Cubby: What You See Is Where You Act

Interlacing the display and manipulation spaces

ISBN 90-9011724-5

Cubby: What You See Is Where You Act

Interlacing the display and manipulation spaces

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft
op gezag van de Rector Magnificus prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,
op maandag 15 juni 1998 te 15:00 uur
door

Johan Partomo DJAJADININGRAT
Master of Design, Royal College of Art, Londen

geboren te Rotterdam

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. G.J.F. Smets

Samenstelling promotiecommissie:

Rector Magnificus,	*voorzitter*
Prof. dr. G.J.F. Smets,	Technische Universiteit Delft, Open Universiteit Nederland, *promotor*
Dr. C.J. Overbeeke,	Technische Universiteit Delft, *toegevoegd promotor*
Prof. dr. ir. C.A. Grimbergen,	Technische Universiteit Delft
Prof. ir. G. Honderd,	Technische Universiteit Delft
Prof. W.S. Green,	Technische Universiteit Delft
Prof. dr. ir. J.F. Gerrissen,	Open Universiteit Nederland
Dr. A. Dunne,	Royal College of Art, Londen

This work is supported by the Dutch Technology Foundation (STW)
under grant STDIO22-2732.

Table of Contents

Acknowledgments

"Don't write this book. You perform a disservice to a field of enquiry that has always struggled for respectability. No man of talent could describe the events that occurred in any realistic fame because they deal with alternative realities which we are yet to comprehend. When presented in the wrong way and the wrong context the incidents and the people involved in them can appear foolish, if not downright psychotic." (Mulder, 1997)

First of all, my special thanks go to Prof. Dr. Gerda 'Bof doen' Smets, for having offered me the unique opportunity to take up a post as a research student, and for showing a real interest in vormgeving. She had the insight that a system like Cubby would be interesting, and dared to invest in its development when others, including myself, were — to put it mildly — still somewhat sceptical.

I am very grateful to Kees 'Zeg kom, mannen' Overbeeke for his daily guidance. Kees, I cannot thank you enough for ploughing through my manuscripts umpteen times, restructuring, refining and rephrasing them until things finally made sense. You have the ability to reduce things to their normal proportions and to really motivate people. It is a joy working with you.

I am greatly indebted to Pieter-Jan 'Moody' Stappers, 'the man who can handle Mathematica with an ease as if it were a pocket calculator'. He had the great icosahedron idea and showed me how to implement collision detection. P.J., without you manipulation in Cubby would never have worked. Most of the time I don't understand what you mean until days later, but once I do, something good comes out! I also thank you for all the work you put into the analysis of the manipulation experiment.

Caroline 'can you hand me the thingy' Hummels gave infinite amounts of moral support in difficult times, helped with generating conceptual interfaces, always showed enthusiasm for discussing design ideas, and did not shrink from defacing a table with an electric saw to build Cubby Mk I. Caroline, you're next in line! Are you positive we have destroyed all evidence of that project for you-know-who?

I thank Wouter 'Pacman' Pasman for his programming magic when the brand new 20,000 guilder Mac stubbornly refused to reliably read its serial ports. Not that I panicked or anything, of course not. Still, I was 'somewhat' grateful when you solved it. I also thank you for carrying out the measurements of the delay of Cubby's video components, and for painstakingly checking Cubby's calibration by staring through a Dynasight reflector.

Maarten ‘Mac Guru’ Gribnau generously shared his knowledge and code. Most importantly, he wrote an excellent driver for the Flock of Birds tracker, thus contributing an indispensable part to manipulation in Cubby. If you hadn’t, Maarten, I would still be trying to write that driver now.

Fred ‘Wolfram Jr.’ Voorhorst co-designed the PDA, Wobbly, a project which of course was dropped much too soon. Fred, one day we'll make that foldable scooter, and that 3DOF ADB mouse, and that killer application for a PDA. I guess that will be the same day that surgeons use Cubby, and airplane mechanics use lime green, lasergun shaped, shadow parallax boroscopes! Maybe now we can finally get our priorities right and play some mahyong.

I am very grateful to Rob Luxen, ‘the man who treats MDF components with the same precision as SMD components’. Thank you for building all those Cubbies (including one at bar height), amplifiers, and video multipliers. I like the way you had one of Cubby’s corners chamfered, it really does lend the set-up that certain elegance.

I also thank Toni Subroto, engineer *and* craftsman, for his great work on the PDA prototype and the mechanical head tracker. “Stop, Toni, there's no need to further optimise the strength-to-weight ratio, it doesn't need to get airborne!”. Oh, and don't go into computers, Toni, mechanical things are so much more interesting!

