants were given as little instruction as possible with each controller. With the Arcball and Virtual Sphere, the experiment began with a practice exercise to encourage the test user to click and drag within the circle (for two degree-of-freedom rotation) as well as outside the circle (for the third degree of freedom). When presenting the 3D Ball or Tracker devices, I placed the device on the table in front of the test user, mentioned that “It is not a mouse, it works differently,” and suggested “Try to discover how to use it.”

6.5.2 Experimental design

A within-subjects latin square design was used to control for order of presentation effects. Test users tried all four interfaces in a single session lasting about 1 hour. Test users performed matches for 15 unique orientations with each interface, but only the last 10 of these were included in the data analysis, to avoid any transient initial learning effects. There was a short break between conditions, during which I interviewed the test user about his or her impressions of the interface technique.

Dependent variables were Time to completion and Accuracy of the match. Accuracy was measured by the shortest-arc rotation between the final user-specified rotation and the ideal matching orientation.

6.5.3 Test users

Twenty-four unpaid test users (12 male, 12 female, all right-handed, mean age 19.1 years) were recruited from the University of Virginia psychology department’s subject pool. All test users had experience with the mouse, while none had any experience with 3D input devices. Two test users had previously tried an interface similar to the Virtual Sphere, in the context of a molecule visualization application.

6.6 Results

Figure 6.4 shows the mean completion times and accuracies which test users achieved. The 3D Ball was 36% faster than the 2D techniques and the Tracker was 33% faster. There was little variation in the mean accuracies obtained.

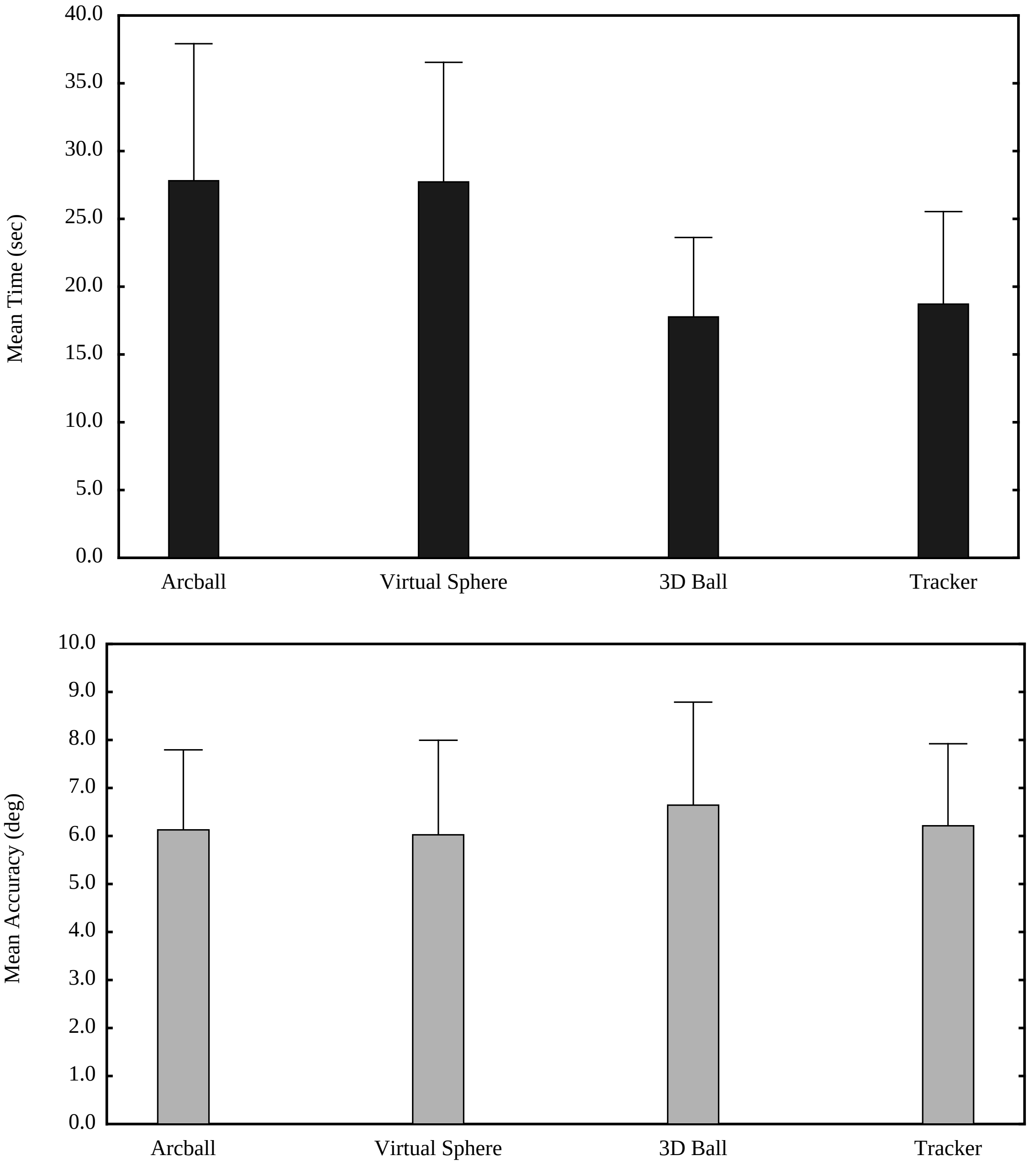


Figure 6.4 Mean times (top) and accuracies (bottom)

Comparing to Chen's results for the Virtual Sphere, test users in the present study had longer times (Chen reported a mean of 17.5 seconds for complex rotations [40], while this study found 27.7 seconds), but in this study test users were more accurate (Chen reported 8 degrees of error[1], while here test users achieved 6 degrees of error).

These discrepancies are probably primarily due to the differing test user populations: Chen used all males, some graduate students, and some students with experience in 3D graphics systems. This study includes females, all test users were undergraduates, and only two test users had any experience with interactive 3D rotation tasks. Another, probably less important, difference between the studies is that Chen's test users also performed more trials (27 trials vs. 15 trials in this study) so they may have become more proficient with the experimental task.

6.6.1 Statistical analysis

I performed an analysis of variance with repeated measures on the within-subjects factor of interface used, with task completion time and accuracy as dependent measures. The interface used was a highly significant factor for completion time ($F_{(3,69)} = 37.89$, $p < .0001$) but not for accuracy ($F_{(3,69)} = 0.92$, $p > .4$, n.s.).

Comparisons for completion time (*fig. 6.5*) revealed that the 3D interfaces were significantly faster than the 2D interfaces, but there was no significant difference between the 3D Ball versus the Tracker, nor was there any significant difference between the Arcball versus the Virtual Sphere.

These results strongly supports hypothesis H1, suggesting that users can perform an orientation matching task significantly faster when the rotation is presented as three integrated degrees of freedom.

---

1. Chen reported accuracy in terms of the sum of squared errors between the user's rotation matrix and the rotation matrix to match. Chen indicates a squared error of 0.04 [40], which converts to 8.1 degrees.

Comparison	**F statistic**	**Significance**
3D Ball vs. Arcball	$F_{(1,23)} = 58.96$	$p < .0001$
3D Ball vs. Virt. Sphere	$F_{(1,23)} = 56.24$	$p < .0001$
3D Ball vs. Tracker	$F_{(1,23)} = 0.83$	$p > .35$, n.s.
Tracker vs. Arcball	$F_{(1,23)} = 47.31$	$p < .0001$
Tracker vs. Virt. Sphere	$F_{(1,23)} = 50.80$	$p < .0001$
Arcball vs. Virt. Sphere	$F_{(1,23)} < 0.01$	$p > .95$, n.s.

Figure 6.5 Interface comparisons for completion time.

Comparisons for accuracy confirmed that there were no significant differences between any of the interfaces. This supports H2, suggesting that any accuracy differences between the interfaces are nonexistent or too small to detect with N=24 test users.

The analysis also revealed that the between-subjects factor of Sex was significant for completion time, as were the Interface × Sex and Interface × Order interactions for both completion time and accuracy (*table 6.6*). This indicates a need to investigate possible biases in the data due to the Sex or Order factors.

Factors for Time	**F statistic**	**Significance**
Sex	$F_{(1,19)} = 9.69$	$p < .005$
Interface × Sex	$F_{(3,57)} = 3.35$	$p < .03$
Interface × Order	$F_{(9,57)} = 2.85$	$p < .01$
Factors for Accuracy	**F statistic**	**Significance**
Interface × Order	$F_{(3,57)} = 4.79$	$p < .02$
Interface × Order	$F_{(9,57)} = 2.01$	$p < .06$

Figure 6.6 Significant between-subjects factors and interactions.

6.6.2 Separate analysis for males and females

This study was not designed to analyze sex differences, yet the results suggest that Sex was a significant factor. This is consistent with the sex differences literature, which has found an advantage for males on some tasks which involve mental rotation [74]. To ensure that Sex was not a distorting factor in the final study, I performed separate analyses with the N=12 male and N=12 female test users (*fig. 6.7*). For completion time, the results of the separate analysis were similar to those obtained in the combined analysis, suggesting that Sex was not a distorting factor.

Males	**Time (sec)**	**Accuracy (deg)**
Arcball	22.1	6.3
Virtual Sphere	23.1	6.4
Ball	14.9	6.3
Tracker	15.9	6.2
Females	**Time (sec)**	**Accuracy (deg)**
Arcball	33.5	5.9
Virtual Sphere	32.4	5.7
3D Ball	20.7	7.0
Tracker	21.5	6.2

Figure 6.7 Means for each interface technique by sex.

For accuracy, there was a relatively small, but significant, effect for females only (*fig. 6.8*). Females were about one degree more accurate (*fig. 6.7*) when using the mouse-based techniques vs. the 3D Ball, but not vs. the Tracker. This suggests a minor qualification to H2, that 3D input (at least with the 3D Ball) may be slightly less accurate than 2D input for females but not for males.

Females	**F statistic**	**Significance**
3D Ball vs. Arcball	$F_{(1,11)} = 4.91$	$p < .05$
3D Ball vs. Virt. Sphere	$F_{(1,11)} = 5.02$	$p < .05$

Figure 6.8 Significant accuracy device comparisons for females only.

6.6.3 Between-subjects analysis

I also performed a between-subjects analysis using only the data from the *first* interface that each test user tried. Thus, the 24 test users were divided into 4 groups with 6 users each, resulting in an analysis which eliminates the Order × Interface effects (*fig. 6.6*) as possible factors, but which also results in an overall less sensitive analysis.

Interface	Time (sec)	Accuracy (deg)
Arcball	30.8	5.1
Virtual Sphere	31.9	7.3
3D Ball	19.9	8.6
Tracker	21.3	7.3

Figure 6.9 Means obtained for the first interface tried.

The accuracy data suggested that test users were less accurate on the first interface tried for all interfaces, except for the Arcball (*figs. 6.9, 6.10*). I can offer no explanation for this result, but the within-subjects analysis strongly suggests that any such initial accuracy differences between conditions evened out as the test user became more experienced with the task. Nonetheless, even though there is only data for N=6 test users in each group, the between-subjects analysis suggests the following qualification to H2: The 3D Ball may be less accurate than the Arcball when users are first exposed to the orientation matching task.

Comparison	F statistic	Significance
3D Ball vs. Arcball	$F_{(1,23)} = 12.60$	$p < .002$
Tracker vs. Arcball	$F_{(1,23)} = 3.75$	$p < .07$
Arcball vs. Virt. Sphere	$F_{(1,23)} = 5.20$	$p < .05$

Figure 6.10 Significant effects for between-subjects analysis of accuracy.

6.7 Qualitative results

Figure 6.11 shows a histogram of the subjective ranks which test users assigned to each interface. The Arcball ranks suggested a generally positive reaction, but this trend was not significantly better than the ranks assigned to the other techniques. The 3D Ball was most commonly selected as the preferred technique for the experimental task, and nobody rated it as the worst technique. Figure 6.12 shows a pie chart illustrating how many subjects chose each interface as their “most preferred technique.”

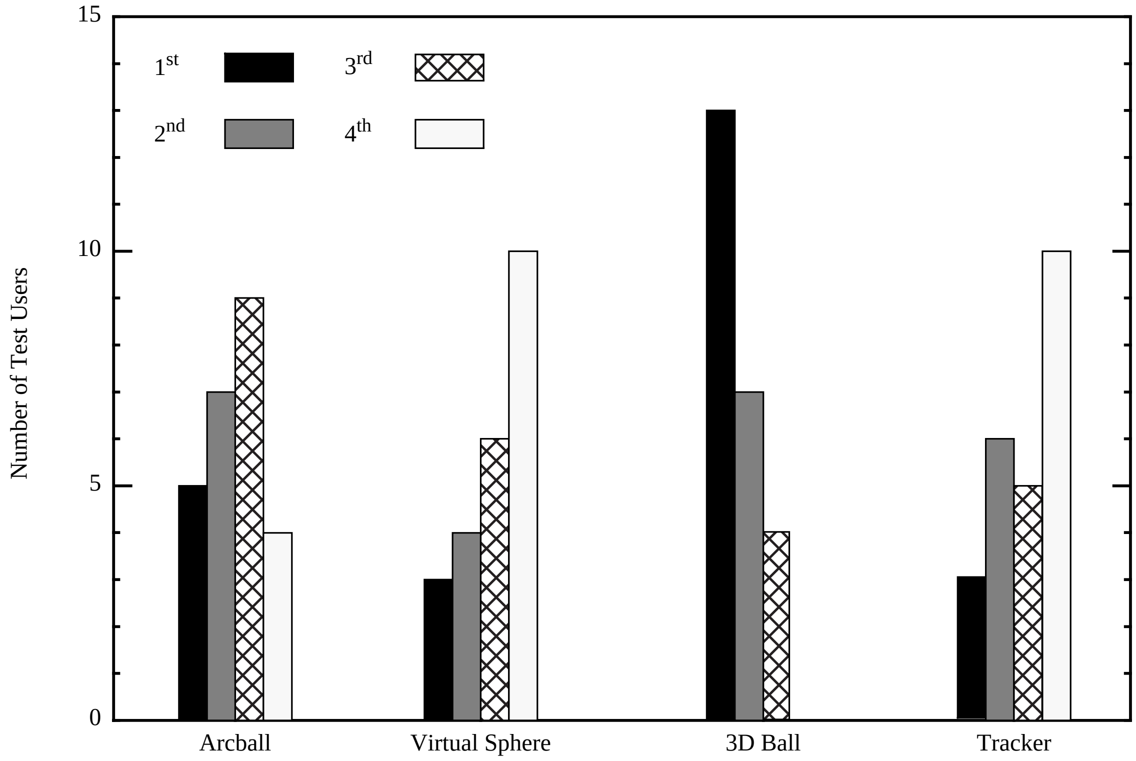


Figure 6.11 Histogram of subjective ranks for each interface.

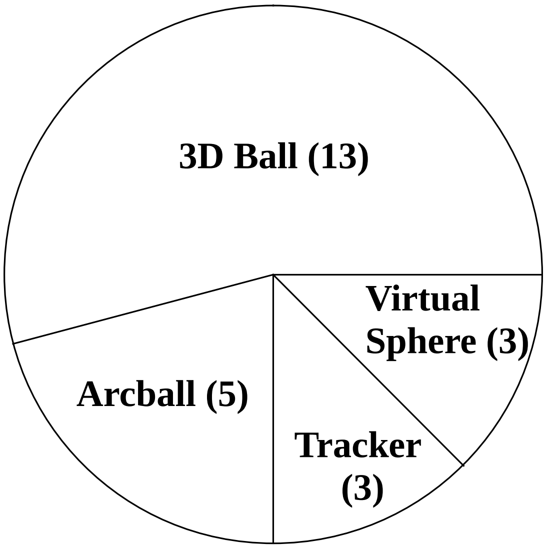


Figure 6.12 Pie chart showing distribution of votes for the favorite technique.

Statistical analysis confirmed that the 3D Ball had significantly higher ratings than the mouse-based 2D techniques or the Tracker 3D input technique (*fig. 6.13*).This provides strong evidence in favor of H3, that the physical form factors of a 3D input device can be an important design consideration. The awkwardness of the Tracker resulted in poor subjective impressions, despite the relatively high task performance which test users were able to achieve with it. Yet the exact same input sensor packaged as a 3D Ball resulted in significantly more favorable reactions.

Subjective Ranks	F statistic	Significance
3D Ball vs. Arcball	$F_{(1,23)} = 7.46$	$p < .015$
3D Ball vs. Virt. Sphere	$F_{(1,23)} = 17.12$	$p < .0005$
3D Ball vs. Tracker	$F_{(1,23)} = 27.94$	$p < .0001$
Tracker vs. Arcball	$F_{(1,23)} = 1.20$	$p > 0.25$, n.s.
Tracker vs. Virt. Sphere	$F_{(1,23)} = 0.01$	$p > 0.90$, n.s.
Arcball vs. Virt. Sphere	$F_{(1,23)} = 2.12$	$p < 0.15$, n.s.

Figure 6.13 Statistical comparison of subjective ranks.

6.7.1 2D techniques: the Arcball and Virtual Sphere

The Arcball & Virtual Sphere shared many qualities, so I will discuss their similarities before contrasting the techniques. The techniques were generally well accepted, with many users commenting that the techniques were “pretty easy” or that they “worked really well once you learned the inside and outside of the circle.”

The most common problems related to the modal distinction between the “inside” and the “outside” of the circle. During practice, I was careful to have users try both dragging inside and outside of the circle. But users see the circle as a target, and not as a border distinguishing two separate modes, so users would frequently attempt to click on the circle itself, and thus mistakenly switch between the inside and outside behaviors. Similarly, when attempting a large single-axis rotation using the outside of the circle, users would mistakenly come inside the circle, and would be surprised when the initial “outside” behavior changed to the “inside” behavior.

Several test users avoided using the outside, and would sometimes become unsure of what to do next if all the rotations matched except the rotation about the third axis perpendicular to the screen. A previous pilot study[1] had revealed that even though this third rotation axis was available, it was a hidden feature of the interface which very few users would discover on their own. In the final study reported here, during initial practice I had each user try dragging the mouse in the outside area of the circle, yet some test users still chose to ignore this feature.

Many test users were uncertain about *where* to click and drag with the mouse, and once they started to drag, they were reluctant to stop. This resulted in the impression that using the mouse was “not as smooth as [the 3D techniques]- you have to start and stop a lot to click the mouse button.” Some test users hesitated to make a large movement which

---

[1] I would like to acknowledge the contribution of Joe Tullio, who implemented the interfaces, ran all of the pilot studies, and assisted with the final study in the course of his senior thesis.

would disturb the progress made so far. As one user commented, “When you get close, you can screw it up -- I didn’t want to mess up what I already had.” Thus, test users sometimes seemed to be unsure of what effect their actions would have, and as a result they would try to plan their motions carefully. This is not a behavior which I observed with the 3D techniques.

6.7.2 Comparison of Arcball and Virtual Sphere

In theory, there are two primary distinctions between the behavior of Arcball and that of the Virtual Sphere: the Arcball avoids “hysteresis effects” and the Arcball uses half-length arcs. Hysteresis occurs when “closed loops of mouse motion may not produce closed loops of rotation” [153]. This means that it may not be possible to “undo” a sequence of drags by reversing their order. Half-length arcs are a property of how rotations combine, and result in a fixed C:D (Control:Display) ratio which is free of hysteresis. For example, with half-length arcs, a sweep across Arcball’s inner circle moves the virtual object 360 degrees. The same sweep would only move the Virtual Sphere 180 degrees. This rotational C:D ratio is fixed by the mathematics underlying Arcball and cannot be changed without reintroducing hysteresis.

Some users did not notice these differences between the Arcball and Virtual Sphere techniques; as one test user put it, “I felt like I was doing the same thing again.” Nonetheless, there was a general preference (16/24 participants) for the Arcball over the Virtual Sphere. A typical reaction was: “The Arcball is a little more responsive, and it gives you more control.” Test users who preferred the Virtual Sphere often commented that they liked its slower rate of movement. The Virtual Sphere’s slow movement suggested that the mouse was “pushing on a specific spot,” whereas “Arcball just rotated around in different axes.”

The Arcball also displays feedback arcs to illustrate how mouse movement affects rotation. These feedback arcs did not seem to be a significant advantage. The feedback arcs were often ignored or even regarded as annoying, although at least a few test users thought they helped at first. The exception was feedback arcs for the outside of the circle (for rotating about the axis perpendicular to the screen), which test users did find to be helpful.