Stephan ‘Naomi’ Wensveen has got the enviable ability to appear really laid-back whilst completing loads of work. He put much work in the visualisation experiment: building the set-up (science is all about ‘Birdy Friend’!), conducting the experiment (“We create a 3D illusion by coupling your head movements to the perspective”), applying his ‘Kalashnikov-Smirnoff’ test method, and designing the icosahedron’s texture mapping. Steeph, at least you learnt that ‘paranimfen’ and naked women descending on parachutes are unrelated!

Finally, special thanks go to my sister Aletta and brother-in-law Matthew for always being there to help with translation problems and for proof-reading parts of the manuscript.

Illustration Acknowledgments

I gratefully acknowledge the following persons, companies and institutions for giving their permission to reproduce their drawings and photographs. All illustrations not listed here are my own.

Figure 1.1	Philips Medical Systems
Figure 1.2	C.A. Grimbergen and J.A. Snel of the Amsterdam Medical Centre.
Figure 1.3	W.P.Th.M. Mali (Afdeling Radiodiagnostiek of the Academische Ziekenhuis Utrecht) and Philips Medical Systems
Figure 1.4	ISG Technologies
Figure 1.5	Philips Medical Systems
Figure 1.6	ISG Technologies
Figure 1.7	ISG Technologies
Figure 2.2	Stereographics Corporation
Figure 2.3	Stereographics Corporation
Figure 3.1l	Odin
Figure 3.1m	Minolta Benelux
Figure 3.1r	Sony Nederland
Figure 3.2	Bang&Olufsen Nederland
Figure 3.4	R. Nakata
Figure 3.5	P. Montgomery could not be found
Figure 3.6	M. Clymer did not respond
Figure 3.7	Apple Nederland
Figure 3.9	D.A. Norman
Figure 4.1l	3COM Nederland
Figure 4.1r	Apple Nederland
Figure 5.1	Philips Medical Systems
Figure 6.13	Uitgeverij Bohn Stafleu van Loghem
Figure 6.16	Stichting Academisch Rekencentrum Amsterdam
Figure 6.17	Stichting Academisch Rekencentrum Amsterdam
Figure 9.1	C. Schmandt and Association for Computing Machinery
Figure 9.2	K. Kameyama

0

An Overview

This thesis forms part of a project entitled “Telepresence: implementation and model-forming of a working principle based on the perceptual meaning of active parallax shifts (STDIO-22.2732)”, sponsored by the Dutch Technology Foundation. The working principle referred to in this description is the Delft Virtual Window System (DVWS) (Overbeeke et al., 1987; Smets et al., 1987; Overbeeke and Stratmann, 1988; Smets et al., 1988). The DVWS gives an observer a three-dimensional (3D) impression of a scene displayed on a conventional monitor screen by coupling the parallax shifts on that screen to his head movements. A generic term to describe systems such as the DVWS is ‘head-coupled movement parallax’ (Figure 0.1).

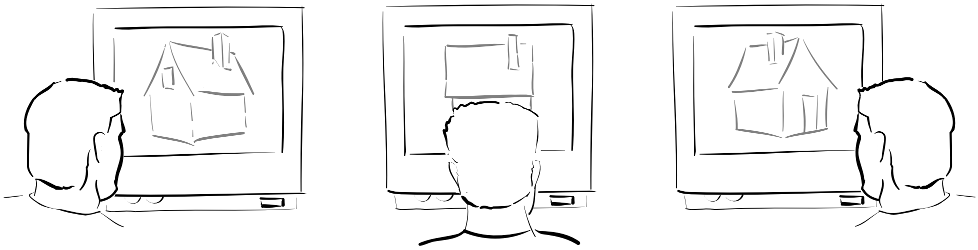


Figure 0.1
An observer looking at a display featuring head-coupled movement parallax. By moving his head he can view the scene — in this case a house — from different perspectives.

Movement parallax is a term from perception psychology. It is one of several depth cues, i.e. types of information which allow a person to perceive depth. Other depth cues include, but are not limited to, linear perspective, textural gradients, occlusion and shadows (Wickens, 1990). Movement parallax is different from motion parallax. In motion parallax, the parallax shifts are not caused by the observer, but by parts of the scene moving relative to each other. An example of motion parallax is when an observer looks at a bracelet which is displayed on a rotating disk in a shop window. In movement parallax, the parallax shifts are caused by body movements made by the observer. Body movements include both hand and head movements. An example of hand-coupled movement parallax is when an observer looks at a bracelet which he turns around in his hand. An example of head-coupled movement parallax is when an observer looks at a bracelet which lies in a display cabinet and moves his head to view the bracelet from various sides. The term 'head-coupled movement parallax' is used to refer explicitly to parallax shifts caused by the observer's head movements. In the user interface and engineering community, displays based on head-coupled movement parallax are often referred to as 'head-tracked displays'.