6.7.3 3D Ball

Overall, test users had very positive reactions to the 3D ball technique. Typical comments were “this makes you have total control,” “it was like holding the object,” and “you could just turn it around instead of repeatedly clicking [with the mouse].”

Unlike the mouse-based techniques, the 3D manipulation techniques had to consider the physical form of the input device, as none of the test users had previous experience with 3D input. The 3D Ball’s form-factors help to convey a clear message: balls are for rolling, spinning, or turning. However, this didn’t always assist learning of the device. Several test users were initially convinced that the 3D Ball should be used by rolling it on the desk surface, rather than by picking it up. This “rolling” strategy was especially problematic for orientations which required the cord to point directly downward.

Although the ball shape conveys a clear message, its smooth, spherical aspect was sometimes problematic, because the surface offered few tactile landmarks or handles. This seemed to prevent some test users from forming a clear concept of the device. One test user commented that “I don’t think I ever figured out how to use it-- I just wiggled it around in my fingers until I found it [the matching rotation].” The smooth surface was also somewhat slippery, making it more difficult for some test users with small hands to tumble the ball.

Most test users found the 3D Ball’s cord annoying, as it would sometimes get in the way or become tangled. Some users also felt that they weren’t as precise with the 3D Ball as with the mouse techniques, and some preferred using the mouse because they were used to it.

6.8 Tracker

The Tracker's physical form-factors do not convey any clear message. Test users were typically quite confused when first encountering the Tracker ("this is a very weird thing"), but they were able to adapt after several trials. One test user explained that "at first I thought it had this really dumb shape, but it turned out to be easier to hold than I thought it would be."

Several test users initially tried to use the tracker by sliding it on the desk, and users often experimented with alternative grips for holding the device in the air. Users commented that they were "unsure how to use it," that it was "an odd shape for manipulation," or that the device seemed to be too small to grasp effectively. Many test users used both hands to hold the tracker, but this did not seem to be by choice. As one test user indicated, "you almost *had* to use two hands" to manipulate the tracker effectively. With the 3D Ball, most test users employed one hand only.

The hindering effect of the cord was the most common verbal complaint. The weight of the cord is comparable to the weight of the tracker itself, which makes it difficult to rotate the device about its center of mass, and results in some awkward postures.

The irregular shape of the tracker caused much confusion, but it also conferred some advantages. The mounting flanges for the tracker (*figure 6.3*) served as handles, and gave some tactile information as to the orientation of the device. One user commented that the "ball was a big smooth object, but now I have some handles and landmarks." Three test users who struggled with the 3D Ball performed quite well with the Tracker: the tactile cues afforded by the device seemed to be essential for these individuals.[1]

---

[1] These are the three test users who rated the Tracker as their favorite technique (*fig. 6.12*).

6.9 Discussion

The qualitative observations from this formal study, as well as prior informal observations from implementing several variants of 3D input devices for orientation [80][164], suggest some general design parameters which can influence how users will employ these devices. These are not well-formulated principles for design, but rather some issues intended to demonstrate how a design can implicitly channel user behavior to differing styles of interaction.

Affordances: The input device needs to suggest that it serves to orient virtual objects, while at the same time suggesting that it should be picked up, and not used on the desk surface. In my work with the neurosurgical interface props, the doll’s head serves this purpose well: the doll’s head just can’t be manipulated effectively on the desk surface, and so strongly encourages being picked up. The 3D Ball presented in this experiment did not suggest this well enough, leading some test users to initially use it by rolling it on the desk surface.

Tactile Cues: For absolute rotation control, the device should not necessarily be completely symmetric and should have clear tactile information which indicates preferred orientation and helps to hold the device securely. In this regard, a ball-shaped device has some shortcomings. For relative rotation control, however, this issue needs to be explored further. The symmetric ball shape might be advantageous in this case, since any landmarks on the device would not correspond to the virtual object, and could be misleading. This is an example of how a drawback for one style of input may be a virtue for another.

Grasp: Many 3D input device designs encourage the user to hold the device against the palm at a fixed orientation. This is known as the *power grasp* [113] because this hand posture emphasizes strength and security of the grip. By contrast, the *precision grasp* involves the pads and tips of the fingers and so emphasizes dexterity and free tumbling of the input device. This issue was formally analyzed by Zhai for a 6 DOF docking task, where he found that using the fine muscle groups emphasized in the precision grasp results in significantly faster performance [187].

The design of a 3D input device can influence which style of grasp users will choose. Integrating buttons with the input device encourages users to adopt a power grasp, holding the device in the palm, while the fingertips maintain contact with the buttons. The power grasp may require relative rotation control to avoid awkward postures, since the biomechanical constraints of the hand and arm can limit the range of rotation. Also, the muscle tension required to press or hold an integrated switch can interfere with the user's ability to manipulate the input device. Using a footpedal to separate the button from the device can be an effective alternative [80], and can help to encourage use of the precision grasp.

Device acquisition time: The results reported here suggest that the mouse-based techniques are slower for the experimental rotation matching task. Nonetheless, in a work routine that uses the mouse frequently, the time to switch over to a 3D device, and then back to the mouse, might become a dominating factor. This could result in longer times for users to accomplish their overall goals, even though individual virtual object rotation tasks could be performed more quickly. This illustrates why one cannot conclude from this study that "3D input is best," but that one must also consider the context of surrounding tasks and the intended users.

It may be possible to eliminate acquisition times in some cases by using a 3D device in the nonpreferred hand. This offers the possibility of working with a 3D device and a mouse at the same time [107], or with a pair of 3D devices simultaneously [80][141][149].

6.10 Conclusion

Revisiting the original hypotheses, this study provides clear evidence that test users were able to take advantage of the integrated control of 3D orientation input to perform a rotation matching task more quickly than with 2D input techniques, despite years of prior experience with the mouse. More importantly, with possible slight qualifications for female test users or during initial exposure to the task, there were no statistically reliable accuracy discrepancies between any of the input techniques, demonstrating that 3D orientation input is fast without necessarily sacrificing accuracy.

The quantitative data did not lend any support to the hypothesis that the physical shape of the 3D input device can be an important design consideration, but the subjective ranks which test users assigned to the interfaces spoke definitively on this point. Users reported diametrically opposed impressions of input sensors which differed only in their physical housing.

The study did not provide evidence that the Arcball performs any better than the Virtual Sphere. Test users tended to prefer the Arcball over the Virtual Sphere, but this advantage was not significant statistically. These non-results do confirm, however, that the Arcball’s possibly confusing concept of half-length arcs (as described in section 6.7.2 on page 124) does not cause any obvious usability problems [153].

This study has focused on manipulative techniques to achieve 3D rotation, but of course direct manipulation (whether 2D or 3D) is not the only approach to orienting and viewing objects. Even with a good interface, specification of orientation can be a difficult task for some users, so it makes sense to obviate the task when appropriate.

For example, the Sketch system [186] (*fig. 2.3 on page 17*) provides a casual, inexact interface and a cleverly chosen set of heuristics which often allows the system to select views of objects based on fast, relatively imprecise 2D mouse clicks. The Jack system

[135] takes a similar approach by automatically generating unobstructed views of selected objects. Additional work is needed to understand what classes of user tasks are appropriate for each technique, and to determine how to best mix the two styles in the same interface.

I have attempted to provide a careful analysis of performance in terms of time and accuracy and to provide detailed observations of usability problems one might expect for novice test users performing tasks similar to the experimental rotation matching task. I have also attempted to synthesize specific qualitative observations from this study and my previous experience with designing 3D input devices for rotation control to suggest some tentative design parameters which can influence how users will employ these devices. These contributions should prove useful to the growing number of interface designers who will incorporate 3D rotation control techniques into their applications.

"The obvious fact that people have found two hands useful, even essential, in daily life has been ignored by the user-interface community, including those working with reality goggles... There is almost no scientific literature on how we use two hands."

Myron Krueger [104], page 172.

Chapter 7

Issues in Bimanual Coordination

7.1 Introduction and goals for the studies

This chapter discusses some experiments which analyze manipulation of physical objects with both hands. This includes an analysis of real-world tasks, such as handwriting and sketching, as well as a formal experimental task which involves using a tool in one hand to point at a target object held in the other hand. An underlying goal of all of these studies is to gather both formal and informal evidence that Guiard's theoretical framework applies to a wide range of tasks of interest for human-computer interaction, as well as the specific tasks supported by the props-based interface.

7.2 Experiment 1: Qualitative analysis of handwriting and sketching

My handwriting and sketching analyses were suggested by Guiard's result that exclusion of the nonpreferred hand from a handwriting task reduces the writing speed of adults by some 20% [67]. I believed that a similar result could be demonstrated for drawing or sketching tasks, which of course are important for the many artists, architects, or engineers who sketch out design concepts using digital media. The general implication is that using two hands is relevant to far more than 3D manipulation. Designers of hand-held computers, sketching software, gesture or pen-based interfaces, and drafting-board style displays such as the Immersadesk [49] or the Immersive Workbench [52] should be designing with a careful consideration for the issues raised by two-handed interaction.

The handwriting and sketching studies emphasize qualitative results. Through subject observation and videotape analysis, my pilot studies quickly led to qualitative insights with immediate application to interface design, but I found it was difficult to formalize these results. There are many interacting within and between-subjects factors that must be carefully controlled to produce a clean quantitative result, but many of these factors (which are outlined in section 7.2.1) have little or no relevance to interface design. As a result I decided that discovering, understanding, and measuring or controlling for these factors in formal experimental designs was beyond the intended scope of this thesis.

The handwriting and sketching studies individually consisted of 6 rounds of experimental work with small numbers of subjects. In total, 29 right-handed subjects (9 friends and colleagues and 20 subjects drawn from the psychology department's subject pool) participated. Each round of experimental work tested one or more variants of the experimental tasks of (1) filling a page with handwriting, (2) filling out a form, (3) sketching an object in the environment, or (4) drawing a circle. Rather than focusing on the specific details of each round of tests, this discussion pools my observations across all subjects, and gives the specifics of the experimental tasks where appropriate.

7.2.1 Handwriting

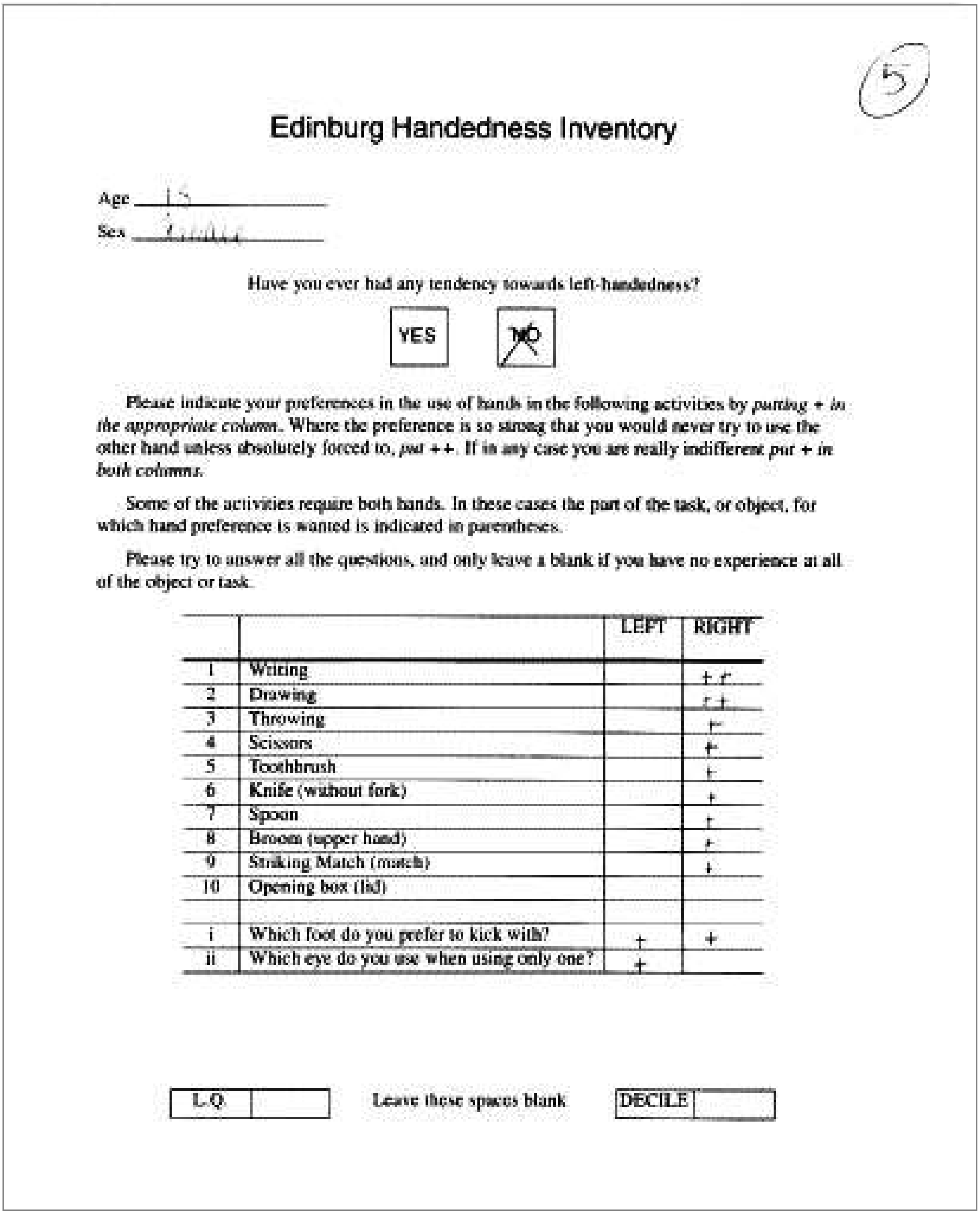
Subjects were asked to make a copy of a passage from a printed page onto a separate blank sheet of paper. I found that subjects almost always place the page for copying to the left of the page for writing, and use the left hand to keep track of their current position in the text. The actual handwriting itself is done unimanually. Clearly, maintaining split attention between the pages is a more difficult task than the handwriting itself, and because of this subjects prefer to recruit the left hand for that task.

The left hand can participate in manipulation of the page when the user is writing on dictation or from memory (such as when writing a memorized sentence or during composition of original prose). For example, of 4 subjects whom I asked to write a memorized phrase on every third line of a college-ruled page, all 4 subjects used the left hand to initially orient the page, and 3 of the 4 subjects used the left hand in a dynamic role to reposition the page at least once during the actual writing. The fourth subject used her left hand to lean on the desk and prop up her head. Most people *can* write unimanually if necessary, and a few even prefer to do so. In this regard, the task of writing on an already selected line is probably an oversimplified representation of the tasks of interest to interface designers. Nonetheless, even for such a simple task, most people do prefer to use both hands.

As part of the experimental protocol, subjects filled out a questionnaire with 14 questions on preferred hand usage (*fig. 7.1*). Subjects did not realize that this doubled as an experimental task which I captured on videotape[1]. Filling out a form is an interesting variant of the handwriting task, because it involves the same basic motor skills, but the task contains both macrometric and micrometric components. Answering a single question requires a small movement to “fill in the blank”, but moving to the next question requires a potentially large movement to another area of the page.

---

[1] I told subjects about the video camera during the experimental debriefing. I obtained signed consent forms from each subject before analyzing the tapes. Subjects had the right to request that the tape be erased immediately.



5

Edinburg Handedness Inventory

Age 15

Sex female

Have you ever had any tendency towards left-handedness?

YES [✓ NO]

Please indicate your preferences in the use of hands in the following activities by *putting + in the appropriate column*. Where the preference is so strong that you would never try to use the other hand unless absolutely forced to, *put ++*. If in any case you are really indifferent *put + in both columns.*

Some of the activities require both hands. In these cases the part of the task, or object, for which hand preference is wanted is indicated in parentheses.

Please try to answer all the questions, and only leave a blank if you have no experience at all of the object or task.

		LEFT	RIGHT
1	Writing		++
2	Drawing		++
3	Throwing		+
4	Scissors		+
5	Toothbrush		+
6	Knife (without fork)		+
7	Spoon		+
8	Broom (upper hand)		+
9	Striking Match (match)		+
10	Opening box (lid)		
i	Which foot do you prefer to kick with?	+	+
ii	Which eye do you use when using only one?	+	

L.Q. ☐ Leave these spaces blank DECILE ☐

Figure 7.1 Questionnaire for the "filling out a form" task.

By analyzing the videotapes, I counted the number of times that 7 subjects moved the page while filling out the questionnaire. Five subjects moved the page at least once; 2

subjects did not move the page at all, while 2 subjects moved the page for essentially every item on the questionnaire (12 or more movements for 14 questions). When subjects moved the page, they moved it in a specific pattern: the preferred hand holding the pen stayed in a very small working area, while the nonpreferred hand shifted the page to bring the work to the preferred hand. Clearly, the nonpreferred hand helps to reduce the working volume of the preferred hand to a small, comfortable, well-practiced region.

In unpublished work, Athenes [6][7] (working with Guiard) has found that all of the following factors can influence the speed of handwriting[1]:

- *Handedness:* There are three distinct classes of handwriters: right handers, left-handed inverters, and left-handed non-inverters (*fig. 7.2*). Each class has its own distinct handwriting strategy [72][6].
- *Paper weight:* Heavier paper has more friction with the desk surface and is easier to use unimanually.
- *Paper size:* Extremely small pages, such as post-it notes, are difficult to write on because the nonpreferred hand can hardly help at all.
- *Line Position on the Page:* The lower the position on the page, the more important the contribution of the nonpreferred hand.
- *Margin Size:* The closer the pen gets to the edge of the paper, the more difficult it is to write with one hand.

As suggested by the above list, isolating the “true” contribution of the nonpreferred hand to handwriting tasks requires consideration of many mundane factors and is an almost

---

[1] I would like to acknowledge private communication in which Yves Guiard discussed these issues.

impossible endeavor. Many similar factors influence the drawing and sketching tasks discussed below. This is why the present studies emphasize my qualitative findings.

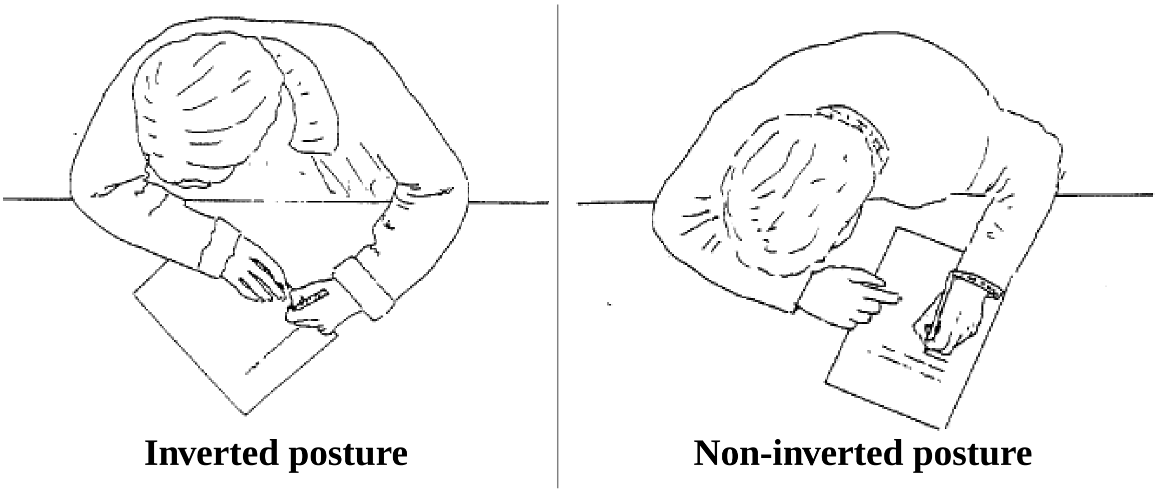


Figure 7.2 Writing posture of left-handed inverters and non-inverters [6].