Head-coupled movement parallax may be of benefit to highly diverse applications, as is illustrated by the subjects of the three theses which form part of this project. Pasman researched luggage inspection (Pasman, 1997; Pasman et al., 1997a), while Voorhorst researched laparoscopy (Voorhorst et al., 1997; Voorhorst, 1998). I have chosen to use head-coupled movement parallax to arrive at an intuitive interface for medical 3D systems.

In the remainder of this introduction I will give an overview of the chapters. Human-computer interface design, computer graphics and medical science all make generous use of jargon. As this thesis borrows from all three subjects, it too suffers from a considerable amount of jargon. I hope the back flap may help to explain the terminology used in this thesis.

Chapter 1, '3D and the Medical Sciences', provides some background information. It discusses how medical 3D data are acquired and the ways in which these data are presented. It also addresses the question as to who would benefit from a medical 3D system with an improved interface. A number of possible applications within the medical sciences for medical 3D systems are given.

In chapter 2, 'Assessing display methods on usability', 3D displays are reviewed. I argue why desktop virtual reality is preferable to immersive virtual reality for medical 3D systems. 3D displays for desktop VR are assessed according to two criteria. The first one is unobtrusiveness. The display method should not hamper the user in terms of mobility and communication. The second criterion is the possibility of unifying the display and manipulation spaces, so that virtual objects can be directly manipulated, either by hand or by means of an instrument. It is this unification of display and manipulation spaces which is the central theme of this dissertation. I think that the non-intuitive manipulation of virtual objects in current 3D systems can be blamed mainly on the separation between the display space and the manipulation space. Many of todays 3D systems tease the user by showing highly lifelike virtual objects locked away behind a screen where he cannot reach them, and frustrate him by forcing him to use input devices which ignore the skills he has developed in everyday life.

Chapter 3 looks at how formgiving can make products intuitive to use. I focus on J.J. Gibson’s theory of affordances, and how it differs from existing ‘good practice’ in product interface design. The idea behind looking at product design is that, to make a human-computer interface more user-friendly, it should become less a computer and more a product. For a medical 3D system, this is particularly relevant, since the user is a non-technical person and the task does not allow for mistakes or delays. The interface should hide the computer as much as possible, allowing the user to focus completely on the task in hand.

In chapter 4, a number of design concepts for medical 3D systems are proposed. These concepts include desktop systems and one hand-held system. While the desktop systems are based on head-coupled movement parallax, the hand-held system uses hand-coupled movement parallax. The benefits of a 3D display for hand-held computers in general are discussed, as well as the advantages of the particular implementation presented in this chapter.

In chapter 5 I turn to input devices. Many currently available medical 3D systems use conventional input devices with two degrees of freedom, such as a mouse or a trackball. Others use special input devices which allow control of one degree of freedom at a time. In an experiment, it is shown that it is the number of simultaneously accessible degrees of freedom which are important for intuitive rotation of virtual objects.

In chapter 6, a new desktop VR system named Cubby is introduced. It uses head-coupled movement parallax on three orthogonal screens which form a concave cubic space. Because the virtual scene appears in front of the screens, it is physically accessible to the user, so that the display and manipulation spaces can be unified. Cubby is compared to a single-screen movement parallax display and to CAVE (DeFanti et al., 1993; Cruz-Neira et al., 1993), Cubby’s larger cousin.

In the first Cubby prototypes, the virtual objects appeared rubbery and distorted. In chapter 7, the possible causes of this deformation are systematically investigated and improvements are made.

In chapter 8, depth perception in Cubby is put to the test. Depth perception in a virtual scene with a headfree tracker and a non-headfree tracker is compared to that in a real scene.

In chapter 9, manipulation is implemented. By means of a stylus, the user can manipulate objects in Cubby with six simultaneous degrees of freedom. Because the display and manipulation spaces are integrated, the virtual objects and the stylus can occupy the same space. In an experiment, a unified and a non-unified Cubby are compared to a unified and a non-unified single-screen movement parallax display.

Finally, in chapter 10, overall conclusions are drawn and some recommendations are made for Cubby’s further development.

3D and the Medical Sciences

Summary

To put the work in this thesis into context I will give some background information on the use of three dimensional computer models in medicine. Through the literature and personal communication an account is given of how radiologists, who are generally considered the primary user group of medical 3D systems, view the use of 3D computer graphics in medicine. I then discuss the literature with regard to the areas of application for medical 3D computing, being visualisation, pre-operative simulation, operative support and education. Finally, I give two examples of medical procedures which already benefit from 3D visualisation, namely stereotactic and craniofacial surgery.

Introduction

Medical scanners allow a patient to be examined in a non-invasive manner. There are different types of scanners on the market, which provide different information. Examples include computer tomography (CT)[1], magnetic resonance imaging (MRI)[2] and positron emission tomography (PET)[3]. Figure 1.1 shows a MRI scanner. Medical scanners use different technologies and have their own strengths and weaknesses in terms of the type of information acquired. These different types of information are

---

[1] computer tomography is an X-ray based technique which makes use of the fact that different types of tissue — depending on their density — exhibit different amounts of absorption of X-ray energy. CT shows anatomical detail. The CT method has three main advantages over the conventional X-ray method. First, because CT images the body in slices it does not suffer from objects situated in depth being superimposed. Second, it has much greater sensitivity allowing differentiation between soft tissues. Third, it allows quantitative measurement of the densities of individual substances (Hounsfield, 1980).

[2] magnetic resonance imaging is a technique in which hydrogen protons react to changes in a magnetic field. An MRI scanner can — depending on its settings — show both anatomical detail and neuronal activity (Raichle, 1994).

Figure 1.1
An MRI scanner.

referred to as imaging modalities. While they use different technologies medical scanners have one important characteristic in common: the data they output is 3D in nature. However, the data typically come as a stack of 2D slices, with the stacking providing the third dimension (Herman, 1993). Consequently, the most common way of presenting this information is not as 3D computer models but as an array of two dimensional slices. Figure 1.2 and Figure 1.3 show such arrays of 2D slices taken in two different directions. This way of presentation requires the person who is to interpret the scans to make a mental 3D reconstruction based on the 2D scans.

It is in fact possible to let a computer make a 3D reconstruction based on the slices (Figure 1.4). One of the difficulties in making 3D reconstructions is segmentation. For a surface-based model, this requires indication of where

---

[3] positron emission tomography is a technique whereby the emission of a previously administered radioisotope is measured. PET shows metabolism and may be used for the visualisation of neuronal activity (Raichle, 1994).

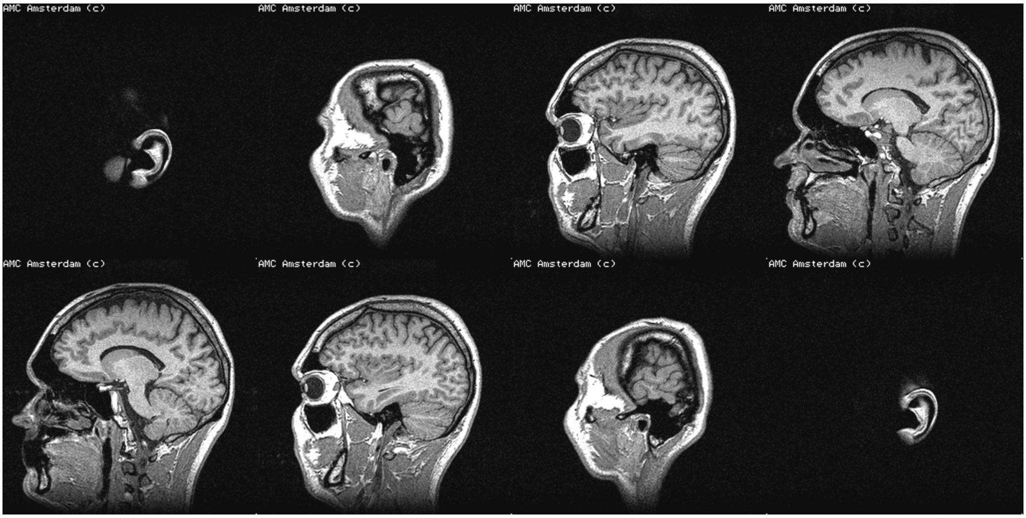


Figure 1.2
A number of slices of a human head taken with an MRI scanner.

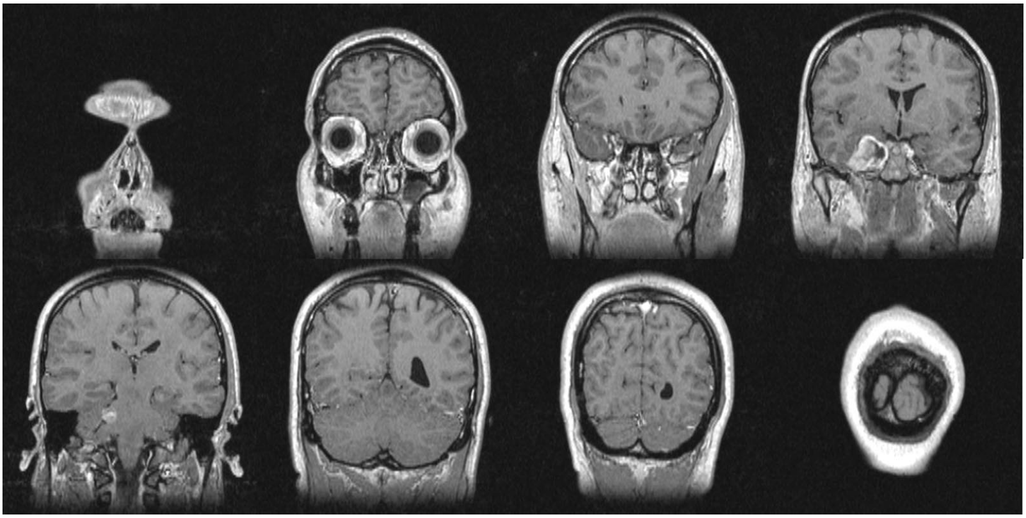


Figure 1.3
A second set of slices taken in a different direction.