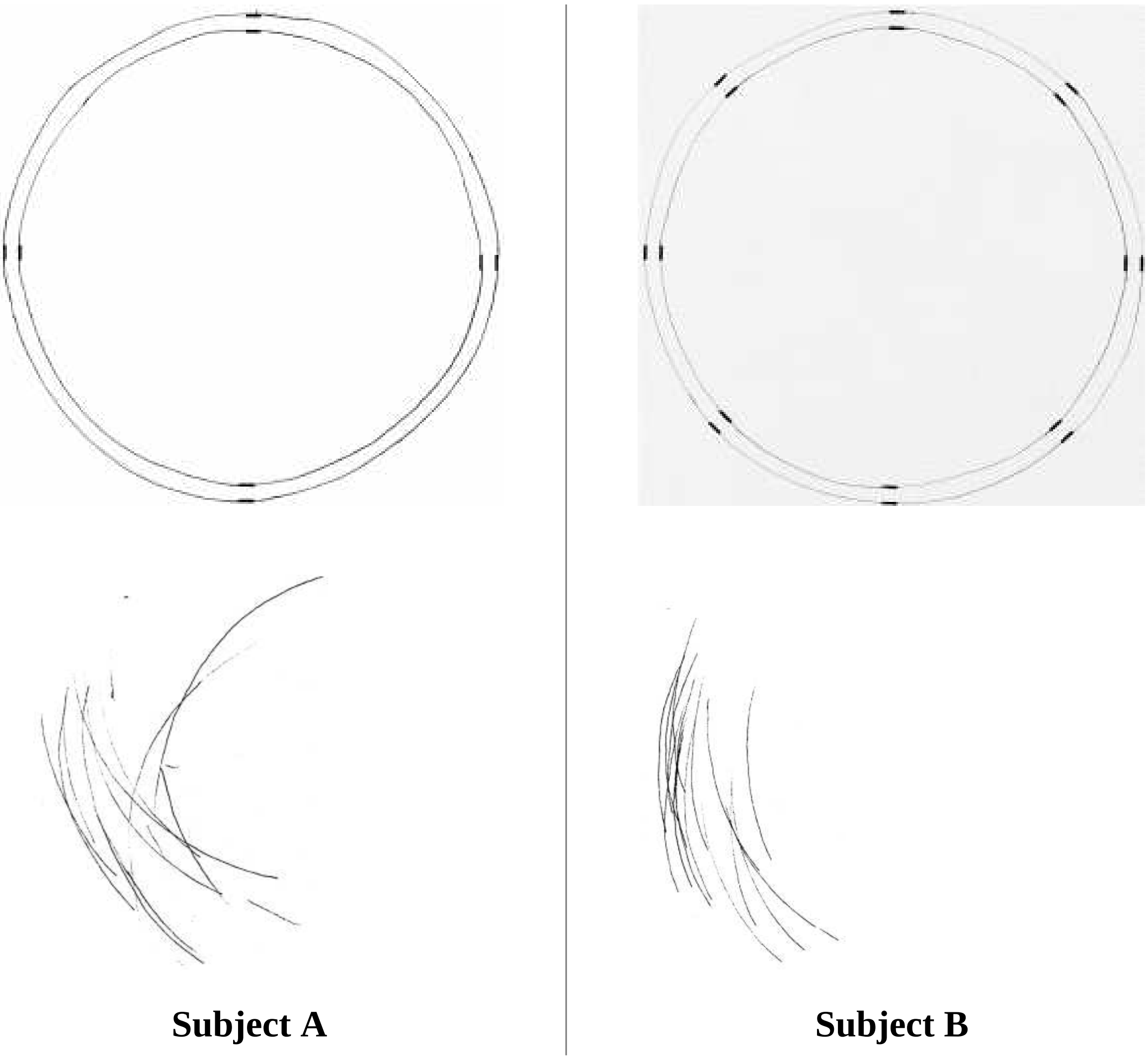
7.2.2 Drawing and sketching

Several experimental rounds studied a circle sketching task. I chose circle sketching because (1) circles are a simple, well-defined geometric form that even non-artists can attempt to draw; and (2) drawing a circle naturally requires strokes of the pen or pencil at a variety of angles. Subjects were asked to sketch circles ranging from 7 to 10 inches in diameter. I tested the following experimental conditions:

- *Unimanual Restricted (UR)*: Subjects were instructed to keep their left hand in their lap. The page was also taped to the desk surface so that it would not move (either accidentally or by the subject's intent) during the task.
- *Bimanual Restricted (BR)*: Subjects were instructed to keep their left hand on the desk surface, but the page was again taped to the desk surface so that the page

could not be rotated. Subjects were allowed to position and move the left hand as desired.

- *Bimanual Free (BF)*: Subjects were instructed to perform the task in “whatever way comes naturally.” The page was not secured to the desk surface.



Subject A

Subject B

Figure 7.3 Examples of circles drawn versus individual pen strokes[1].

---

[1] The circles shown in this figure were 7.5 inches in diameter (outer circle) and 7.0 inches (inner circle) as originally drawn by the subject.

As expected, in the BF condition most subjects did naturally use their left hand to rotate the page and thereby extend the working range of the right hand. This strategy allowed subjects to repeatedly draw an almost identical stroke until the circle was complete. Subjects would perform most of the work on the left hand side of the page. To illustrate this tendency, I placed transfer paper (carbon paper) on the back of a specially prepared desk blotter; subjects sketched a circle (*fig. 7.3, top row*) on a page on top of this blotter. The transfer paper captured the impression of the subject's pen strokes relative to the blotter (*fig. 7.3, bottom row*). As the examples shown in figure 7.3 demonstrate, most of the arcs are in the same direction and take up a small area relative to the size of the circle itself. Note that figure 7.3 shows examples from two separate subjects performing slightly different tasks (subject A had 4 registration marks to guide the drawing of each circle, while subject B had 8 such registration marks).

In the BR and UR conditions where rotating the page was not possible, subjects often had difficulty in the upper right-hand quadrant of the circle. This is exactly the point where the inflection of the circle goes against the natural biomechanical curves generated by the wrist and fingers of the right hand. To compensate for this biomechanical awkwardness, subjects would try to move their entire torso relative to the page to produce a better angle for the action of the right hand. Having the left hand in the lap, as required by the UR condition, made it even more difficult for subjects to counter-rotate their body to adjust for the fixed page. As one subject commented, "I usually put my hand on the table for support." Thus, the UR condition reveals that the left hand normally plays a postural role as well.

There was a tendency for the bimanual conditions (especially the BF condition) to produce better quality results (the circle looked better), as self-assessed by the subjects or by an independent judge. However, the overall difference between conditions was not that impressive. The times to complete the task varied tremendously between subjects; some who were in a hurry would complete every circle in 20 seconds, no matter how bad they looked, while others would spend 3 minutes perfecting their work every time. Thus, the experimental tasks did not adequately control for time/accuracy trade-offs.

While the above results suggest that using both hands for sketching tasks may have some benefits, using a second hand might also be a liability, because each hand occludes part of the circle that the subject is sketching: the subject loses the overall context of the entire circle. In the UR condition, subjects seemed to have more difficulty precisely controlling hand movements, but they could see more of the circle, since their right hand (for at least half of the time) is to the right of the circle and does not occlude it at all. In the BF condition, it was easier to make the outside of the circle “smooth,” but it was harder to make it truly circular. When subjects tried to draw free-hand circles in the BF condition without any registration marks to guide their work, many of the resulting circles were egg-shaped.

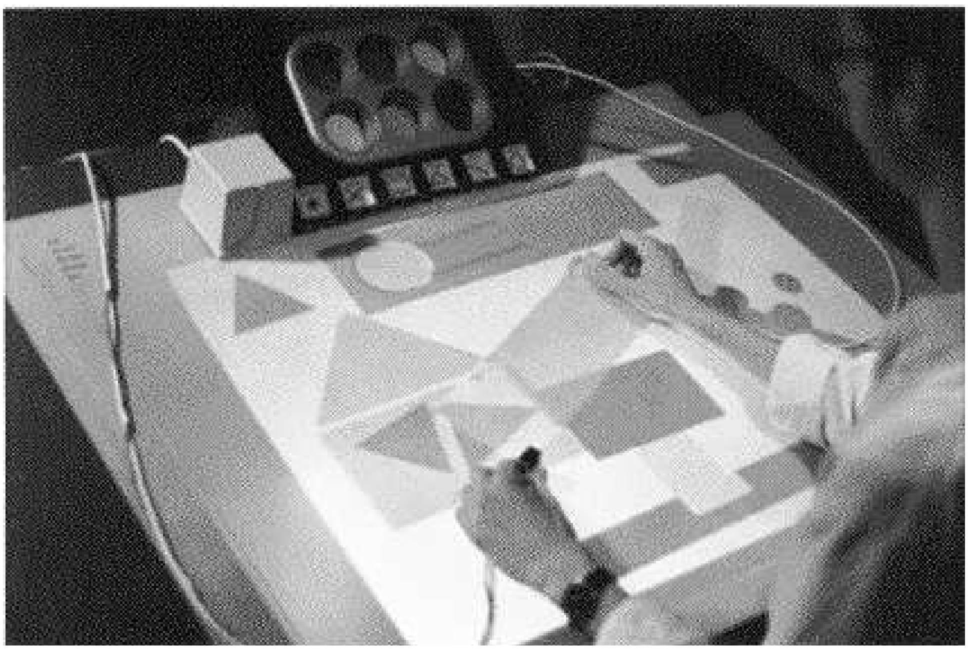


Figure 7.4 An example two-handed application [59] using the Active Desk.

The same issue arises in interface design when using both hands for direct input on a large display surface such as the Active Desk (fig. 7.4). In unpublished work, Buxton tried to reproduce the results of Kabbash [95] using direct input on the Active Desk display. Kabbash (as originally discussed in section 2.6.2.2 on page 36) found that users could connect and color a series of randomly appearing dots much faster with the two-handed ToolGlass technique [15] than with one-handed techniques; Kabbash produced these results using indirect input devices in conjunction with a standard monitor. But on the Active Desk, Buxton found that the targets would sometimes randomly appear in parts of the display occluded by the user's arms, causing subjects to become extremely confused. Thus, interface designs which use two hands for direct input need to be aware of the problems introduced by occlusion of the display surface and potential problems caused by the resulting loss of context[1].

As a final example, during one experimental session, I asked two subjects with minors in art to sketch an object which I placed at the rear of the desk surface. Although these subjects were allowed to use both hands, neither subject rotated the page while they were capturing the general outline of the object. However, once they started to shade it and "overstrike" some of their initial strokes, they began to frequently rotate the page with the left hand. Thus, when drawing an object from the environment, rotating the page is a detriment because it would require a mental rotation to align the page and the environment; but when attention is focused primarily on the page itself, these subjects found rotating the page to be a natural and efficient task strategy.

Although these two subjects had some artistic skill, this was not the case for most of the subjects which I recruited for these studies. I speculate that for skilled artists performing a sketching task, using the left hand to rotate the page or otherwise assist the

---

[1] A simple reformulation of Buxton's experiment to produce targets at predictable locations did manage to reproduce the results from Kabbash. I would like to acknowledge personal communication during which Bill Buxton discussed this experiment on the Active Desk.

activities of the right hand is even more important than for unskilled subjects. When asked directly about this issue, one professional artist (who was *not* a participant in this study) commented, “I can’t draw unless I can rotate the page.”

7.2.3 Summary of handwriting and sketching observations

When the task constraints allow for use of both hands, people will naturally involve both hands in the execution of handwriting and sketching tasks. Even when the left hand is not directly involved in the manipulation itself, it often plays a role in postural support or maintaining split attention, such as the observed use of the left hand as a place holder when copying text from one page to another.

Furthermore, as illustrated by the impressions of individual pen strokes relative to the desk (*fig. 7.3*), when the left hand helps to manipulate the page, the patterns of hand use are quite consistent and predictable across subjects. By dynamically positioning and orienting the page, the nonpreferred hand extends the working range of the “sweet spot” for the preferred hand.

There are problems which these studies do not address. Many uninteresting factors obscure the time, precision, or postural advantages (if any) of using two hands for these tasks, and a formal experiment would need to control for these factors. The tasks which I tested were also somewhat open-ended, and did not adequately control for time/accuracy trade-offs. This, in addition to the large between-subject variability in terms of handwriting speed, artistic ability, and so forth, means that these studies were not well suited to produce quantitative results.

7.3 Experiment 2: Cooperative bimanual action

7.3.1 Overview

Experiment 2 explores cooperative bimanual action. Right-handed subjects manipulated a pair of physical objects, a *tool* and a *target object*, so that the tool would touch a target on the object (*fig. 7.5*). For this task, there is a marked specialization of the hands. Performance is best when the left hand orients the target object and the right hand manipulates the tool, but is significantly reduced when these roles are reversed. This suggests that the right hand operates relative to the frame-of-reference of the left hand.

Furthermore, when physical constraints guide the tool placement, this fundamentally changes the type of motor control required. The task is tremendously simplified for both hands, and reversing roles of the hands is no longer an important factor. Thus, specialization of the roles of the hands is significant only for skilled manipulation.

Figure 7.5 A subject performing the experimental task.

7.3.2 Introduction

Two-handed interaction has become an accepted technique for “fish tank” 3D manipulation, for immersive virtual reality, and for 2D interfaces such as ToolGlass [15]. Unfortunately, there is little formal knowledge about how the two hands combine their action to achieve a common goal.

The present experiment was motivated by my experiences with the props-based interface. Informally, I observed that the operation of the interface was greatly simplified when both hands were involved in the task. But the early design stages had to consider many possible ways that the two hands might cooperate. An early prototype allowed users to use both hands, but was still difficult to use. The nonpreferred hand oriented the doll’s head, and the preferred hand oriented the cross-sectioning plane, yet the software did not pay any attention to the relative placement between the left and the right hands. Users felt like they were trying to perform two separate tasks which were not necessarily related.

I modified the interface so that relative placement mattered. The software interpreted all motion as relative to the doll’s head in the user’s nonpreferred hand, resulting in a far more natural interaction: users found it much easier to integrate the action of the two hands to perform a cooperative task. Informally, this suggests that two-handed coordination is most natural when the preferred hand moved relative to the nonpreferred hand. The current experiment formalizes this hypothesis and presents some empirical data which suggests right-to-left reference yields quantitatively superior and qualitatively more natural performance.

Beyond this specific example, interface designers in general have begun to realize that humans are two-handed, and it is time to develop some formal knowledge in support of such designs. In this spirit, the present experiment, which analyzed right-handed subjects only, contributes the following pieces of such formal knowledge:

- The experimental task, which represents a general class of 3D manipulative tasks involving a tool and a reference object, requires an asymmetric contribution of the two hands.

- For such tasks, performance is best when the right hand operates relative to the left. Reversing the roles of the hands significantly reduces performance both in terms of time and accuracy.

- When physical constraints help to perform the task, the type of motor control required fundamentally changes. This is as an especially telling finding for virtual manipulation, where the user is often equipped with a glove offering no haptic feedback, and speaks strongly for the possible benefits of using either passive or active haptic feedback (using either "props" or a force-feedback arm akin to the Phantom [147], respectively).

- Specialization of the roles of the hands is significant only for skilled manipulation. This does not imply that two-handed input will be ineffective for tasks which afford symmetric manipulation, but instead restricts the scope of tasks where asymmetry factors will have important design implications.

- These findings are quite robust with respect to a subject's level of motor skill. Given the wide range of skill exhibited by the participants, it is remarkable that all 16 showed the above effects to some extent.

- Qualitatively, these results held despite the strong tendency for subjects to adopt coping strategies which attempted to maintain the natural roles for the hands. For example, when roles were reversed, some subjects tried to hold the tool stationary in the left hand while moving the target object to meet it. Clearly, the constraints of the task limited the effectiveness of this strategy.

The study only included right-handed subjects because hand usage patterns in left-handers tend to be somewhat more chaotic than those in right-handers, which complicates experimental design. The issues posed by handedness are surprisingly complicated [69][74], and without a clear understanding of bimanual action in right-handers, it seems premature to address the unique behavioral strategies employed by left-handers. For example, as shown in figure 7.2 on page 136, left-handers can adopt at least one of two distinct postures, inverted or non-inverted, for handwriting. Nonetheless, I speculate that left-handers should exhibit a similar (but less consistent) pattern of findings to those reported here for right-handers.

7.4 Related work on bimanual action

In the HCI, psychology, and motor behavior literatures, experiments studying hand lateralization issues have typically been formulated in terms of hand superiority by contrasting unimanual left-hand performance versus unimanual right-hand performance [3][94][140]. While such experiments can yield many insights, they do not reveal effects which involve simultaneous use of both hands.

For truly bimanual movement, most psychology and motor behavior experiments have studied tasks which require concurrent but relatively independent movement of the hands. Example tasks include bimanual tapping of rhythms [44][134][185] and bimanual pointing to separate targets [87][117][184]. Since the hands are not necessarily working together to achieve a common goal, it is uncertain if these experiments apply to *cooperative* bimanual action.[1]

There are a few notable exceptions, however. Buxton and Myers [27] demonstrated that computer users naturally use two hands to perform compound tasks (positioning and

---

[1] For bimanual rhythm tapping, conceptually the two hands *are* working together to produce a single combined rhythm. This task, however, does not address the hypothesis of right-to-left reference in bimanual manipulation.

scaling, navigation and selection) and that task performance is best when both hands are used. Buxton [29] has also prepared a summary of issues in two-handed input.

Kabbash [95] studied a compound drawing and selection task, and concluded that two-handed input techniques, such as ToolGlass [15], which mimic everyday "asymmetric dependent" tasks yield superior overall performance. In an asymmetric dependent task, the action of the right hand depends on that of the left hand [95][67]. This experiment did not, however, include any conditions where the action of the left hand depended on the right hand.

Guiard performed tapping experiments with a bimanually held rod [69]. Subjects performed the tapping task using two grips: a preferred grip (with one hand held at the end of the rod and the other hand near the middle) and a reversed grip (with the hands swapping positions). The preferred grip yielded better overall accuracy, but had reliably faster movement times only for the tapping condition with the largest amplitude. Guiard also observed a distinct partition of labor between the hands, with the right hand controlling the push-pull of the rod, and the left hand controlling the axis of rotation.

A number of user interfaces have provided compelling demonstrations of two handed input, but most have not attempted formal experiments. Three-dimensional virtual manipulation is a particularly promising application area. Examples previously described in chapter 2 include the Virtual Workbench [138], 3Draw [141], Worlds-in-Miniature [164], PolyShop [1], and work by Shaw [149] and Multigen, Inc. [121]. There is also some interest for teleoperation applications [163]. In two dimensions, examples include Toolglass and Magic Lenses [15], Fitzmaurice's Graspable User Interface [59] (shown in figure 7.4 on page 139), and Leganchuk's bimanual area sweeping technique [108]. Bolt [17] and Weimer [180] have investigated uses of two hands plus voice input. Hauptmann [76] showed that people naturally use speech and two-handed gestures to express spatial manipulations.

7.5 The Experiment

7.5.1 Task

The subject manipulates a *tool* (either a plate or stylus) in one hand and a *target object* (either a puck, a triangle, or a cube) with the other hand (*fig. 7.6*). Each target object has a rectangular slot cut into in it, at the bottom of which is a small gold-colored target area. There are two versions of the task, a Hard task and an Easy task.

For the Hard task, the subject must mate the tool and the target object so that the tool touches only the target area (*fig. 7.7*). The target area is wired to a circuit that produces a pleasant beep when touched with the tool; if the tool misses the target area, it triggers an annoying buzzer which signals an error. The target area is only slightly larger than the tool, so the task requires dexterity to perform successfully. I instructed each subject that avoiding errors was more important than completing the task quickly.

For the Easy task, the subject only has to move the tool so that it touches the bottom of the rectangular slot on the target object. The buzzer was turned off and no “errors” were possible: the subject was allowed to use the edges of the slot to guide the placement. In this case, the subject was instructed to optimize strictly for speed.

Each subject performed the Hard and the Easy task using two different grips, a Preferred grip (with the left hand holding the target object and the right hand holding the tool) and a Reversed grip (with the implements reversed). This resulted in four conditions: Preferred Hard (PH), Preferred Easy (PE), Reversed Hard (RH), and Reversed Easy (RE).

Subjects were required to hold both objects in the air during manipulation (*fig. 7.5*), since this is typically what is required when manipulating virtual objects. Subjects were allowed to rest their forearms or wrists on the table, which most did.