surfaces need to be put. For a voxel-based model, it entails adding data to the model through which organs can be distinguished from each other (Schubert et al., 1993; Zubal

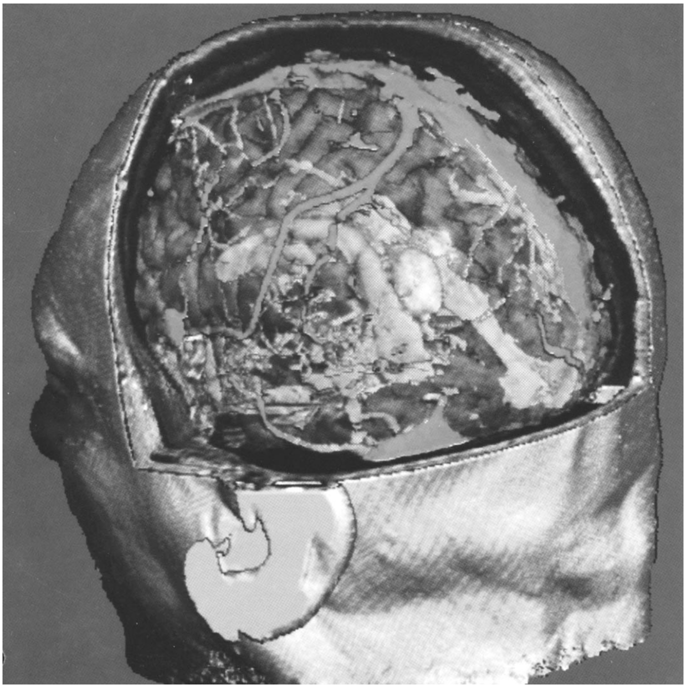


Figure 1.4
A 3D computer reconstruction.

et al., 1993). Through segmentation, organs and structures are delimited. They can then be shown or hidden as required. For example, in Figure 1.4 part of the skull is hidden to reveal the brain surface underneath. Interest in 3D visualisation is increasing as the problem of mentally reconstructing a 3D image becomes more pressing with the current trend to integrate different image modalities (Zuiderveld, 1995).

This thesis concentrates on the design of a 3D system which allows intuitive interaction with such a 3D computer reconstruction of the human body. Scanner design, integration of different imaging modalities and reconstruction of a 3D model are outside the scope of this thesis.

In this chapter I illustrate why I think there is a need to improve the interface of medical 3D systems. To do so I first explain who the users of such a medical 3D system could be and which areas lend themselves for application. I finish this chapter with two examples of medical procedures which already benefit from the application of 3D computer graphics, namely stereotactic and craniofacial surgery. Abstracted versions of these medical procedures will later be used as tasks to test experimental interfaces.

The potential user of a medical 3D system

In this section the users of medical 3D systems are introduced. Medical imaging has traditionally been performed by radiologists and thus they are often seen as the primary target group for medical 3D systems. Contrary to this notion radiologists do not whole-heartedly embrace a 3D computer reconstruction as the ultimate diagnostic aid for each and every medical case. Based on personal communication and the literature I attempt to give a more differentiated approach of their views on medical 3D systems and the way in which they use such systems. It is shown that users who benefit the most from 3D computer reconstructions are likely not to be the radiologists, who are predominantly concerned with diagnostics, but surgeons and other therapists.

Personal communication

I gratefully acknowledge the following radiologists for sharing their time and opinions on the use of 3D systems in their field: Professor Dr. H.E. Schütte MD and A.I.J. Klooswijk MD (Department of Radiology, Academic Hospital Dijkzigt, Rotterdam), Dr. G. Voorhout DVM (Faculty of Veterinary Science, University of Utrecht) and E.C.R. Wijffels MD (Department of Radiology, ‘Merwede Hospital’, Dordrecht).

Radiologists are confronted on a daily basis with two dimensional representations of the human anatomy. They are highly skilled in the ‘mental 3D reconstruction’ on the basis of 2D scans. Thus many of them view computer 3D reconstruction as unnecessary (Schütte, personal communication). An argument in favour of using the original 2D scans rather than the 3D reconstruction is that latter results in loss of detail (Voorhout, personal communication).

Klooswijk (personal communication) pointed out that a 3D reconstruction contains the same information as 2D scans, but presented in a different manner. He also remarked that radiologists are highly knowledgeable in anatomy, can relate 2D images to 3D reality without trouble and “can build a 3D image in their minds”. However, Klooswijk noted that there are situations in which 3D computer visualisation can be helpful. In general these are situations in which the patient’s anatomy has been disturbed to such an extent that recognition based on comparison with the normal anatomy is made difficult. This may be the case, for example, with severe trauma resulting from an accident.

All of the radiologists I spoke to claimed that computer 3D reconstruction is mainly useful for medical doctors who are not confronted with 2D scans on a daily basis and thus lack the skills to reliably ‘read’ such scans. Wijffels (personal communication) remarked that surgeons who have to relate the patient on the operating table to the images provided by the department of radiology benefit the most from 3D computer reconstructions.

Literature

In the literature the impression gained from personal communication with radiologists is a recurring theme.

The above quote by Herman (1993) summarises both the fact that 3D computer graphics are not wholeheartedly embraced by radiologists and the disbelief in this attitude by computer scientists. Although the radiology department does the actual scanning, imaging and first interpretation, other medical professions may yield more benefit from 3D computer reconstructions.

Merril (1993) reinstates this by saying: “While radiologists have spent years gaining the ability to perform such reconstruction, the general practitioner usually does not have the same skills to allow for the same degree of expertise”.

The ability of radiologists to make ‘mental’ 3D reconstructions may be based on familiarity with the images. An experiment showed that none of the medical doctors who were shown 2D CT scans of a fabric toy elephant could recognise the shape of an elephant (Evenblij, 1995, Bentum, 1995). This would be consistent with the assertion of the

“Are most radiologists, who appear to be not convinced by all these studies indicating the diagnostic usefulness of 3D displays privy to some super-scientific information so that they can justifiably say with Hamlet “there are more things in heaven and earth, Horatio, than are dreamt of in your philosophy”, or are we seeing an example of establishment rejecting something new and useful?” (Herman, 1993)

radiologists in the previous section that 3D reconstructions are mainly useful when the anatomy has been disturbed beyond recognition.

Areas of application for medical 3D systems

Having discussed the possible users of medical 3D systems, I now turn to the areas of application in which medical 3D systems may be of use. These areas can be broadly divided into four categories: visualisation, pre-operative simulation, operative support and education. An explanation of these areas is given.

Visualisation of a virtual body

A medical 3D system may be specifically aimed at the viewing of 3D reconstructions. Such a system is used for the viewing of a 3D reconstruction or part there of. In concrete terms it allows to the user to rotate a virtual body and to view cross-sections. It is not possible to manipulate parts of the body in the sense that they can be picked up, moved and placed else where.

Pre-operative simulation

In pre-operative simulation systems the user cannot only change his viewpoint on the model, but also manipulate parts of the virtual body. Such a system is not used during the actual operation but prior to it. It can be used for familiarisation with the problem, trying out different approaches and practising the preferred operation approach again and again.

Potentially such a system can have a beneficial impact on the quality of the operation and also reduce its duration. In turn this may lead to reduction of anaesthetic related problems and of costs. Note that operation may include non-invasive surgery such as radiation therapy.

Operative support

With a 3D system for operative support I mean a system which is an aid during the actual operation. Through the use of a 3D reconstruction it is possible to assist the surgeon in placing his instruments. Such image guided surgery is a way of ensuring that the point of entry into the body is as small and as near to the target area as possible.

Education

Traditional medical education

The most informative way to learn human anatomy is dissection of a subject. However, for most students such practical experience is available only in the early phases of their study (Schubert et al., 1993). Access to cadavers is becoming harder to get (Burdea and Coiffet, 1993).

Further information for study and reference is gained by means of anatomical textbooks and atlases. Contrary to dissection of a subject, atlases offer a limited number of drawings and photographs from which the anatomy needs to be mentally reconstructed. Another disadvantage of atlases is that they offer morphological information only. The student needs to be familiar with functional information since it is very difficult, if not impossible to deduce functional information from the morphological structures.

Modern preparation methods, such as the one by Von Hagen (Veldhoen, 1998), make possible the preservation of a complete human body or parts thereof for educational purposes. While such methods rival dissection of a cadaver in showing morphological structures in three dimensions, they cannot preserve the feel of the tissue and their potential for interaction is limited.

Computer based medical education

Hyper atlases or multimedia atlases integrate functional and morphological information. They combine information which traditionally was offered separately in the form of text books, atlases, video and audio tapes. Nilsson and Khakhar (1993) give an example in which the movements of the larynx are shown during breathing and the shaping of sounds. It integrates a graphical animation of the vocal chords, the produced sounds and a graphical representation of the resulting vibrations. Interactivity can allow the student to change parameters and observe the results.