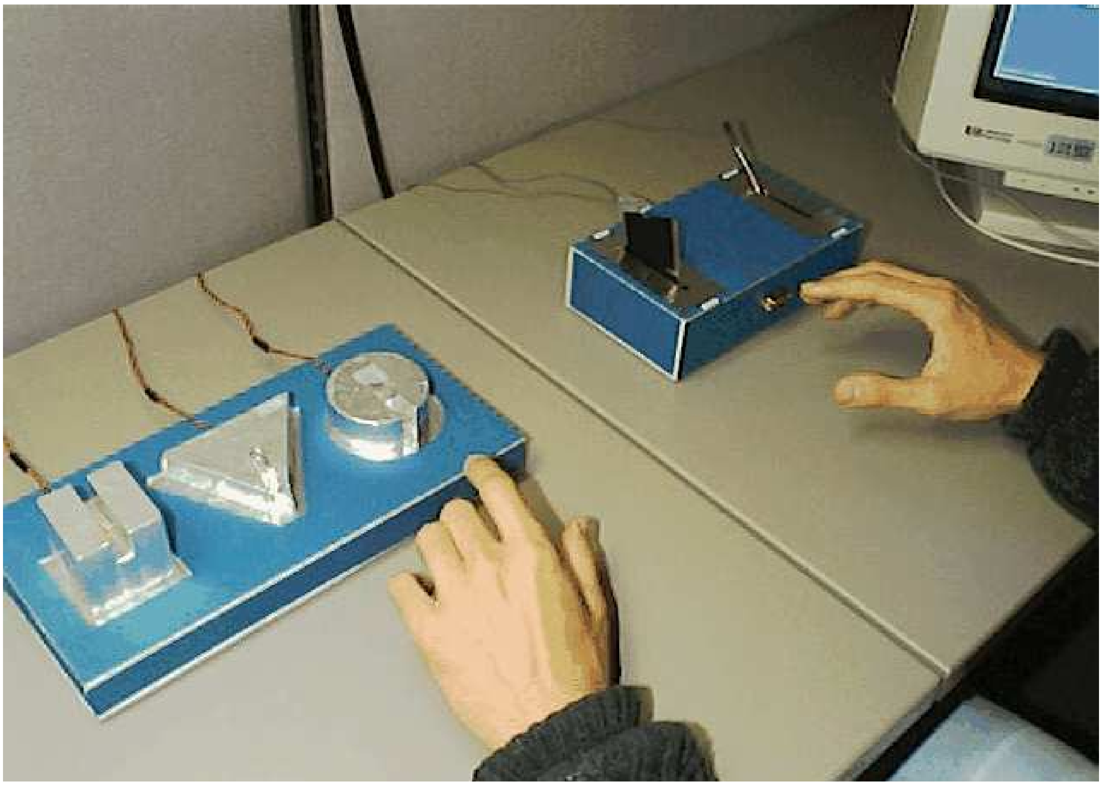


Figure 7.6 Configuration for Experiment 2.

7.6 Experimental hypotheses

The experimental hypotheses were suggested by my experiences with the props-based interface and formalized with the help of Guiard's Kinematic Chain (KC) model. The high-level working hypothesis for this experiment is that the KC model can be used to reason about two-handed 2D or 3D tasks and interface design.

The specific hypotheses for this experiment are as follows:

H1: The Hard task is asymmetric and the hands are not interchangable. That is, the Grip (preferred, reversed) used will be a significant factor for this task.

H2: For the Easy task, the opposite is true. Reversing roles of the hands will not have any reliable effect.

H3: The importance of specialization of the roles of the hands increases as the task becomes more difficult. That is, there will be an interaction between Grip (preferred, reversed) and Task (easy, hard).

H4: Haptics will fundamentally change the type of motor control required.

7.6.1 Subjects

Sixteen unpaid subjects (8 males, 8 females) from the Psychology Department subject pool participated in the experiment. Subjects ranged from 18 to 21 (mean 19.1) years of age. All subjects were strongly right-handed based on the Edinburgh Handedness Inventory [127].

7.6.2 Experimental procedure and design

Figure 7.6 shows the overall experimental set-up. The experiment was conducted using instrumented physical objects, rather than virtual objects. Since the purpose of the experiment is to look at some basic aspects of bimanual motor control, using physical objects helps to ensure that the experiment is measuring *the human*, and not artifacts caused by the particular depth cues employed, the display frame rate, device latency, or other possible confounds associated with virtual manipulation. The physical objects also provided the haptic feedback needed to test hypothesis H4.

The experiment began with a brief demonstration of the neurosurgical props interface to engage subjects in the experiment. I suggested to each subject that he or she should “imagine yourself in the place of the surgeon” and stressed that, as in brain surgery, accurate and precise placement was more important that speed. This made the experiment more fun for the subjects, who would sometimes joke that they had “killed the patient” when they made an error.

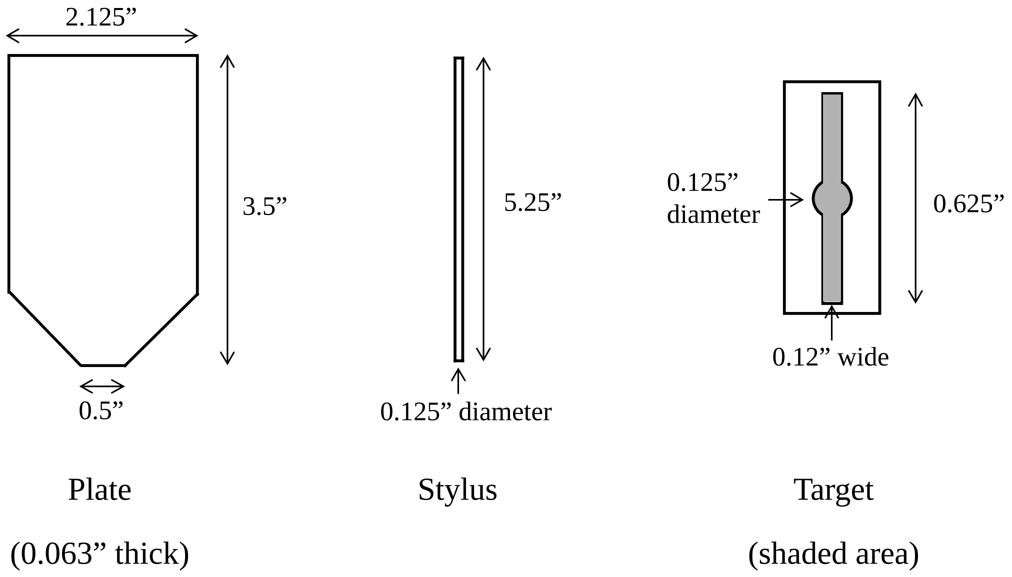


Figure 7.7 Dimensions of the plate tool, the stylus tool, and target areas.

There were two tools, a plate and a stylus, and three target objects, a cube, a triangle, and a puck (*figs. 7.7, 7.8*). Using multiple objects helped to guarantee that the experimental findings would not be idiosyncratic to one particular implement, as each implement requires the use of slightly different muscle groups. Also, the multiple objects served as a minor ruse: I did not want the subjects to be consciously thinking about what they were doing with their hands during the experiment, so they were initially told that the primary purpose of the experiment was to test which shapes of input devices were best for two-handed manipulation.

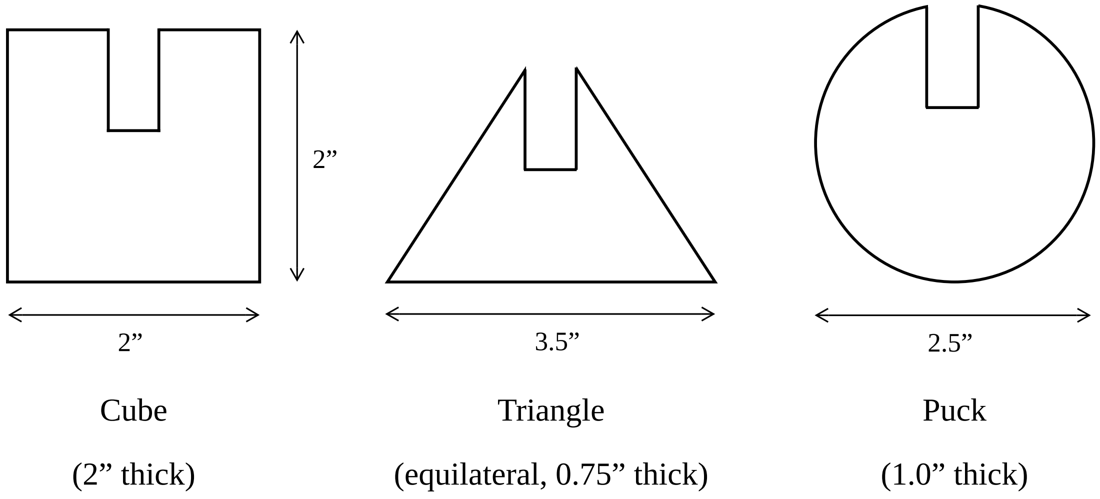


Figure 7.8 Dimensions of the Cube, Triangle, and Puck target objects.

The subject next performed a practice session for the Hard task, during which I explained the experimental apparatus and task. This session consisted of 6 practice trials with the Preferred grip and 6 practice trials with the Reversed grip[1].

For the experimental trials, a within-subjects latin square design was used to control for order of presentation effects. For each of the four experimental conditions, subjects performed 24 placement tasks, divided into two sets of 12 trials each. Each set included two instances of all six possible tool and target combinations, presented in random order. There was a short break between conditions.

For the Hard task, the dependent variables were time (measured from when the tool is picked up until it first touches the target area) and errors (a dichotomous pass / fail variable). For the Easy task, since no errors were possible, only time was measured.

---

[1] This also doubled as a lateral preferences assessment, to ensure that each subject actually did prefer the “Preferred” grip to the “Reversed” grip.

7.6.3 Details of the experimental task and configuration

For each trial, the computer display (at the right of the working area) simultaneously revealed a pair of images on the screen, with the objects for the left and right hands always displayed on the left and right sides of the screen (*fig. 7.9*).

Two platforms were used, one to hold the tools and one to hold the target objects (*fig. 7.6*). The tool platform was instrumented with electrical contact sensors, allowing the apparatus to detect when the tool was removed from or returned to the platform. Returning the tool to the platform (after touching the target) ended the current trial and displayed a status report. The subject initiated the next trial by clicking a footpedal.

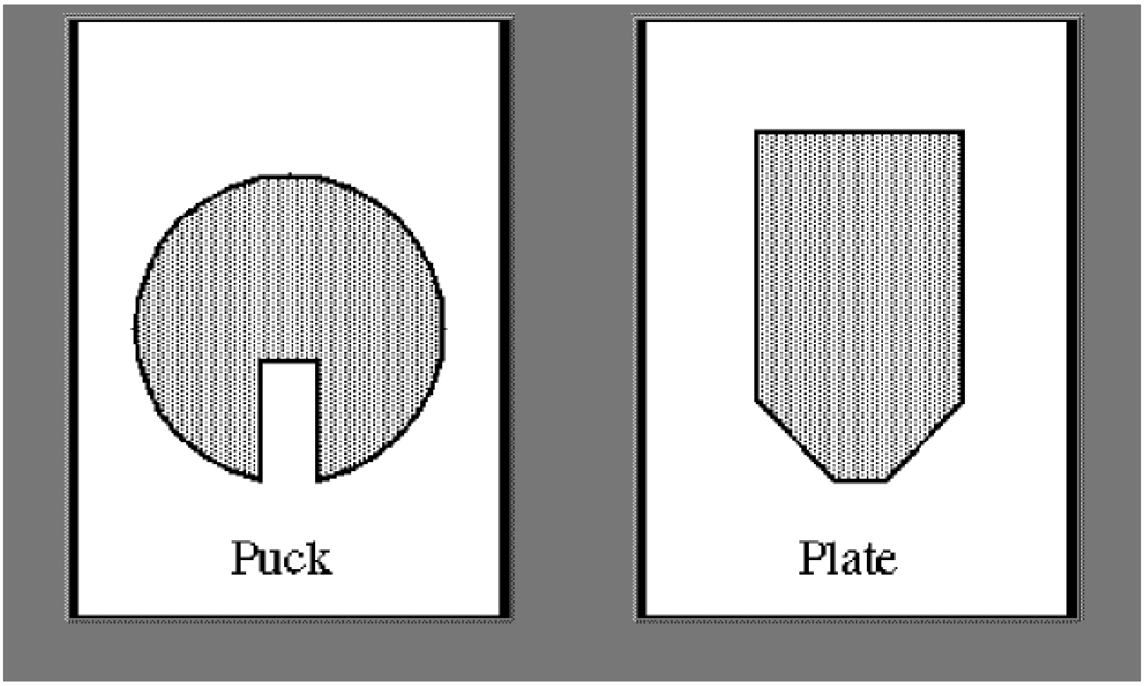


Figure 7.9 Sample screen showing experimental stimuli.

Each subject was seated so that the midline of his or her body was centered between the two platforms. The tool platform was flipped 180° during the Reversed conditions, so that the plate was always the closest tool to the objects. The platforms were positioned one foot back from the front edge of the desk, and were spaced 6” apart.

Figure 7.8 shows the dimensions for the cube, triangle, and puck target objects. Each object was fitted with an identical target (*fig 7.7, right*) which was centered at the bottom of the rectangular slot (0.75” deep by 0.375” wide) on each object. The objects were machined from delrin (a type of plastic) and wrapped with foil so they would conduct electricity. The target area and the foil were wired to separate circuits; some capacitance was added to each circuit to ensure that even slight contacts would be detected.

When using the plate, subjects were instructed to use the entire 0.5” wide tip of the plate to touch the target. For the stylus, the subject was told to touch the rounded part of the target area (the stylus was thicker than the other parts of the target, as shown in figure 7.7, and thus only the central rounded part of the target area could be touched without triggering an error).

7.6.4 Limitations of the experiment

There are a couple of factors which limit the sensitivity of this experiment. First, ideally the experiment would present a range of controlled difficulties analogous to the Index of Difficulty (ID) for Fitts’ Law [114]. Fitts’ law relates the movement time for one hand to two quantities, the amplitude A of the movement and the width W of the intended target. Together, A and W can be used to compute ID, the index of difficulty for the movement. But Fitts’ Law applies to movement of one hand, and I am not aware of any adaptations which could handle movement of both hands together. Instead, the experiment uses an *easy* versus *hard* difficulty distinction.

Second, the accuracy measurements yield a dichotomous pass / fail outcome. Thus, the apparatus captures no quantitative information about the magnitude of the errors made when the subjects missed the target in the Hard conditions. Even given these limitations, the experimental results are quite decisive. Therefore, I decided to leave resolution of these issues to future work, and to demonstrate some effects with the simplest possible experimental design and apparatus.

7.7 Results

For each condition, only the second set of 12 trials was used in the data analysis, to minimize any confounds caused by initial learning or transfer effects across conditions.

A straightforward analysis of the Hard task shows a strong lateral asymmetry effect. For both the plate and the stylus tools, 15/16 subjects performed the task faster in the PH condition than in the RH condition (significant by the sign test, $p < .001$). The difference in times is not due to a time / accuracy trade-off, as 15/16 subjects (using the plate) and 14/16 subjects (using the stylus) made fewer or the same amount of errors in the PH condition vs. the RH condition.

For the Easy task, as predicted by Hypothesis 2, the lateral asymmetry effect was less decisive. For both the plate and the stylus tools, 11/16 subjects performed the task faster in the PE condition than in the RE condition (not a significant difference by the sign test, $p > .20$). For at least one of the tools, 6/16 subjects performed the task *faster* in the RE condition vs. the PE condition.

Figure 7.10 summarizes the mean completion times and error rates. No errors were possible in the Easy conditions. In the Hard conditions, the relatively high error rates resulted from the difficulty of the task, rather than a lack of effort. I instructed subjects that “avoiding errors is more important than speed,” a point which I emphasized several times and underscored by the analogy to performing brain surgery.

Condition	**Mean**	**Std. dev.**	**Error rate**
Preferred Easy (PE)	0.76	0.15	--
Reversed Easy (RE)	0.83	0.19	--
Preferred Hard (PH)	2.33	0.77	43.9%
Reversed Hard (RH)	3.09	1.10	61.1%

Figure 7.10 Summary of mean completion times and error rates.

7.7.1 Qualitative analysis

Before proceeding with a full statistical analysis, it seems appropriate to first discuss some of the qualitative aspects of the experiment. Some of the subjects were videotaped; the qualitative observations presented here were based on these tapes as well as handwritten notes.

I observed three patterns of strategies in experimental subjects when they were performing the Hard task:

- *Maintaining natural roles of hands*: In the RH condition, some subjects tried to perform the task by "holding the [left-hand] tool steady and bringing the [right-hand] object to meet it."

- *Tool stability*: Also in the RH condition, many subjects adjusted their left-hand grip to be as close to the tip of the tool as possible. This helped to reduce the effect of any left-hand unsteadiness.

- *Having the right view of the objects*: The target was at the bottom of a slot, so there was a restricted set of views where the subject could see the target. For the Hard task, subjects often performed the task with edge-on or overhead views, sometimes holding one eye closed to get the best view of the tool tip and target.

Subjects usually performed the RH task differently than the PH task. When using the Preferred grip, the left hand would first orient the object, and the right hand would then move in with the tool, so that at the time of contact the target object was usually stationary and only the tool was in motion. But in the Reversed grip, there often were several phases to the motion. The right hand would first orient the object, and the left hand would approach with the tool; but then the left hand would hesitate and the right hand would move towards it. During actual contact with the target, both the tool and the object were often in motion.

At first glance, it would seem that the primary difference between the RH and the PH conditions was the left hand's unsteadiness when handling the tools. For at least some of the subjects, however, it also seemed that the right hand had difficulty setting the proper orientation for the action of the left hand. So the right hand was best at fine manipulation, whereas the left hand was best at orientating the target object for the action of the other hand.

For the Easy tasks, subjects performed the task quickly and without much concentration, since they could rely on the physical constraints of the slot to guide the tool. Subjects were divided about whether or not the RE task was unnatural. Some thought it was "definitely awkward," others thought it was "fine." At least one subject preferred the Reversed grip; this preference was confirmed by a small Reversed grip advantage in the quantitative data.

Finally, when switching to the Hard task after performing a block of the Easy task, subjects often took several trials to adjust to the new task requirements. Once subjects became used to relying on physical constraints, it required a conscious effort to go back. To assist this transition, I instructed subjects to "again emphasize accuracy" and to "focus initially on slowing down."

7.7.2 Detailed statistical analysis

I performed a 2 × 3 × 2 × 2 analysis of variance (ANOVA) with repeated measures on the factors of Tool (plate or stylus), Object (cube, puck, or triangle), Task (easy or hard), and Grip (preferred or reversed), with task completion time as the dependent variable. Significance levels are summarized in figure 7.11.

Factor	**F statistic**	**Significance**
Grip	$F_{(1,15)} = 38.73$	$p < .0001$
Task	$F_{(1,15)} = 66.60$	$p < .0001$
Tool	$F_{(1,15)} = 5.22$	$p < .05$
Object	$F_{(2,30)} = 3.33$	$p < .05$
Grip × Task	$F_{(1,15)} = 24.83$	$p < .0005$
Tool × Task	$F_{(1,15)} = 16.57$	$p < .001$
Object × Task	$F_{(2,30)} = 2.52$	$p < .10$, n.s.
Grip × Task × Tool	$F_{(1,15)} = 5.11$	$p < .05$

Figure 7.11 Significance levels for Main effects and Interaction effects.

Overall, the preferred Grip was significantly faster than the reversed Grip and the easy Task was significantly faster than the hard Task. The Tool and Object factors were also significant, though the effects were small. The plate Tool was more difficult to position than the stylus: this reflects the requirement that the subject must align an additional degree of freedom with the plate (rotation about the axis of the tool) in order to hit the target. The cube Object was somewhat more difficult than the other Objects.

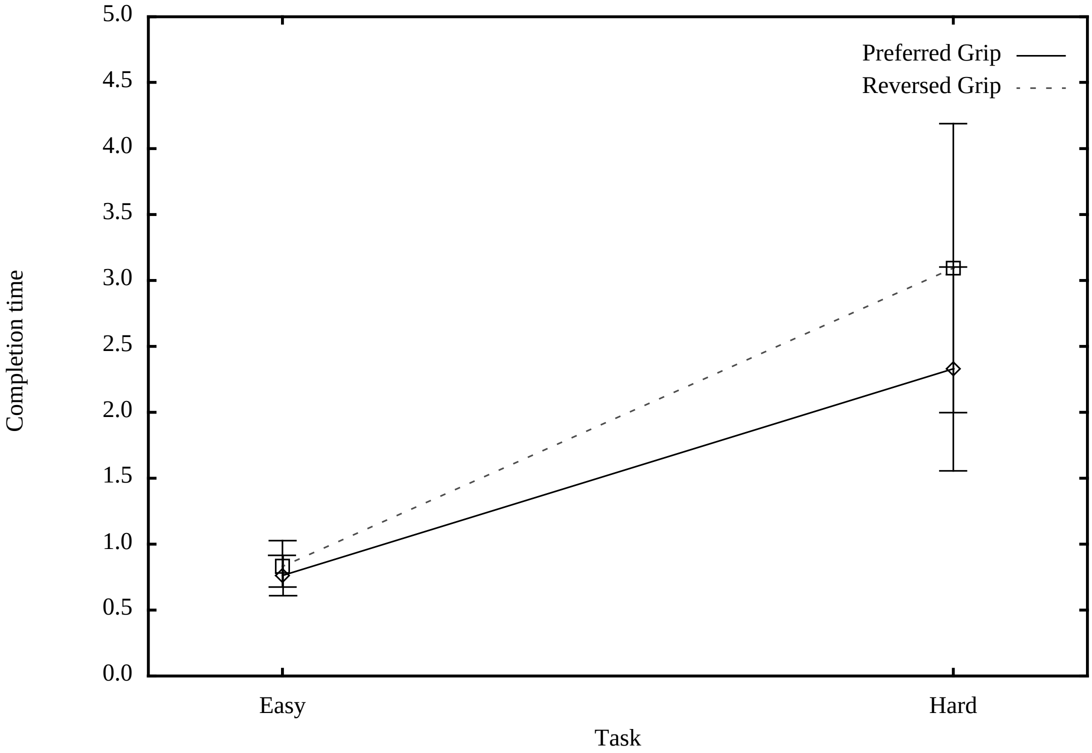


Figure 7.12 The Task X Grip interaction.