Medical 3D systems

While multimedia atlases are generally limited to 2D interactivity, medical 3D systems can offer the student the possibilities to explore and manipulate the anatomy from various points of view. Merril (1993) points out that educational software developments have been hampered by using computers as ‘electronic page turning devices’. Applying computer technology in this way foregoes the essential aspects of medical education: anatomy is three dimensional and the processes in the body are dynamic.

A way to think about educational medical 3D systems is as a combination of the benefits of multimedia atlases and those of a 3D system for pre-operative simulation. For medical students who are trying to acquire operative skills and operative planning skills, it would be good to start with 3D virtual patients to explore various possibilities (Thalmann and Thalmann, 1993).

Medical tasks which benefit from 3D visualisation and manipulation

Craniofacial surgery

Craniofacial surgery is involved with operations of the face. Mostly such surgery is performed on children with congenital defects of the skull (Hattem, 1995). In craniofacial surgery parts of the skull may need to be removed, relocated or replaced by prostheses. It is one area of medicine in which planning by means of 3D computer models has become accepted. The reasons for this are the 3D complexity of the bone structure of the skull, the near presence of critical parts and the importance of fine control over the cosmetic end result.

Stereotactic tasks

In a stereotactic task a particular location in the body is targeted by means of equipment which provides a fixed frame of reference. Most often this concerns head operations or radiation therapy. In such procedures the spatial relationships between objects have to be taken into account in order to avoid destruction of vital brain structures and rupture of vasculature (Ehricke et al., 1992).

Head surgery

Before the patient is scanned a ring with so-called fiducial markers is firmly attached to the head or the fiducial markers are attached directly to the head (Figure 1.5). As these fiducial markers remain attached while the scan is being made they also show up on the scan and the resulting 3D reconstruction. When the patient is on the operation table — still with the markers present — the surgeon can calibrate his tracked instruments by touching the fiducial markers. The computer system can then display both the 3D reconstruction and the current placement of the surgeon’s instruments. This combination makes image-

Figure 1.5
A fiducial marker being attached directly to the scalp.

guided neurosurgery possible. The computer graphics show the 3D reconstruction with the neuroanatomy, pathological structures and the vasculature, as well as the stereotactic trajectories. By looking at the computer screen while moving his instruments the surgeon can choose the best point of entry and the best path. However, it should be noted that when the skull is opened, the brain deforms, and no longer exactly matches its computer-reconstructed counterpart. A similar problem results from the application of drugs, which are administered to make the brain shrink to allow room for operation. Figure 1.6 depicts a set-up with a probe pointing to a model of a human head and a monitor showing the placement of the probe in relation to a 3D computer reconstruction. Figure 1.7 shows a close-up of a computer screen with three orthogonal views and one perspective view.

Radiation beam therapy

Radiation beam therapy is also a form of a stereotactic task. In the case of radiation beam therapy the path does not represent a trajectory for a probe but that of one or

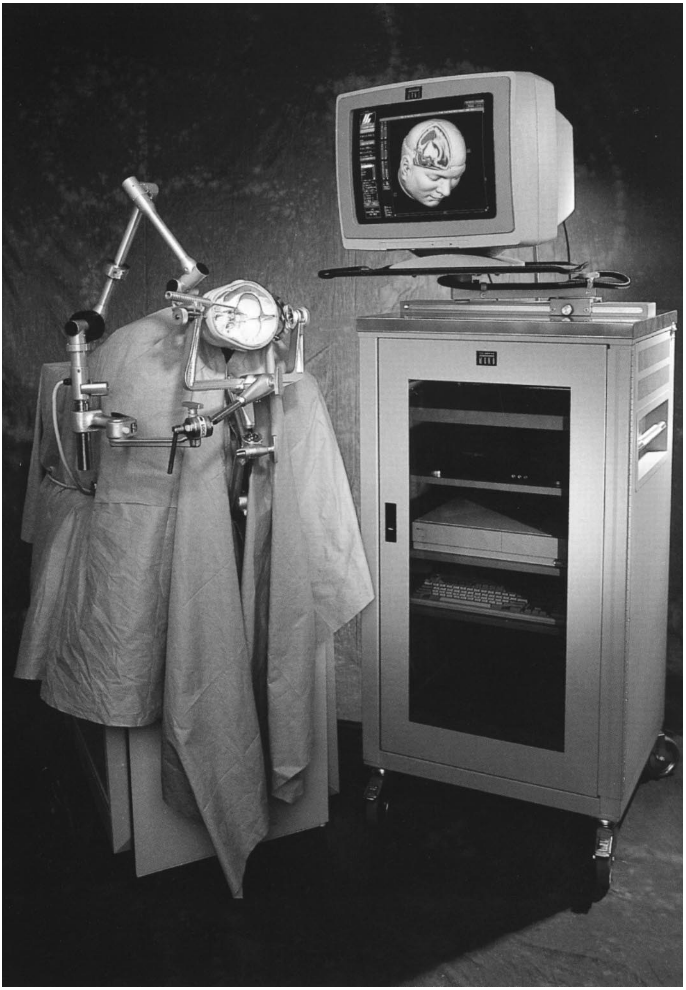


Figure 1.6
The stereotactic probe is clamped to a table and points to a model of a head. On the monitor the placement of the probe in relation to the 3D computer reconstruction can be seen.

more radiation beams. The essence is that the tumour receives a lethal dose of radiation while the surrounding healthy tissue does not.