The ANOVA revealed a highly significant Grip ✕ Task interaction, which suggests that the difference between the Preferred and the Reversed grips increases as the task becomes more difficult. This speaks eloquently in favor of Hypothesis 3: the importance of specialization of the roles of the hands increases as the task becomes more difficult (*fig. 7.12*).

There was also a significant three-way Grip ✕ Task ✕ Tool interaction (*fig. 7.15*). This indicates that the extent of the Grip ✕ Task interaction varied with the tool being used (there was a larger distinction between the preferred and reversed postures with the stylus). Finally, the Tool ✕ Task interaction (*fig. 7.13*) was significant and the Object ✕ Task interaction approached significance. This suggests that the Tools and Objects differed only for the hard Task, not the easy Task.

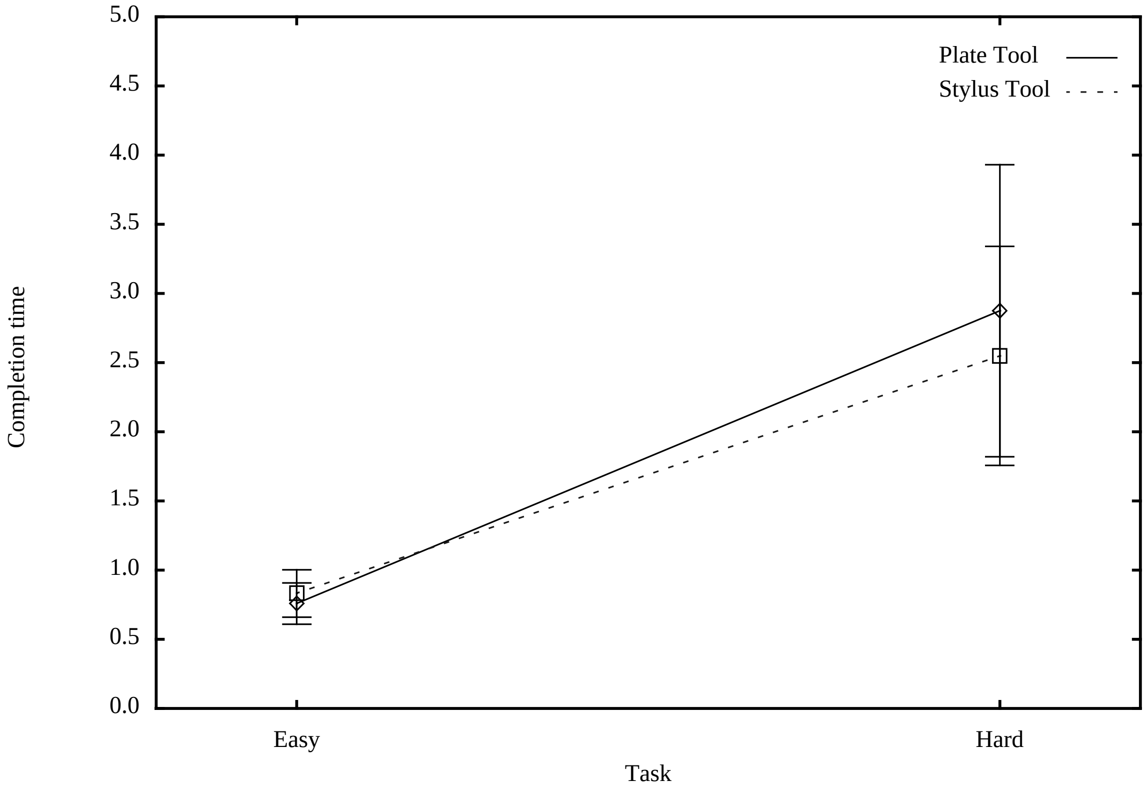


Figure 7.13 Tool X Task interaction.

Figure 7.11 reported pooled effects across the easy and hard Task and the preferred and reversed Grip. Based on the experimental hypotheses, I also compared the individual experimental conditions. These are summarized below in figure 7.14.

Contrast	**F statistic**	**Significance**
PE vs. RE	$F_{(1,15)} = 3.94$	$p < 0.10$, n.s.
PH vs. RH	$F_{(1,15)} = 33.56$	$p < 0.0001$

Figure 7.14 Significance levels for comparisons of experimental conditions.

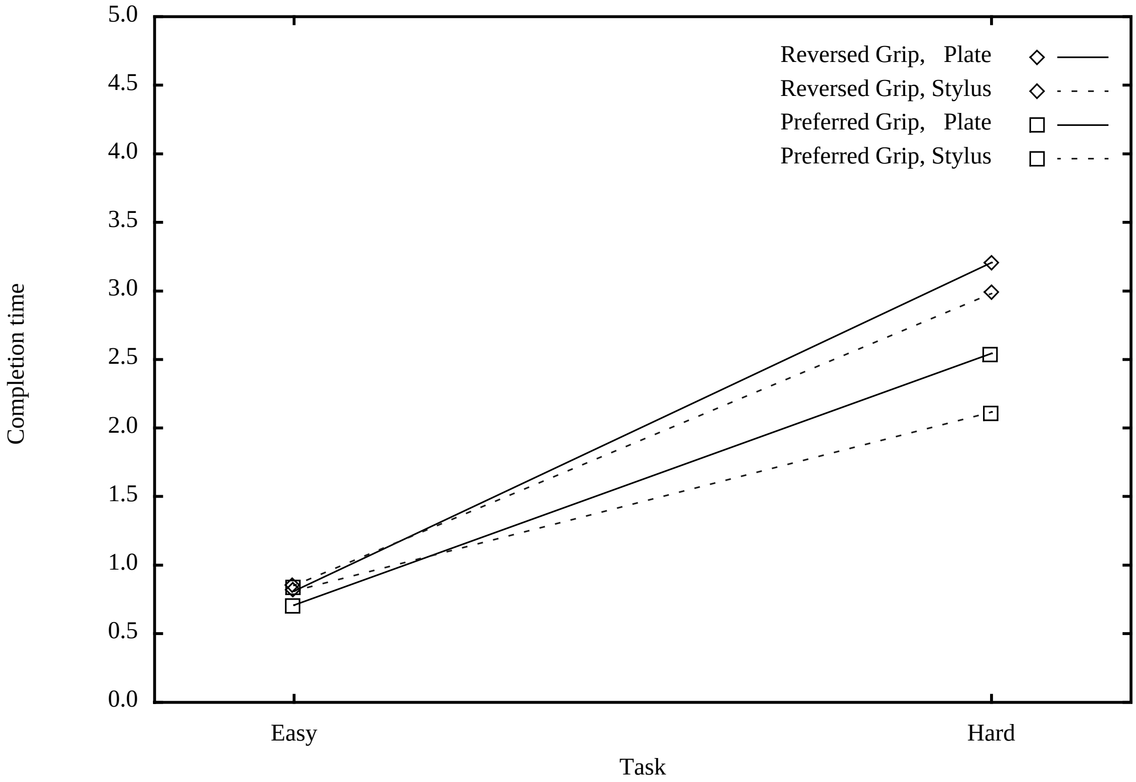


Figure 7.15 The Task X Grip X Tool interaction.

The Grip factor is significant for the Hard task (PH vs. RH), but not the Easy task (PE vs. RE). This supports Hypothesis 1: the task is asymmetric and reversing the roles of the hands has a significant effect. The Grip factor is also nearly significant for the Easy task (PE vs. RE), suggesting that Grip probably is a minor factor even in this case. This partially supports Hypothesis 2; reversing the roles of the hands has a much smaller effect for the easy task, but the study cannot confidently conclude that there is *no* effect.

7.7.3 Possibility of Order or Sex biases

I repeated the ANOVA with between-subject factors of Sex and Order of presentation to ensure that the above experimental results were not biased by these factors. The Order of the experimental conditions did not approach statistical significance, nor did the Order × Condition interaction, indicating that the results are not biased by transfer or asymmetrical transfer effects.

There was a small, but significant, main effect of Sex, along with several significant interactions (*fig. 7.16*). Although this experiment was not designed to detect sex differences, this finding is consistent with the literature, which suggests that females may be better at some dexterity tasks [74].

Factor	**F statistic**	**Significance**
Sex	$F_{(1,14)} = 5.55$	$p < .05$
Tool × Sex	$F_{(1,14)} = 12.80$	$p < .005$
Task × Sex	$F_{(1,14)} = 5.23$	$p < .05$
Tool × Task × Sex	$F_{(1,14)} = 20.90$	$p < .0005$

Figure 7.16 Overall sex difference effects.

To ensure that Sex is not a distorting factor, separate analyses were performed with N=8 male and N=8 female subjects. This is a less sensitive analysis, but the previous pattern of results still held: Grip, Task, and the Grip × Task interaction were all significant for both groups (*fig. 7.17*). Males tended to be more sensitive to which Tool was being used for manipulation, which accounts for the Tool × Sex and Tool × Task × Sex interactions (*fig. 7.16*). The Task × Sex interaction results from females being faster than males for the Hard task, but not the Easy task. Therefore, on the basis of these analyses, one can confidently conclude that the differences between the experimental conditions are not biased by Order or Sex effects.

MALES **Factor**	**F statistic**	**Significance**
Grip	$F_{(1,7)} = 13.69$	$p < .01$
Task	$F_{(1,7)} = 44.59$	$p < .0005$
Grip × Task	$F_{(1,7)} = 9.24$	$p < .02$
FEMALES **Factor**	**F statistic**	**Significance**
Grip	$F_{(1,7)} = 29.93$	$p < .001$
Task	$F_{(1,7)} = 47.41$	$p < .0005$
Grip × Task	$F_{(1,7)} = 24.79$	$p < .002$

Figure 7.17 Results of separate analyses for males and females.

7.8 Discussion

On the whole, the experimental results strongly supported the experimental hypotheses as well as the high-level hypothesis that Guiard's Kinematic Chain model can be used to reason about bimanual performance for skilled 3D manipulative tasks. Reviewing this evidence:

> *H1: The Hard task is asymmetric and the hands are not interchangable.* This hypothesis was supported by the overall Grip effect and the Preferred Hard vs. Reversed Hard contrast, both of which were highly significant. This suggests that maniplation is most natural when the right hand works relative to the left hand.

There are several qualities of the experimental task which may have led to the lateral asymmetry effects:

- *Mass asymmetry*: When holding the tool, some subjects had visible motor tremors in the left hand; but when they held the target object, the greater mass helped to damp out this instability.

- *Having the right view of the objects:* As mentioned previously, in the Reversed condition, some subjects tried to hold the tool at a fixed orientation in the left hand and move the target object to the tool. But as the subject moved the target object, he or she would no longer have the best view to see the target, and performance would suffer.

- *Referential task:* The task itself is easiest to perform when the manipulation of one object can done relative to a stationary object held in the other hand.

Under virtual manipulation, one can overcome some of these factors (such as mass asymmetry), but not all of them. For example, many virtual manipulation tasks (such as the neurosurgeon's task of cross-sectioning volumetric medical image data) will require a specific view to do the work and will have a referential nature.

H2: For the Easy task reversing roles of the hands will not have any reliable effect. The Grip effect was much smaller for the Easy task, but was significant at the $p < 0.10$ level, so one cannot confidently conclude there was *no* Grip effect. Nonetheless, for practical purposes, lateral asymmetry effects are much less important here.

H3: The importance of specialization of the roles of the hands increases as the task becomes more difficult. The predicted Grip × Task interaction was highly significant, offering strong evidence in favor of H3.

H4: Haptics fundamentally change the type of motor control required. Taken together, the experimental evidence for H1-H3 further suggests that the motor control required for the Easy conditions, where there was plentiful haptic feedback in the form of physical constraints, fundamentally differed from the Hard conditions.

The evidence in support of this final hypothesis underscores the performance advantages that are possible when there is haptic feedback to guide the task. Subjects devoted little cognitive effort to perform the Easy task, whereas the Hard task required concentration and vigilance.

This suggests that passive haptic feedback from supporting surfaces or active haptic feedback from devices such as the Phantom [147] can have a crucial impact for some tasks. This also underscores the difficulty of using a glove to grasp a virtual tool: when there is no physical contact, the task becomes a hand-eye coordination challenge, requiring full visual attention. With haptic feedback, it can be an automatic, subconscious manipulation, meaning that full visual attention can be devoted to a high-level task (such as monitoring an animation) instead of to the “tool acquisition” sub-task. For example, the PolyShop VR application [1] (*fig. 2.5 on page 18*) uses a physical drafting table as a support surface for 2D interactions, allowing the user to just touch the table to make selections.

These issues underscore the design tension between physical and virtual manipulation. The design challenge is find ways that real and virtual objects can be mixed to produce something better than either can achieve alone.

> *"...many people think about input only at the device level, as a means of obtaining improved time-motion efficiency through use of the sensorimotor system... this is a big mistake. Effectively structuring the pragmatics of input can also have a significant impact on the cognitive level of the interface."*
>
> *Bill Buxton [29]*

Chapter 8

The Bimanual Frame-of-Reference

8.1 Overview

This experiment investigates the synergy of the two hands for virtual object manipulation. The results suggest that the two hands together provide sufficient perceptual cues to form a frame of reference which is independent of visual feedback. The same is not true for one hand moving in empty space. My interpretation is that users may not have to constantly maintain visual attention when both hands can be involved in a manipulation.

The data suggest a transfer of skill from this experiment's bimanual condition to the unimanual condition. Subjects who performed the bimanual condition first seemed to learn a more effective task strategy for the unimanual condition. This further suggests that involving both hands in the physical articulation of a task can influence cognitive aspects of performance, in terms of the task strategy used.

8.2 Introduction

The central hypothesis of this study is that the combined action of the two hands can play a vital role for the interactive manipulation of virtual objects. Certainly, there is much informal evidence to support this position. Most everyday manipulative tasks involve both hands: for example, striking a match; unscrewing a jar; sweeping with a broom; writing on a piece of paper; dealing cards; threading a needle; or painting on a canvas, where the preferred hand holds the paintbrush, the nonpreferred hand the palette.

A few virtual reality applications and demonstrations, both on the desktop [141][149][80] and in fully immersive situations [121][119][164], have recognized the design possibilities for the two hands, and for years Buxton [27][95] has argued that one can improve both the naturalness and degree of manipulation of interfaces by employing both hands. Yet, beyond a possibly improved efficiency of hand motion, there has been little formal evidence of precisely what advantages, if any, the two hands can bring to virtual manipulation.

Moreover, just because a behavioral strategy is exhibited in the real world, this does not necessarily mean that it will be useful in a virtual environment. Virtual environments offer opportunities to violate the limitations of physical reality, and one only needs to mimic those qualities of physical reality which facilitate skill transfer or which form essential perceptual cues for the human participant to perform his tasks.

To establish the utility of the two hands in virtual environments, an experiment needs to formally demonstrate what users can do with two hands that they can’t easily do with one, and address some questions of when, and why, a bimanual interface might offer some advantages. I do not claim to answer all of these questions, but the current study offers some data which suggests areas where involving both hands may have some advantages.

The experiment suggests that the two hands together form a hand-relative-to-hand frame of reference. A frame of reference is a centered and oriented perceptual coordinate system which is specified by a center point plus three directional axes. An interesting property of the bimanual frame of reference is that the information can be encoded by the hands themselves, and as such does not necessarily rely on visual feedback. As an intuitive example, it is easy to touch your index fingers behind your head, but this action is clearly not guided by vision.

8.3 Hypotheses

This study investigates the following specific hypotheses:

H1. The two hands together provide sufficient perceptual cues to form a frame of reference which is independent of visual feedback.

H2. When using just one hand, subjects can employ other body-relative cues (such as sense of joint angles or sense of torso midline) to make an unbiased estimate of a remembered hand position, but these cues are less precise. Thus, unimanual control is more dependent on visual feedback.

H3. The physical articulation of a task can influence cognitive aspects of performance, in terms of the task strategy used. Using two hands together encourages exploration of the task solution space, and this will allow subjects to get a better sense of what a good strategy is for the experimental task.

8.4 Cognitive aspects of performance

Hypothesis 3, which asserts that using two hands can influence cognitive aspects of performance, has previously been articulated by Buxton. Working with his Input Research Group, Leganchuk [108] has provided some preliminary evidence which suggests that “representation of the task in the bimanual case reduces cognitive load.”

Leganchuk’s experiment (*fig. 2.18 on page 38*) studied an “area sweeping” task in which subjects selected an area encompassing a target. This is similar to sweeping out a rectangle to select a set of targets in a graphics editing application. Using both hands allowed subjects to complete the task significantly faster than using just one hand.

Furthermore, the difference in times could not be attributed to the increased time-motion efficiency alone. This was interpreted as evidence that the bimanual technique "reduces cognitive load."

Another way to investigate the hypothesis that bimanual control can influence cognitive aspects of performance is to take direct measures of cognition, such as quantifiable metrics of learning, memory, or transfer of skill. Leganchuk's strategy of taking differences between one and two-handed techniques relies on the assumption that differences beyond those clearly accounted for by increased time-motion efficiency can be attributed to differences in cognitive load. But if one can demonstrate a direct metric of cognition, this assumption does not have to be introduced.

8.5 The Experiment

8.5.1 Subjects

Seventeen unpaid subjects (13 female, 4 male) were recruited from the University of Virginia psychology department's subject pool. One subject (male) was left-handed. No subjects had experience with 3D input devices or two-handed computer interfaces.

8.5.2 Task

The task consisted of two phases. In the primary phase, users attempted to align two virtual objects. The purpose of this phase was to engage the user in an initial task which would require moving and placing the hand(s) in the environment. The second phase consisted of a "memory test" where users tried to reproduce the placement of their dominant hand without any visual feedback.

I used the neurosurgical interface props as input devices. In the current experiment, the doll's head controls the orientation and depth of a target object. The target object is an extruded triangular shape (*fig. 8.1, left*). The plate tool controls the position and orientation of a blue semi-transparent rectangular plane on the screen.

For the primary task, users were instructed to align and intersect the triangle and the plane so that they were coplanar (*fig. 8.1, right*). The triangle would highlight in yellow when the plane was aligned with it. The plane was considered to be aligned with the triangle if the plane was within 13 millimeters of all three corners of the triangle. Each edge of the triangle was 50 millimeters long. The triangle appeared at a new initial random orientation for each trial. The "stage" seen in the background of figure 8.1 served only as a simple perceptual aid and never moved.

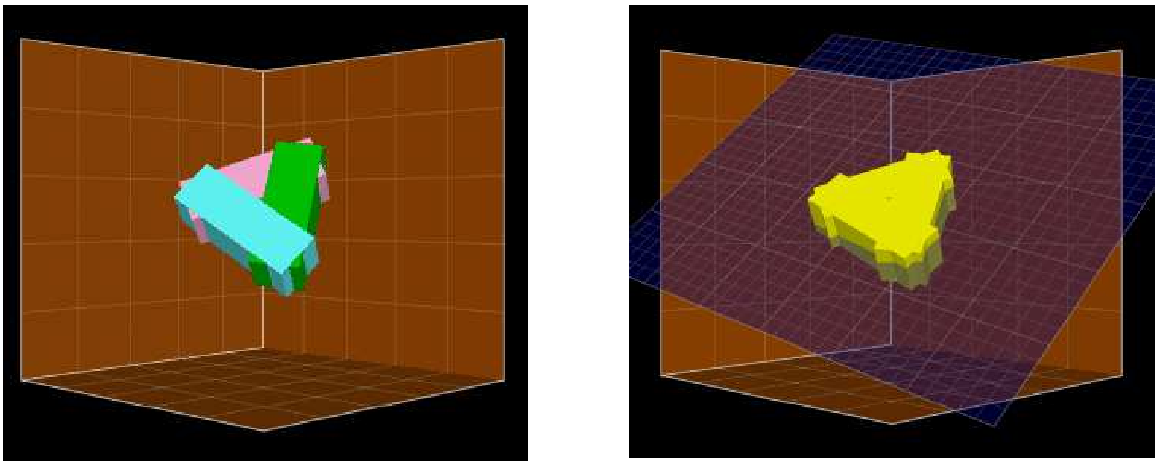
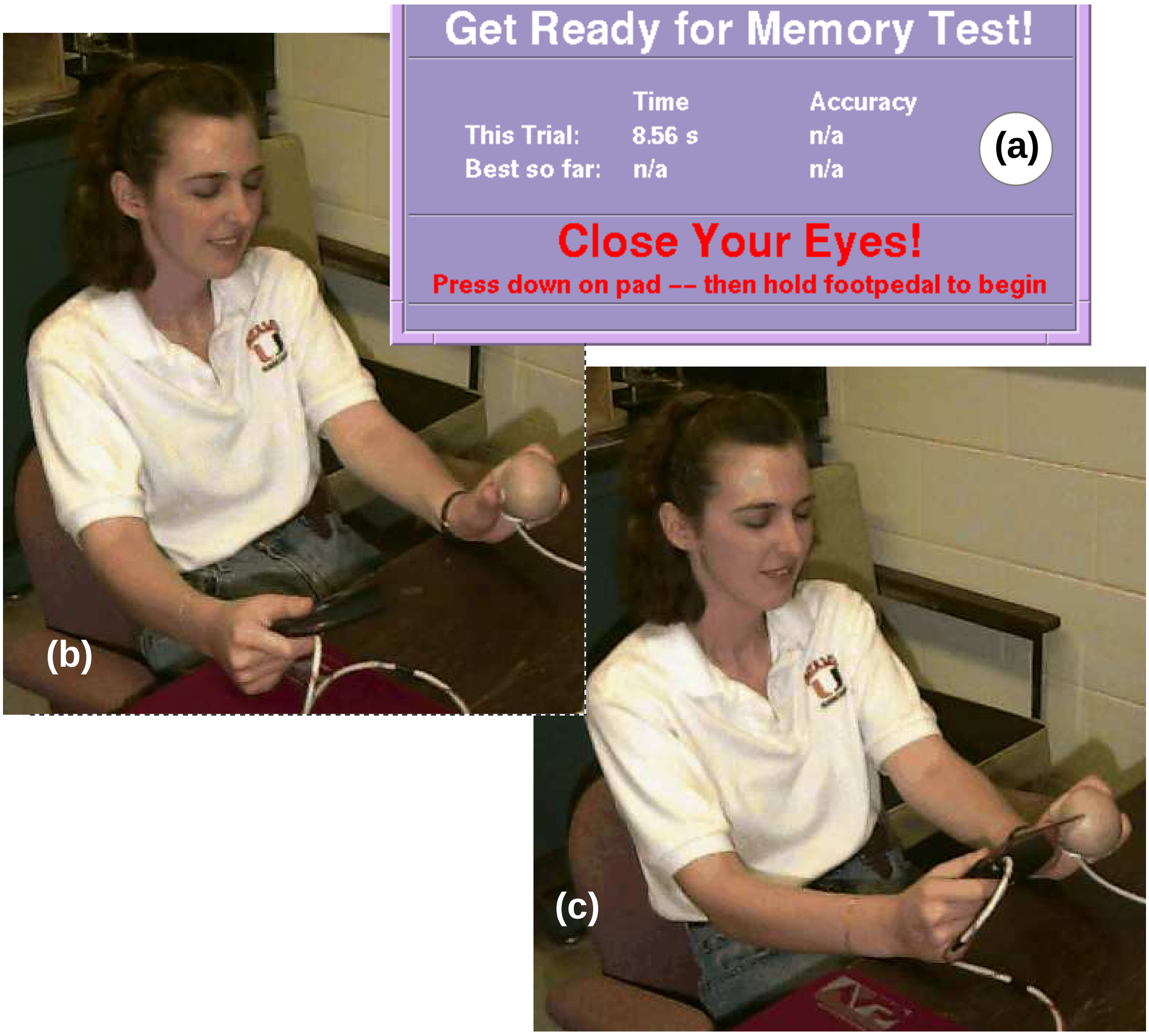


Figure 8.1 Stimuli for the primary task.

A footpedal was used as a clutch for the plate tool. When subjects held the pedal down, the plane could move freely relative to the target object. When the pedal was released, the plane would stay embedded in the target object. If the two were aligned when the pedal was released, this ended the primary task.



Get Ready for Memory Test!

	Time	Accuracy
This Trial:	8.56 s	n/a
Best so far:	n/a	n/a

Close Your Eyes!

Press down on pad –– then hold footpedal to begin

Figure 8.2 The memory test.

At the end of the primary task, the computer recorded the position and orientation of the preferred hand (which was always holding the plate tool). A dialog then appeared telling the subject to “Get Ready for Memory Test!” (*fig. 8.2, a*). For the memory test, subjects were instructed to put their preferred hand down on a mouse pad at the side of the work space, close their eyes (*fig. 8.2, b*), and then to attempt to exactly reproduce the position and orientation of the plate tool without any visual feedback (*fig. 8.2, c*). At the end of each trial, the computer displayed the subject’s best accuracy so far.

8.5.3 Experimental conditions

The experiment compared two conditions, a bimanual condition and a unimanual condition. In the bimanual condition, simultaneous motion of both input devices was possible. The doll’s head was held in the nonpreferred hand, the plate in the preferred hand. Since the distance between the two objects does not affect the alignment required for the primary task, it is always possible for the doll’s head and the plate tool to make physical contact when completing the task. Subjects were instructed to use this technique in the bimanual condition, since the purpose of the experiment was to test how well subjects could use the nonpreferred hand as a reference. The bimanual condition is shown in figure 8.2.

In the unimanual condition, subjects were instructed to always keep their nonpreferred hand in their lap. Subjects were only allowed to grasp one device at a time, using only their preferred hand. There was a definite device acquisition time required when switching input devices, but subjects were instructed that time to complete the task was not important -- only their accuracy on the memory test mattered. For the memory test, the unimanual condition was identical to the bimanual condition, except that the nonpreferred hand was no longer available as a reference.

Clearly, both conditions utilized a space-multiplexed design, with a separate input device for each function, as opposed to a time-multiplexed design, where a single device controls multiple functions by changing modes. Brooks [19] reports that overloading a device with multiple functions can often cause confusion. Thus, I chose a space-multiplexed design for the unimanual condition because I did not want the possible issue of users becoming confused over which “mode” they were in to interfere with the experiment itself.

For the unimanual condition only, a second clutch footpedal was needed to allow subjects to rotate the doll’s head and leave it “parked” at a particular orientation, thus allowing them to put down the doll’s head and pick up the plate tool. Users had no difficulty in using the two pedals: there were no experimental trials where a user clicked the wrong pedal in the unimanual condition.

Originally, I had planned to use two footpedals in the bimanual condition as well, but pilot studies suggested this was problematic. If the footpedal is used to “park” the target object in the bimanual condition, the user is again moving relative to the environment, not relative to the reference frame specified by the nonpreferred hand. Once pilot subjects developed some experience with the task, they would essentially always hold down the second footpedal to maintain the doll’s head as a reference. Thus in the bimanual case the second footpedal seemed to introduce confusion without adding any new or helpful capabilities.

8.5.4 Experimental procedure and design

A within-subjects latin square design was used. Eight subjects (six female, two male) performed the unimanual condition first and nine subjects (seven female, two male) performed the bimanual condition first. Subjects performed 12 experimental trials for each condition.

During a practice session subjects were introduced to the equipment[1] and allowed to become familiar with it. I gradually introduced each element of the experimental procedure and made sure that subjects could perform the task before moving on. Practice sessions lasted 10-20 minutes (prior to the first experimental condition) and 5-10 minutes (prior to the second experimental condition). There was not a fixed number of trials or set time limit for practice, but rather each subject practiced until he or she felt completely comfortable with the equipment and experimental procedure.

---

[1] All experimental data was collected using a Polhemus FASTRAK [137] six degree-of-freedom magnetic tracking system.

8.6 Results

Accuracy on the memory test was the only dependent measure. Accuracy was measured in terms of angle (shortest-arc rotation to align the remembered reference frame with the ideal reference frame) and distance (the translational offset between the reference frames). Distance was also logged as single-axis offsets in the reference frame of the plate tool (offsets along the left-right axis, front-back axis, and up-down axis). The left-right axis is the primary axis along which the hands made contact.

Accuracy metric	Bimanual	Unimanual
Angle (degrees)	11.2	10.4
Distance (mm)	16.6	32.6
Left-right distance (mm)	3.7	17.5
Up-down distance (mm)	8.3	14.2
Front-back distance (mm)	11.5	16.6

Figure 8.3 Overall means obtained in each experimental condition.

Figure 8.3 reports the overall means obtained in each experimental condition. An analysis of variance with repeated measures was performed on the within-subjects factor of Condition (unimanual vs. bimanual). Condition was not a significant factor for angle but was highly significant for measures of distance (*fig. 8.4*).

Accuracy metric	F statistic	Significance
Angle	$F_{(1,16)} = 0.74$	$p > .40$, n.s.
Distance	$F_{(1,16)} = 44.21$	$p < .0001$
Left-right distance	$F_{(1,16)} = 51.95$	$p < .0001$
Up-down distance	$F_{(1,16)} = 10.22$	$p < .006$
Front-back distance	$F_{(1,16)} = 12.04$	$p < .0035$

Figure 8.4 Significance levels for main effects.

This evidence strongly supports hypothesis H1, suggesting that subjects were able to utilize the perceptual cues provided by the nonpreferred hand to reproduce their six degree-of-freedom posture independent of visual feedback. Subjects were significantly more accurate with both hands than with just one, supporting H2.

The analysis also revealed a significant Condition × Order interaction for Distance ($F_{(1,15)} = 9.09$, $p < .01$). It is often assumed that alternating the order of two conditions across subjects automatically controls for order effects caused by transfer of skill. But this is not true if there is a one way (or asymmetric) transfer of skill between the two conditions. A Condition × Order interaction is the statistical evidence for such an asymmetrical transfer effect [139].

The means grouped by order (*fig. 8.5*) show this effect. When performing unimanual first, subjects' distance was 5% better on the subsequent bimanual condition than those subjects who completed bimanual first. But when performing bimanual first, subjects performed 28% better on the subsequent unimanual condition than those subjects who completed unimanual first. My interpretation is that subjects learned a more effective task strategy in the bimanual condition, and were able to transfer some of this skill to the unimanual condition.

My qualitative observations also support this position. When performing the unimanual condition first, subjects had a tendency to avoid using the doll's head: only 2 out of 8 of these subjects consistently reoriented the target object with the doll's head. Subjects would instead adapt the plate tool to the initial (randomly generated) orientation of the target object. But for 8 out of the 9 subjects who tried the bimanual condition first, during the unimanual condition they would re-orient the doll's head on essentially every trial. As one subject explained, during the bimanual condition she had learned that "instead of accepting what it gave me, I did better when I moved [the doll's head]."

All of this evidence supports H3, suggesting that bimanual control can affect performance at the cognitive level by influencing a subject’s task-solving strategy. To definitively demonstrate that there is a bimanual to unimanual transfer effect, future experimental work should compare the results obtained in this experiment to data from subjects who perform two blocks of the unimanual condition or two blocks of the bimanual condition.

	Bimanual First		**Unimanual First**	
Means by Order	**Biman**	**Uniman**	**Biman**	**Uniman**
Angle	11.5	9.8	11.0	11.0
Distance	17.2	27.6	16.0	38.2
Left-right distance	4.4	13.8	3.0	21.6
Up-down distance	7.8	12.1	8.9	16.6
Front-back distance	12.3	15.0	10.6	18.4

Figure 8.5 Means grouped by order of experimental conditions.

Finally, an analysis of signed distance errors supported H2: with just one hand, subjects could make unbiased estimates of a remembered hand position. The estimate is said to be “unbiased” because in the unimanual condition the means of the signed errors along each axis did not significantly differ from zero, although means along the up-down and front-back axis did approach significance (*fig. 8.6*).

The analysis shows a highly significant bias for the bimanual condition along the left-right axis (fig. *8.6*), and a bias nearing significance along the up-down axis. Although significant, the left-right bias effect is caused by an overall bias of less than 2 millimeters (*fig. 8.7*). This small bias effect is difficult to interpret, but for practical purposes, the small magnitude of the effect means there is essentially no bias in the bimanual condition either.

Accuracy metric	F statistic	Significance
Unimanual		
Left-right distance	$F_{(1,16)} = 0.08$	$p > .75$
Up-down distance	$F_{(1,16)} = 4.23$	$p < .06$
Front-back distance	$F_{(1,16)} = 2.83$	$p > .10$
Bimanual		
Left-right distance	$F_{(1,16)} = 12.16$	$p < .005$
Up-down distance	$F_{(1,16)} = 3.37$	$p < .10$
Front-back distance	$F_{(1,16)} = 2.22$	$p > .15$

Figure 8.6 Tests for bias in remembered hand position.

Accuracy metric	Mean (mm)	Std. Deviation
Unimanual		
Left-right distance	-1.1	16.6
Up-down distance	-5.8	11.6
Front-back distance	-4.6	11.2
Bimanual		
Left-right distance	-1.7	2.1
Up-down distance	-2.9	6.4
Front-back distance	+2.3	6.3

Figure 8.7 Means of signed distance errors.

8.7 Qualitative results

The unimanual condition seemed to impose two chief difficulties for users. First, when orienting the target object with the doll’s head, the subject had to think ahead to the next step to anticipate which orientations of the target would be easiest for the action of the plate tool. Second, since the unimanual condition requires movement relative to the environment, the user had to remember approximately where he or she had “parked” the doll’s head.

The bimanual condition avoids both of these difficulties. When using both hands, it is much easier to see what orientations of the target will be easy to select with the plane. As one subject commented, “it was easier to get them to come together, and faster too.” Another noted that “two hands have much more flexibility for how you solve the problem.” And since the doll’s head is always part of the interaction, subjects had no need to remember where they had “parked” it.

The bimanual condition does introduce the possibility of the two objects colliding. For example, if the plate is directly behind the doll’s head and the subject needs to move it forward, he or she cannot do this directly since the doll’s head is in the way. But the virtual plane is bigger than the physical plate, so one can solve this problem by (for example) holding the plate immediately to the right of the doll’s head. On the screen, one sees the blue plane intersecting the target object even though the two input devices don’t physically intersect. Two subjects initially found this to be confusing, but quickly adapted after they realized that one does not have to *physically* intersect the two objects[1].

I asked subjects about their task strategies and the cues they had used to perform the memory test. Subjects generally tried to hold as many of the variables constant as possible, and then memorized the rest. For example, subjects often kept the elbow and wrist angles fixed and would try to maintain an invariant hand posture with respect to the input device. Remaining variables such as the height and depth of the hand placement were then estimated from memory.

In the bimanual condition, subjects seemed to have an innate sense of where they had touched (either on the doll’s head or on their hand)-- as one subject explained, “the touch knew the position”-- and many subjects thought of the angle of the plane as a separate thing to memorize. Subjects certainly also made use of the physical landmarks on the doll’s

---

[1] I have occasionally seen this same problem in the context of the neurosurgery application.

head (such as the ears or the features of the face). Without these landmarks, subjects probably would have been less accurate, but many subjects seemed to zero in on the exact spot even before physical contact was made. As one subject commented, “even before you touch the spot you know it.”

In the Unimanual condition, the edge and surface of the desk served as a physical reference and most subjects tried to use this to their advantage, for example by resting their forearm against the desk and remembering the touch point. Subjects would often attempt to estimate the left-right placement of their hand using body-relative cues such as the torso midline or the positions of their legs. One subject commented that “because there was nothing to land on, your sort of lose your sense of balance.” Another described her hand as “just floating in space, but using both hands gave you something else to reference.”

8.8 Discussion

The experimental results have clear design implications for the role of visual feedback and attention in human-computer interaction. Users maintain a fairly precise, body-relative representation of space which does not depend on visual feedback. A relatively inaccurate environment-relative representation of the space is also maintained. My interpretation is that two hands and split attention go well together, opening up new possibilities for eyes-free interaction.

When using two hands, the user’s attention does not necessarily have to constantly monitor the manipulation itself, and attention can be directed towards a secondary task, such as watching an animation or a representation of the manipulation from a second viewpoint. Unimanual control is more dependent on visual feedback and can therefore impede the user’s ability to split attention between multiple tasks.

The Worlds-in-Miniature (WIM) interface metaphor [164] (*fig. 2.12 on page 28*) provides an example of these issues in action. The WIM provides the virtual reality user with a hand-held miniature representation of the immersive life-size world. Users manipulate the WIM with both hands, using the nonpreferred hand to orient the clipboard and the preferred hand to manipulate object on the WIM.

By using both hands, users can quickly develop a sense of the space represented by the WIM. Stoakley [164] reports that some users have manipulated objects in the WIM without even looking at it (that is, users have manipulated objects while holding the clipboard below the field-of-view seen in the immersive display). This is convenient for tasks such as hanging a picture, where it is useful to manipulate objects in the WIM while attending to the 1:1 scale view to check the picture’s alignment and to evaluate how it fits the architectural space.

This is perhaps the strongest qualitative evidence that interaction techniques based on hand-relative-to-hand manipulation can allow users to focus attention on their tasks without necessarily becoming distracted by the interface technology.

8.9 Conclusion and future work

This study has provided some initial evidence which helps to support the high-level hypothesis that using both hands can help users gain a better sense of the space they are working in [66]. Immersive VR systems which use just one hand often do not offer users any physical reference points. Making use of both hands provides a simple way to increase the degree of manipulation and to let the user’s own hand act as a physical reference. Another technique is to introduce a grounding object such as a drafting table; but even in this situation, using two hands plus the grounding object allows interesting design possibilities [1].

A second high-level hypothesis is that in some cases using both hands can change the way users think about a task. This experiment also provided some initial evidence in favor of this hypothesis, suggesting that it may be easier to explore alternative strategies for problem solving when both hands can be used. This was reinforced by the qualitative observation that subjects were more likely to “take what they were given” in the unimanual condition because they had difficulty anticipating which orientation of the target object would be easy to select with the plane.

The input devices used in this study were rich in tactile orientation cues and this probably helped subjects to perform the experimental task more precisely. If the experiment had used featureless spheres as input devices, for example, subjects probably would have had a less acute sense of the orientation of each device [85]. I also believe that allowing contact between the two hands was a factor in the experiment, but not the only factor. When using two hands, subjects could often come quite close to the original position even before contact was established. Further study is necessary to determine if this differs significantly from moving a single hand relative to the environment.

As a thought experiment, one can imagine using a single hand to move the plate tool relative to a doll’s head mounted on a pivot (or a similar mechanical assemblage, such a set of gimbal rings). This would be analogous to using one hand on a tablet fitted with a physical template, which works well [30]. But the current experimental data suggest that the dynamic role of the nonpreferred hand also led to a cognitive performance benefit in terms of task strategy chosen. The task syntax supported by moving one hand relative to a reference object on a pivot is quite similar to that required by this experiment’s unimanual condition. As such I speculate that using the pivot with just one hand would have some of the same limitations: users might have difficulty anticipating what orientation of the pivot object would be most facile for the action of the plate tool.

When using two hands, using a mechanical pivot (as opposed to the current free-moving doll’s head) might have advantages for some tasks. A jewelers or a technician’s workbench provides an example from the real world: a clamp or armature is often used to hold an object being worked on, so that both hands can perform detailed work on the object being held. An interesting issue for future research is to see if a mechanical apparatus is really necessary, or whether a virtual solution using some combination of constraint modes and clutching operations would be viable. I speculate that a combination of the two approaches could provide a powerful yet highly usable interface metaphor.

"... the scientific base of user technology is necessary in order to understand why interaction techniques are (or are not) successful, to help us invent new techniques, and to pave the way for machines that aid humans in performing significant intellectual tasks."

Stu Card [38]

Chapter 9

Conclusions

9.1 Approach

The WIMP (Windows, Icons, Menus, and Pointer) interface metaphor was designed to match the computing capabilities available in the late 1970's and early 1980's. Computers are no longer limited to small black-and-white displays and an impoverished input bandwidth. Continued vigorous growth in processor power and memory capacity, in addition to modern innovations such as large flat-screen color displays, real-time 3D graphics on the PC platform, and high degree-of-freedom input devices make the limitations of the WIMP interface paradigm increasingly apparent.

A host of devices, displays, and interaction techniques have been proposed as candidates for the post-WIMP interface. A key challenge is to discover interaction techniques that do not necessarily behave like the real world, yet nonetheless seem natural. Directly mimicking reality will not necessarily produce the desired results. A demonstration of a new interaction technique that does not directly mimic reality may not be enough to fully take advantage of the technique. To understand why interaction techniques do or do not work, and to design new interaction techniques which meet these criteria, the interface designer needs to understand the human. In essence, my approach focuses not on a particular technology, but rather on the capabilities of the human participant.

9.2 Contributions

My work contributes a system which has undergone extensive informal usability testing in the context of real domain experts doing real work, but it also presents the results of experimental evaluations which illustrate general behavioral principles. Some of the informal design issues which my user observations have suggested are now supported by formal experimental evaluations. And the selection of experimental evaluations has been tempered by design experience, so there is some assurance that the experiments apply to design in a meaningful way. Together, these approaches make a decisive statement that using both hands for virtual manipulation can result in improved user productivity.

The primary contributions of my work are:

- *Revisiting passive haptic issues:* Facile virtual manipulation requires studying the feel of the interface, and not just the look of the interface.

- *Application to neurosurgical visualization:* I contribute the design, implementation, and usability testing of a three-dimensional volume visualization application which has been well accepted by neurosurgeons.

- *Two-handed virtual manipulation techniques:* Using the nonpreferred hand as a dynamic frame of reference is an important design principle for virtual manipulation. The asymmetrical hierarchical organization of two-handed manipulation is not generally known by interface designers, nor will it be obvious how to apply this knowledge when they design two-handed interfaces without benefit of the knowledge contributed by this dissertation.

- *Basic knowledge about two hands:* I contribute experimental data which shows that performance is best when the preferred hand operates relative to the frame-of-reference of the nonpreferred hand.

- *Two hands are not just faster than one hand:* Two hands together provide the user with information which one hand alone cannot. Using both hands can influence how users *think* about a task.

- *Research methodology:* As a whole, the dissertation provides a case study for an interdisciplinary approach to the design and evaluation of new human-computer interaction techniques.

9.3 High level conclusions

This work has demonstrated the inherently interdisciplinary nature of human-computer interaction. The collaboration of neurosurgeons, computer scientists, and psychologists has enabled a decisive contribution to the virtual manipulation and two-handed interaction research fields. Without any one part of this collaboration, the work would not be as convincing.

This work also illustrates the synergy between behavioral principles, usually proposed by the psychology community, and design principles, sought after by the human-computer interaction community. Behavioral science has direct application to design: human-computer interaction can profit from increased collaboration with perceptual psychologists, cognitive psychologists, and motor behavior researchers.

9.4 Future Work

Multimodal input has been an underlying theme of this thesis. An exploration of multimodal input which integrates both hands, the voice (voice recognition), and possibly other effectors such as the head, legs, and feet is a natural extension of this work. It seems quite likely that Guiard’s kinematic chain model can be profitably applied to more than just the two hands working together. For example, an organist uses both hands on the organ keyboard, uses both feet to control pedals, and sometimes sings as well; during microsurgery, the surgeon holds an implement in each hand, issues verbal orders to the surgical team, and sometimes even uses the feet to steer the surgical microscope and to control the zoom factor. Can computer interfaces take advantage of these skills as well?

Much previous work has discussed all of these capabilities as if they were independent input channels, but as Guiard’s work suggests they are not independent at all, but rather form a hierarchy of interdependent effectors. And there are certainly other constraints which interface designers have only begun to understand: for example, Karl [96] has found that issuing voice commands can interfere with short term memory under some conditions. As a second example, this time from the psychology literature, Cremer [44] has reported that during a finger tapping task, concurrent speech activity can disrupt right-hand performance but not left-hand performance, while a concurrent visuospatial task can disrupt left hand performance. Clearly, there are some constraints on multimodal human performance which are not yet understood at the level of interface design principles.

Interfaces which combine both active force feedback and two-handed interaction, for applications such as telemanipulation [163] or exploration of computer-generated virtual spaces, seem like a promising research area. This might involve using an armature such as the Phantom [147] in one hand with a virtual object held in the other hand, or using a pair of armatures, one for each hand. For example, one possible application the user interface group [173] has envisioned is to add active force feedback to the Worlds-in-

Miniature application (*fig. 2.12 on page 28*) [164], so that the user can feel objects held in a miniature world in the nonpreferred hand. Such a configuration might also allow one to explore behavioral issues with two hands and active haptic feedback. For example, in a prototype of this configuration, George Williams (working with the user interface group [173]) found that when the preferred hand holds a force reflecting stylus up to a virtual object held by the passive nonpreferred hand, there is a compelling “ghost sensation” that the nonpreferred hand is also receiving force feedback.

Large flat-panel display technology is rapidly coming to market. As scale increases, the need arises for tasks with both macrometric and micrometric movement scales, which would seem to make such a configuration ideal for two-handed interaction. Large, flat displays can be simulated now using back-projection techniques (for example, the ActiveDesk (*fig. 7.4 on page 139*), the Immersadesk (*fig. 2.14 on page 30*) [49], or the Reactive Workbench [52]) and researchers have already started to explore two-handed interaction techniques for these systems.

In medical imaging in general, there are still tremendous opportunities to improve physician user interfaces. A current trend is the “filmless hospital” which replaces the traditional physical film media with purely digital patient images, and has led to the introduction of so-called PACS systems (Picture Archiving and Communications Systems). Working with digital images raises a host of interface, security, archiving, and telecommunication issues which never arose with film-based imaging, and as a result, physicians find many of the commercially available PACS systems very difficult to use.

In terms of the neurosurgery application, there is still much to be learned by turning the props-based interface into a clinical tool. Multimedia Medical Systems [122] is currently exploring the issues of turning my interface design into a commercial application, for neurosurgery as well as other medical applications. By deploying a commercial version of the system at a significant number of hospitals, there is the potential for this effort to provide a real success story for spatial interaction technologies.

9.5 Epilogue

For me, perhaps the most striking element of all of this work is how much a resourceful design can get away with, in terms of helping the user to see things that aren't really there, while using some very simple technology. My interface design is perhaps more notable for the technologies it does *not* use than for those it does use: it does not use head tracking; it does not use stereoscopic projection; it does not use gloves; it does not use an immersive head-mounted display. Nonetheless, it allows the user to see and explore the interior of a virtual object in interesting new ways. The point I am trying to make is that most applications of interactive 3D graphics do not need a full scale virtual reality implementation, but often some rather simple technology in the desk top setting can do the trick.

The technology trap is an easy one to fall into. Fred Brooks has recounted the story of the world's first molecular visualization system in virtual reality [22], which he and his colleagues designed to allow a molecular chemist to walk around a room-filling molecule and inspect it from different standpoints. What is one of the first things Brooks's molecular chemist collaborator asked for? A chair!

Indeed, my chief neurosurgeon collaborator had initially talked to me about doing an immersive walkthrough of the human brain. It probably would have failed for many of the same reasons that Brooks's room-filling molecule was not completely satisfactory for molecular visualization. This teaches another valuable lesson: domain experts are expert primarily at the work they do, and not at the technology that might best help them to do that work better. The interface I designed uses six degree-of-freedom input devices and two-handed manipulation not because these are interaction techniques that I or my neurosurgeon collaborators had *a priori* chosen to investigate, but rather because they are well suited to some of the problems that neurosurgeons are trying to solve.

“A man will turn over half a library to make one book.”

Samuel Johnson

quoted in Boswell’s ***Life of Johnson****, Chap. viii, 1775.*

Bibliography

1. Abel, K., Alic, M., Moshell, J. M., “Rapid Construction of Model Building for Urban Combat Mission Rehearsal Databases (PolyShop),” URL: http://www.vsl.ist.ucf.edu/~polyshop/polyshop.html.

2. Adelstein, B., Johnston, E., Ellis, S., “A Testbed for Characterizing Dynamic Response of Virtual Environment Spatial Sensors,” UIST’92, 15-22.

3. Annett, J., Annett, M., Hudson, P., Turner, A., “The Control of Movement in the Preferred and Non-Preferred Hands,” Quarterly Journal of Experimental Psychology, 31, 641-652.

4. Arvo, J., “Random Rotation Matrices.” in ***Graphics Gems II***, Arvo, J., ed., Academic Press, pp. 355-356.

5. Ascension Technology Corporation, PO Box 527, Burlington, VT 05402. (802) 655-7879.

6. Athenes, S., “Contribution de la main droite a l’acte graphique chez le scripteur gaucher adulte: une comparaison entre postures ‘inversee’ et ‘non-inversee” (*Contribution of the right hand in handwriting for left-handed adults: a comparion of the ‘inverted’ and ‘non-inverted’ postures*). Universite de Provence, France. Yves Guiard, advisor, June 1983.

7. Athenes, S., “Adaptabilite et developpement de la posture manuelle dans l’ecriture: etude comparative du droitier et du gaucher” (*Adaptability and development of manual posture in handwriting: a study comparing right-handers and left-handers*). Universite D’Aix-Marseille II, France. Yves Guiard, advisor, June 1984.

8. Badler, N., Manoochehri, K., Baraff, D. “Multi-Dimensional Input Techniques and Articulated Figure Positioning by Multiple Constraints,” ACM Workshop on Interactive 3D Graphics, 1986, pp. 151-170.

9. Baecker, R., Grudin, J., Buxton, W., Greenberg, S., eds., “The Emergence of Graphical User Interfaces,” in ***Readings in Human-Computer Interaction: Toward the Year 2000***, Morgan Kaufmann Publishers, 1995, pp. 49-70.

10. T. Baudel, M. Beaudouin-Lafon, “Charade: Remote Control of Objects Using Hand Gestures,” Communications of the ACM, 36 (7), 1993, 28-35.

11. Bergman, L., Fuchs, H., Grant, E., “Image Rendering by Adaptive Refinement,” Computer Graphics, 20 (4), 1986, pp. 29-37.

12. Berry, W. D., ***Understanding Regression Assumptions***. Quantitative Applications in the Social Sciences, Vol. 92, Sage Publications: Newbury Park, CA, 1993.

13. Bier, E. A., Stone, M. C., “Snap-Dragging,” Computer Graphics, 20 (4), 1986, pp. 233-240.

14. Bier, E. A., “Snap-Dragging In Three Dimensions,” Proc. 1990 Symposium on Interactive 3D Graphics, Computer Graphics, 24 (2), pp. 193-204.

15. Bier, E., Stone, M., Pier, K., Buxton, W., DeRose, T., “Toolglass and Magic Lenses: The See-Through Interface,” SIGGRAPH ‘93, pp. 73-80.

16. Bolt, R., “Put-That-There: Voice and Gesture at the Graphics Interface,” SIGGRAPH ‘80, 262-70.

17. Bolt, R. A., Herranz, E., “Two-Handed Gesture in Multi-Modal Natural Dialog,” UIST ’92, pp. 7-13.

18. Brooks, F. P. Jr., “Walkthrough--a Dynamic Graphics System for Simulating Virtual Buildings,” Proc. ACM Workshop on Interactive 3D Graphics, 1986, pp. 9-21.

19. Brooks, Jr., F. P., “Grasping reality through illusion: interactive graphics serving science,” ACM CHI’88 Conference on Human Factors in Computing Systems, pp. 1-11.

20. Brooks, Jr., F.P. “The Computer Scientist as Toolsmith II.” Keynote/Newell Award address at SIGGRAPH 94, Orlando, July 1994. Communications of the ACM, March 1996, Vol. 39, No. 3, pp.61-68. Also Computer Graphics, Vol. 28, No. 4 (November 1994), pp. 281-287.

21. Brooks, F., Ouh-Young, M., Batter, J., Kilpatrick, P., “Project GROPE--Haptic Displays for Scien-tific Visualization,” Comp. Graph. 24 (4), 1990.

22. Brooks, F., “3-D User Interfaces: When Results Matter,” Keynote speech at UIST’96, November 7th, 1996.

23. Brutzman, D., Pesce, M., Bell, G., van Dam, A., AbiEzzi, S., “VRML: Prelude and Future,” SIGGRAPH’96, Panel Discussion, pp. 489-490.

24. Bryson, S., Levit, C., “The Virtual Wind Tunnel,” IEEE Computer Graphics and Applications, July 1992, pp. 25-34.

25. Bryson, S, “Virtual Reality in Scientific Visualization,” Communications of the ACM, May 1996, 39 (5), pp. 62-71.

26. Butterworth, J., Davidson, A., Hench, S., Olano, T. M., "3DM: A Three Dimensional Modeler Using a Head-mounted Display," Proc. 1992 Symp. on Interactive 3D Graphics, pp. 135-138.

27. Buxton, W., Myers, B., "A Study in Two-Handed Input," CHI'86, pp. 321-326.

28. Buxton, W., "Chunking and Phrasing and the Design of Human-Computer Dialogues," Information Processing '86, Proc. of the IFIP 10th World Computer Congress, pp. 475-480.

29. Buxton, W., "Touch, Gesture, and Marking," in ***Readings in Human-Computer Interaction: Toward the Year 2000***, Morgan Kaufmann Publishers, 1995.

30. Buxton, W., Hill, R., Rowley, P., "Issues and Techniques in Touch-Sensitive Tablet Input," Computer Graphics, July 1985, pp. 215-224.

31. Buxton, W, ***The Pragmatics of Haptic Input***. ACM CHI'90 Tutorial 26 Notes.

32. Buxton, W., "There's More to Interaction Than Meets the Eye," Norman, D. and Draper, S., eds., ***User Centered System Design: New Perspectives on Human-Computer Interaction***, Hillsdale, N.J., 1986, pp. 319-337.

33. Buxton, W., "Human skills in interface design," L. MacDonald and J. Vince, eds., ***Interacting With Virtual Environments***, New York, Wiley, pp. 1-12.

34. Buxton, W., "Lexical and Pragmatic Considerations of Input Structure," Computer Graphics, 17 (1), 1983, pp. 31-37.

35. Buxton, W., Fiume, E., Hill, R., Lee, A., Woo, C., "Continuous hand-gesture driven input," Proceedings of Graphics Interface '83, pp. 191-195.

36. Card, S., Mackinlay, J., Robertson, G., "The Design Space of Input Devices," CHI'89, 117-124.

37. Card, S., Robertson, G., Mackinlay, J., "The Information Visualizer, an Information Workspace," CHI'91, pp. 181-187.

38. Card, S., Moran, T., "User Technology: From Pointing to Pondering," in ***Readings in Human-Computer Interaction: Toward the Year 2000***, R. Baecker, J. Grudin, W. Buxton, & S. Greenberg, eds., pp. 587-602.

39. Card, S., English, W., Burr, B., "Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT," Ergonomics, 21, pp. 601-613.

40. M. Chen, S. J. Mountford, A. Sellen, "A Study in Interactive 3-D Rotation Using 2-D Control Devices," Computer Graphics, 22, 4, August 1988, 121-129.

41. Chung, J. C., "A comparison of Head-tracked and Non-head-tracked Steering Modes in the Targeting of Radiotherapy Treatment Beams," Proc. 1992 Symp. on Interactive 3D Graphics, 193-196.

42. Cohen, P., Sullivan, J., "Synergistic Use of Direct Manipulation and Natural Language," CHI'89, pp. 227-233.

43. Conner, D., Snibbe, S., Herndon, K., Robbins, D., Zeleznik, R., van Dam, A., “Three-Dimensional Widgets,” Proc. 1992 Symposium on Interactive 3D Graphics, pp. 183-188, 230-231.

44. Cremer, M., Ashton, R., “Motor performance and concurrent cognitive tasks”, J. Motor Behavior, 13, pp. 187-196, 1981

45. Cruz-Neira, C., Sandin, D., DeFanti, T., “Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE,” SIGGRAPH ‘93, pp. 135-142.

46. Deering, M., “High resolution virtual reality,” Computer Graphics, July 1992, 26 (2), pp. 195-202.

47. Digital Image Design Incorporated, Cricket input device, http://www.didi.com/www/cricket.shtml

48. Drake, J., J. Rutka, J. Hoffman, “The ISG Viewing Wand,” Neurosurgery, 34, 1994, pp. 1094-1097.

49. Electronic Visualization Laboratory at the University of Illinois at Chicago. Creators of the Immersadesk, demonstrated at SIGGRAPH in 1994. Information available at URL: http://evlweb.eecs.uic.edu/EVL/VR/systems.html as well as URL: http://www.eecs.uic.edu/~mczernus/Immersadesk/immersadesk.html

50. Elographics, 105 Randolph Road, Oak Ridge, TN 37830. (615) 482-4100.

51. Evans, K., Tanner, P., Wein, M., “Tablet-based Valuators that Provide One, Two, or Three Degrees of Freedom,” Computer Graphics, 15 (3), August 1981, pp. 91-97.

52. Fakespace, Inc. URL: http://www.fakespace.com/.

53. Feiner, S., MacIntyre, B., Haupt, M., Solomon, E., “Windows on the World: 2D Windows for 3D Augmented Reality,” Proc. ACM Symposium on User Interface Software and Technology, 1993, pp. 145-155.

54. Feiner, S., Macintyre, B., Seligmann, D., “Knowlege-Based Augmented Reality,” Comm. of the ACM, 36 (7), 1993, pp. 53-61.

55. Feiner, S., Shamash, A., “Hybrid User Interfaces: Breeding Virtually Bigger Interfaces for Physically Smaller Computers,” UIST ‘91, pp. 9-17.

56. S.S. Fisher, M. McGreevy, J. Humphries, W. Robinett, “Virtual Interface Environment for Telepresence Applications,” Oct. 1988, Proc. Human Factors Society 32nd Annual Meeting.

57. Fitzkee, D., ***Magic By Misdirection***, Lee Jacobs Productions, P. O. Box 362, Pomeroy, OH, 45769-0362, 1975.

58. Fitzmaurice, G. W., “Situated Information Spaces and Spatially Aware Palmtop Computers,” Comm. of the ACM, 36 (7), 1993, pp. 39-49.

59. Fitzmaurice, G., Ishii, H., Buxton, W., “Bricks: Laying the Foundations for Graspable User Interfaces,” Proc. CHI’95, pp. 442-449.

60. Foley, J. D., Wallace, V., Chan, P., “The Human Factors of Computer Graphics Interaction Techniques,” IEEE CG&A, Nov. 1984, pp. 13-48.

61. Galyean, T. A., Hughes, J. F., “Sculpting: An Interactive Volumetric Modeling Technique,” Computer Graphics, 25 (4), pp. 267-274.

62. Gibson, J., ***The Ecological Approach to Visual Perception***. Lawrence Erlbaum, Hillsdale, NJ.

63. Gibson, J., “Observations on Active Touch,” Psychological Review, 69 (6), November 1962, pp. 477-491.

64. Gleicher, M., “Supporting Numerical Computations in Interactive Contexts,” Graphics Interface ‘93.

65. Goble, J., Hinckley, K., Snell, J., Pausch, R., Kassell, N., “Two-handed Spatial Interface Tools for Neurosurgical Planning,” *IEEE Computer,* special issue on virtual reality (to appear), 1995.

66. Goble, J., J. Snell, K. Hinckley, N. Kassell, “A real-time system for 3D neurosurgical planning,” *Visualization in Biomedical Computing 1994,* Robb, R. A., ed., Proc. SPIE 2359, October 1994, pp. 552-573.

67. Guiard, Y., “Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model,” The Journal of Motor Behavior, 19 (4), 1987, pp. 486-517.

68. Guiard, Y., “Failure to Sing the Left-Hand Part of the Score during Piano Performance,” J. Music Perception, 6 (3), 1989, pp. 299-314.

69. Guiard, Y., T. Ferrand, T., “Asymmetry in Bimanual Skills,” in ***Manual asymmetries in motor performance***, Elliott & Roy, eds. Boca Raton, FL: CRC Press, 1995.

70. Guiard, Y., “The Kinematic Chain as a Model for Human Asymmetrical Bimanual Cooperation,” in ***Cognition and Action in Skilled Behavior***, A. Colley & J. Beech, eds., Amsterdam: North-Holland, 1988, pp. 205-228.

71. Guiard, Y., “The Distal-to-Proximal Increase of Link Length along the Human Kinematic Chain: An Exponential Progression of Workspace Extension,” Annales des Sciences Naturalles, 1994, in press.

72. Guiard, Y., Millerat, F., “Writing Postures in Left-handers: Inverters are Hand-crossers,” Neuropsychologia, 22 (4), pp. 535-538, !984.

73. Guiard, Y., Diaz, G., Beaubaton, D., “Left-hand Adgantage in Right-handers for Spatial Constant Error: Preliminary Evidence in a Unimanual Ballistic Aimed Movement,” Neuropsychologia, 21 (1), pp. 111-115, 1983.

74. Halpern, D., ***Sex differences in cognitive ability***. Lawrence Erlbaum Associates, 1992.

75. Harre, R., “J. J Gibson: The Mechanism of Perception”, in ***Great Scientific Experiments: 20 Experiments that Changed our View of the World***, Phaidon Press Ltd, Oxford, England, 1981, pp. 147-153.

76. Hauptmann, A., “Speech and Gestures for Graphic Image Manipulation,” Proc. CHI’89, pp. 241-245.

77. Herndon, K., Zeleznik, R., Robbins, D., Conner, B., Snibbe, S., van Dam, A., “Interactive Shadows,” Proc. ACM Symposium on User Interface Software and Technology, 1992, pp. 1-6.

78. Herndon, K., Meyer, T., “3D Widgets for Exploratory Scientific Visualization,” Proc. ACM Symposium on User Interface Software and Technology, 1994, pp. 69-70.

79. Hewlett-Packard Company, 3000 Hanover St., Palo Alto CA 94304.

80. Hinckley, K., R. Pausch, J. Goble, N. Kassell, “Passive real-world interface props for neurosurgical visualization,” ACM CHI’94 Conference on Human Factors in Computing Systems, 1994, pp. 452-458.

81. Hinckley, K., Pausch, R., Goble, J. C., Kassell, N. F., “A Survey of Design Issues in Spatial Input,” Proc. ACM Symposium on User Interface Software and Technology, 1994, pp. 213-222.

82. Hinckley, K., Pausch, R., Goble, J. C., Kassell, N. F., “A Three-Dimensional User Interface for Neurosurgical Visualization,” Proc. SPIE Conf. on Medical Imaging, 1994.

83. Hinckley, K., Pausch, R., Proffitt, D., Kassell, N., “Attention and Visual Feedback: The Bimanual Frame-of-Reference,” currently submitted to ACM 1997 Symposium on Interactive 3D Graphics.

84. Hinckley, K., Pausch, R., Proffitt, D., Patten, J., Kassell, N., “Cooperative Bimanual Action,” currently submitted to ACM CHI’97 Conference on Human Factors in Computing Systems, 1997.

85. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell, N., “Usability Analysis of 3D Rotation Techniques,” currently submitted to ACM CHI’97 Conference on Human Factors in Computing Systems, 1997.

86. Hinckley, K., Goble, J., Pausch, R., Kassell, N., “New Applications for the Touchscreen in 2D and 3D Medical Imaging Workstations,” Proc. 1995 SPIE Conference on Medical Imaging.

87. Honda, H., “Rightward superiority of eye movements in a bimanual aiming task,” Quarterly J. Exp. Psych., 34A, pp. 499-513, 1982.

88. Interlink Electronics, (805) 484-8855.

89. International Technologies, Inc., (401) 781-5595.

90. Iwata, H., “Artificial Reality with Force-feedback: Development of Desktop Virtual Space with Compact Master Manipulator,” Computer Graphics, 24 (4), pp. 165-170.

91. Jacob, R., Sibert, L., “The Perceptual Structure of Multidimensional Input Device Selection,” CHI’92, pp. 211-218.

92. Jackson, T. R., Goble, J. C., Laws, E. R., Kassell, N. F., “Tissue Selective Magnetic Resonance Subtraction Venography,” Proc. 1995 SPIE Conference on Medical Imaging.

93. Judd, C. M., McClelland, G. H., ***Data Analysis: A Model-Comparison Approach***, Harcourt Brace Jvanovich: San Diego, CA., 1989.

94. Kabbash, P., MacKenzie, I. S., Buxton, W., “Human Performance Using Computer Input Devices in the Preferred and Non-Preferred Hands,” INTERCHI’93, pp. 474-481.

95. Kabbash, P., W. Buxton, A. Sellen, “Two-handed input in a compound task,” ACM CHI’94 Conference on Human Factors in Computing Systems, 1994, pp. 417-423.

96. Karl, L., Pettey, M., Shneiderman, B., “Speech-Activated versus Mouse-Activated Commands for Word Processing Applications: An Empirical Evaluation,” Intl. J. Man-Machine Studies, 1993.

97. Katz, W. T., “Segmentation of the Brain from 3-D Magnetic Resonance Imaging of the Head,” Proc. of the Satellite Symposium on 3D Advanced Image Processing in Medicine, 1, 1992, pp. 213-216.

98. Kaufman, A., Yagel, R., “Tools for Interaction in Three Dimensions,” Proc. 3rd International Conf. on HCI (Boston, MA), Vol. 1, Sept. 1989, pp. 468-475.

99. Kelly, P. J., B. A. Kall, eds., ***Computers in Stereotactic Neurosurgery***, Blackwell Scientific Publications Inc., Cambridge, MA, 1992.

100. Kikinis, R., P.L. Gleason, W.E. Lorensen, W.M. Wells III, W.E.L. Grimson, T. Lozano-Perez, G. Ettinger, W. White and F.A. Jolesz, “Image guidance techniques for neurosurgery,” Visualization in Biomedical Computing 1994, Robb, R. A., ed., Proc. SPIE 2359, October 1994, pp. 537-540.

101. Kimura, D., Humphrys, C, “A Comparison of Left- and Right-Arm Movements During Speaking,” Neuropsychologia, 19 (6), 1981, pp. 807-812.

102. Kimura, D., “Manual Activity During Speaking- I. Right-Handers and II. Left-Handers,” Neuropsychologia, 11, 1973, pp. 45-55.

103. Krueger, M., “Environmental Technology: Making the Real World Virtual,” Communications of the ACM, 36 (7), 1993, pp. 36-37.

104. Krueger, M., ***Artificial Reality II***, Addison-Wesley: Reading, MA, 1991.

105. Krueger, M., Gionfriddo, T., Hinrichsen, K., “VIDEOPLACE--An Artifical Reality,” Proc. of ACM CHI’85, pp. 35-40.

106. Lasseter, J., “Principles of Traditional Animation Applied to 3D Computer Animation,” Computer Graphics (Proc. ACM SIGGRAPH ‘87), 21 (4), July 1987, pp. 35-44.

107. Leblanc, A., Kalra, P., Magnenat-Thalmann, N., Thalmann, D., Sculpting with the “Ball and Mouse” Metaphor, Graphics Interface ‘91, pp. 152-159.

108. Leganchuk, A., Zhai, S. Buxton, W., “Bimanual direct manipulation in area sweeping tasks,” 1995. URL: http://www.dgp.toronto.edu/people/andrea/bimanual.html.

109. Leksell, L. “A stereotactic apparatus for intracerebral surgery,” Acta Chir Scand 99, pp. 229-233, 1949.

110. Liang, J., Shaw, C., Green, M., “On Temporal-Spatial Realism in the Virtual Reality Environment,” UIST’91, pp. 19-25.

111. Liang, J. , Green, M., “JDCAD: A Highly Interactive 3D Modeling System,” 3rd International Conference on CAD and Computer Graphics, Beijing, China, Aug. 1993, pp. 217-222.

112. Liang, Jiandong, ***Interaction Techniques for Solid Modeling with a 3D Input Device***, Ph.D. Dissertation, University of Alberta, Dept. of Computing Science, Edmonton, Alberta, Fall 1995.

113. Mackenzie, C., Iberall, T., ***The Grasping Hand***, Advances in Psychology 104, Stelmach, G., Vroon, P., eds., North Holland, Amsterdam, 1994.

114. MacKenzie, I. S., “Fitts’ law as a research and design tool in human-computer interaction,” Human-Computer Interaction, 7, pp. 91-139, 1992.

115. Mackinlay, J., Card, S., Robertson, G., “Rapid Controlled Movement Through a Virtual 3D Workspace,” Comp. Grap., 24 (4), 1990, 171-176.

116. Mapes, D., Moshell, J. M., “A Two-Handed Interface for Object Manipulation in Virtual Environments,” Presence, 4 (4), Fall 1995, pp. 403-416.

117. Marteniuk, R., MacKenzie, C., Baba, D., “Bimanual movement control: Information Processing and interaction effects,” Q. J. Exp. Psych., 36A, 335-365.

118. McKenna, M., “Interactive Viewpoint Control and Three-dimensional Operations,” Proc. 1992 Symposium on Interactive 3D Graphics, pp. 53-56.

119. Mine, M., http://www.cs.unc.edu/~mine/chimp.html.

120. Minsky, M., Ouh-young, M., Brooks, F. P., Behensky, M., “Feeling and Seeing: Issues in Force Display,” Comp. Graph., 24 (2), 234-244.

121. Multigen Inc., http://www.multigen.com/smart.htm.

122. Multimedial Medical Systems, Inc., http://www.mmms.com.

123. Nielsen, J., “Noncommand User Interfaces,” Communications of the ACM, 36 (4), pp. 83-99.

124. Nielsen, J., ***Usability Engineering***, Academic Press: Morristown, NJ, 1993.

125. Nielson, G., Olsen, D., “Direct Manipulation Techniques for Objects Using 2D Locator Devices,” 1986 ACM Workshop on Interactive 3D Graphics, Chapel Hill, NC, October 1986, pp. 175-182.

126. Norman D., ***The Design of Everyday Things***, Doubleday: New York, New York, 1990.

127. Oldfield, R., “The assessment and analysis of handedness: The Edinburgh inventory,” Neuropsych-ologia, 9, pp. 97-113, 1971.

128. Osborn, J., Agogino, A., “An Interface for Interactive Spatial Reasoning and Visualization,” Proc. CHI’92, pp. 75-82.

129. Ostby, E., “Describing Free-Form 3D Surfaces for Animation,” Proc. ACM Workshop on Interactive 3D Graphics, Oct. 1986, pp. 251-258.

130. Pausch, R., ***Support for Rapid Prototyping of Two- and Three-Dimensional User Interfaces***, Proposal for ARPA BAA 93-42. Comp. Science Department, University of Virginia, March, 1994.

131. Pausch, R., Shackelford, M. A., Proffitt, D., “A User Study Comparing Head-Mounted and Stationary Displays,” Proc. IEEE Symposium on Research Frontiers in Virtual Reality, Oct. 1993.

132. Pausch, R., Vogtle, L., Conway, M., “One Dimensional Motion Tailoring for the Disabled: A User Study,” Proc. CHI’92, pp. 405-411.

133. Pausch, R., Burnette, T., Brockway, D., Weiblen, M., “Navigation and Locomotion in Virtual Worlds via Flight into Hand-Held Miniatures,” Computer Graphics (SIGGRAPH ‘95), pp. 399-400.

134. Peters, M., “Constraints in the performance of bimanual tasks and their expression in unskilled and skilled subjects,” Q. J. Exp. Psych., 37A, pp. 171-196.

135. Phillips, C., Badler, N., Granieri, J., “Automatic Viewing Control for 3D Direct Manipulation,” Computer Graphics (Proc. 1992 Symposium on Interactive 3D Graphics), pp. 71-74.

136. Pixsys Inc., 3522 22nd St., Boulder, CO 80304. (303) 443-0771.

137. Polhemus Navigation Sciences, Inc., P. O. Box 560, Colchester, VT 05446. (802) 655-3159.

138. Poston, T., Serra, L., “Dextrous Virtual Work,” Communications of the ACM, 39 (5), pp. 37-45, 1996.

139. Poulton, E., Freeman, P., “Unwanted asymmetrical transfer effects,” Psychlogical Bulletin, 66 (1), 1966.

140. Provins, K., Glencross, D., “Handwriting, typewriting and handedness,” Q. J. Exp. Psych., 20, pp. 282-289.

141. Sachs, E., Roberts, A., Stoops, D., “3-Draw: A Tool for Designing 3D Shapes,” IEEE Computer Graphics & Applications, Nov. 1991, pp. 18-26.

142. Schmandt, C. M., “Spatial Input/Display Correspon-dence in a Stereoscopic Computer Graphic Work Station,” Computer Graphics, 17 (3), 1983, pp. 253-262.

143. Sears, A.,Plaisant, C., Shneiderman, B., “A New Era for High Precision Touchscreens,” in Advances in Human-Computer Interaction, Hartson, Hix, eds., Vol. 3, 1992, pp. 1-33.

144. Sears, A., Shneiderman, B., “High Precision Touchscreens: Design Strategies and Comparisons with a Mouse,” International Journal of Man-Machine Studies, 34 (4), 1991, pp. 593-613.

145. Sellen, A. J., Kurtenbach, G. P., Buxton, W. A. S., “The Role of Visual and Kinesthetic Feedback in the Prevention of Mode Errors,” Proc. IFIP INTERACT’90, pp. 667-673.

146. Sellers, M., “Designing for Demanding Users,” ACM Interactions, July 1994, pp. 55-64.

147. SensAble Devices, Inc., The PHANToM Haptic Interface. URL: http://www/tiac.net/users/sensable/

148. Serra, L., Poston, T., Ng, H., Chua, B., Waterworth, J., “Interaction techniques for a virtual workspace,” Proc. ICAT/VRST’95 - Virtual Reality Software and Technology 1995, Chiba, Japan.

149. Shaw, C., Green, M., “Two-Handed Polygonal Surface Design,” ACM UIST’94 Symposium on User Interface Software & Techn*ology*, pp. 205-212.

150. Shaw, C., “Two-Handed 3D Editing,” ACM CHI’96 Workshop on the Psychological Issues of Virtual Environment Interfaces, April 1996. URL: http://www.cs.uregina.ca/~cdshaw/papers/twohand/twohand.html.

151. Shepard, R. N., Metzler, J., “Mental Rotation of Three-Dimensional Objects,” Science, Vol. 171, 1971, pp. 701-703.

152. Shoemake, K., “Animating Rotations with Quaternion Curves,” Computer Graphics, 19 (3), 1985, 245-254.

153. Shoemake, K., “ARCBALL: A User Interface for Specifying Three-Dimensional Orientation Using a Mouse,” Graphics Interface, 1992, 151-156.

154. Shoemake, K., “Arcball Rotation Control,” in *Graphics Gems IV*, Heckbert, P., ed., Academic Press, pp. 175-192.

155. Shoemake, K., ***Quaternions***, May 1994, available at URL ftp://ftp.cis.upenn.edu/pub/graphics/shoemake/quatut.ps.Z.

156. Smith, R. B., “Experiences with the Alternate Reality Kit: An Example of the Tension between Literalism and Magic,” Proceedings of ACM CHI+GI’87 Conference on Human Factors in Computing Systems and Graphics Interface, 1987, pp. 61-67.

157. Snell, J., M. Merickel, J. Ortega, J. Goble, J. Brookeman, N. Kassell, “Segmentation of the brain from 3D MRI using a hierarchical active surface template,” SPIE Conference on Medical Imaging, 1994, pp. 2-9.

158. Snell, J. W., Merickel, M. B., Ortega, J. M., Goble, J. C., Brookeman, J. R., Kassell, N. F., “Model-Based Boundary Estimation of Complex Objects Using Hierarchical Active Surface Templates,” Pattern Recognition, 1995.

159. Snell, J. W. et al, “Model-Based Segmentation of the Brain from 3-D MRI Using Active Surfaces,” Proc. of the SPIE Conference on Medical Imaging, Newport Beach, CA, 1993.

160. Snell, J. W., Jackson, T. R., Katz, W. T., Hinckley, K., Goble, J. C., Kassell N. F., “A Three-Dimensional Stereotactic Neurosurgical Planner/Simulator”, Proc. 1995 SPIE Conference on Medical Imaging.

161. Spaceball Technologies, Inc. (508) 970-0330.

162. Spectra Symbol, (801) 972-6995.

163. Stassen, H., Smets, G., “Telemanipulation and Telepresence,” 6th Symp. on Analysis, Design, and Evaluation of Man-machine Systems, 1995, pp. 13-23.

164. Stoakley, Conway, Pausch, “Virtual Reality on a WIM: Interactive Worlds in Miniature,” CHI’95, pp. 265-272.

165. Sturman, D., Zeltzer, D., Pieper, S., “Hands-On Interaction with Virtual Environments,” UIST’89, pp. 19-24.

166. I. E. Sutherland, “A Head-mounted Three Dimensional Display,” Proc. the Fall Joint Computer Conference, 1968, pp. 757-764.

167. Takemura, H., Tomono, A., Kayashi, Y., “An Evaluation of 3-D Object Pointing Using a Field Sequential Stereoscopic Display,” Proc. Graphics Interface ‘88, June 1988, pp. 112-118.

168. Taylor, R., Robinett, W., Chi, V., Brooks, F., Wright, W., Williams, R., Snyder, E., “The Nanomanipulator: A Virtual-Reality Interface for a Scanning Tunneling Microscope,” SIGGRAPH’93, pp. 127-134.

169. Tognazzini, B., “Principles, Techniques, and Ethics of Stage Magic and Their Application to Human Interface Design,” Proc. INTERCHI’93, pp. 355-361.

170. Torborg, J., Kajiya, J., “Talisman: Commodity Realtime 3D Graphics for the PC,” Computer Graphics (Proc. SIGGRAPH’96), pp. 353-363.

171. Trident Systems Inc., 10201 Lee Highway, Suite 300, Fairfax, VA 22030. (703) 273-1012.

172. University of North Carolina, Dept. of Computer Science, Chapel Hill, NC. Unpublished “Virtual Golf” virtual reality demonstration.

173. University of Virginia user interface group. http://www.cs.virginia.edu/~uigroup/

174. Verbex Voice Systems, Inc., 1-800-ASK-VRBX.

175. Waccom Technology Corporation, 501 S.E. Columbia Shores Blvd. #300, Vancouver, WA 98661, 360-750-8882.

176. Ware, C., “Using Hand Position for Virtual Object Placement,” Visual Comp., 6 (5), 1990, 245-253.

177. Ware, C., Osborne, S., “Exploration and Virtual Camera Control in Virtual Three Dimensional Environments,” Comp. Graph., 24 (2), 175-183.

178. Ware, C., Arthur, K., Booth, K. S., “Fish Tank Virtual Reality,” INTERCHI’93, pp. 37-41.

179. Colin Ware, D. R. Jessome, “Using the Bat: A Six-Dimensional Mouse for Object Placement,” IEEE Computer Graphics and Applications, November 1988, 65-70.

180. D. Weimer, S. K. Ganapathy, “A Synthetic Visual Environment with Hand Gesturing and Voice Input,” Proc. ACM CHI’89 Conference on Human Factors in Computing Systems, 1989, pp. 235-240.

181. Weiser, M., “The Computer for the 21st Century,” Scientific American, Sept. 1991, pp. 94-104.

182. Wellner, P., “Interacting with Paper on the DigitalDesk,” Communications of the ACM, 36 (7), 1993, pp. 87-97.

183. Wexelblat, A., “Gesture at the User Interface,” Workshop at the ACM CHI’95 Conference on Human Factors in Computing Systems. URL: http://wex.www.media.mit.edu/people/wex/workshop.html.

184. Wing, A., “Timing and co-ordination of repetitive bimanual movements,” Quarterly. J. Exp. Psych., 34A, pp. 339-348, 1982.

185. Wolff, P., Hurwitz, I., Moss, H., “Serial organization of motor skills in left- and right-handed adults,” Neuropsychologia, 15, pp. 539-546, 1977.

186. Zeleznik, R., Herndon, K., Hughes, J., “SKETCH: An Interface for Sketching 3D Scenes,” Proc. SIGGRAPH’96, pp. 163-170.

187. Zhai, S., Milgram, P., Buxton, W., “The Effects of Using Fine Muscle Groups in Multiple Degree-of-Freedom Input,” Proc. ACM CHI’96, pp. 308-315.

188. Zhai, S., Buxton, W., Milgram, P., “The “Silk Cursor”: Investigating Transparency for 3D Target Acquisition,” CHI’94, pp. 459-464.

189. Zimmerman, T., Lanier, J., Blanchard, C., Bryson, S., Harvill, Y., “A Hand Gesture Interface Device,” Proc. of ACM CHI+GI’87 Conference on Human Factors in Computing Systems and Graphics Interface, pp. 189-192.

190. Zimmerman, T., Smith, J. R., Paradiso, J., Allport, D., Gershenfeld, N., “Applying Electric Field Sensing to Human-Computer Interfaces,” Proc. ACM CHI’95 Conference on Human Factors in Computing Systems, pp. 280-287.