The beam is directed in such a way that it always passes through the tumour. If a single beam is used it is moved so that tumour is constantly exposed to radiation while the normal tissue is not. When using multiple beams their positioning is such that they overlap and thus add radiation at the location of the tumour (Cook et al, 1987). Systems used to set out a strategy for radiation beam therapy are called radiation treatment planning systems (TPS).

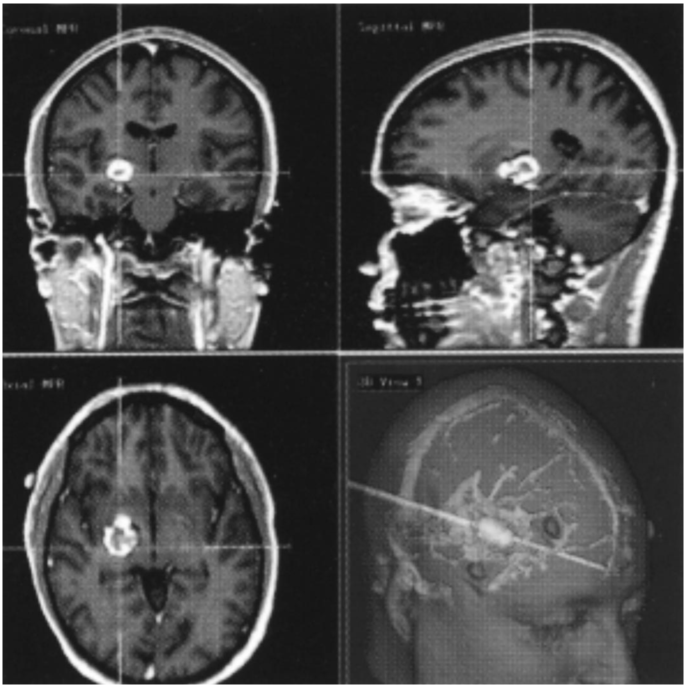


Figure 1.7
A monitor screen showing three orthogonal views and one perspective view. The location of the probe can be seen in all views.

Important aspects to be represented graphically are the radiation beam and its relationship to the normal anatomy and the tumour, and the isodose surfaces. For successful treatment registering is essential: the coordinate systems of the radiation source and the TPS need to coincide.

In the next chapter...

Having introduced the user of a medical 3D system, the areas of application and some tasks typical to the medical sciences, the next chapter considers 3D displays. This includes systems which are commercially available, as well as experimental systems. The systems which are considered are not limited to those designed for medical use. Instead systems designed for other areas of 3D graphics are also considered.

2

Assessing Display Methods on Usability

Summary

An intuitive 3D system requires careful consideration of the display method. Accordingly, in this chapter 3D display methods are reviewed, not only those specifically developed for the medical sciences but also those developed for other applications. As we are going to look at non-medical systems I point out what sets medical 3D systems apart from generic 3D systems and from 3D Computer Aided Design (CAD) in particular.

The suitability of 3D displays for use in a medical environment is assessed according to two criteria. The first one is that a medical 3D system ought to be as unobtrusive as possible. It should not hamper the user in his mobility. Neither should it hinder him in his communication with others. The second criterion is that the display method should allow the display and manipulation space to be unified, so that virtual objects can be directly manipulated, either by hand or through an instrument. A number of people (Schmandt, 1983; Kameyama et al., 1993a, 1993b; Ishii et al., 1994), including myself, think that this unification of display and manipulation space could allow the user to manipulate virtual objects with more confidence and higher accuracy. In chapter 9 I will come back to unified systems.

First I make the distinction between desktop virtual reality systems and immersive virtual reality systems and argue the advantages of desktop VR for medical applications. Desktop VR systems based on stereoscopy, movement parallax or both, and their relative merits are then considered. Two movement parallax systems, the Delft Virtual Window System and Fish Tank VR, are compared with regard to the second criterion. It is argued that which of these two systems is superior depends upon the task. I finish the section on 3D displays by setting out the reasons why I chose to base the systems in the remainder of this thesis on movement parallax only, rather than a combination of stereoscopy and movement parallax.

Difference between 3D CAD systems and medical 3D systems

At first sight medical 3D systems may seem to require similar considerations as 3D Computer Aided Design systems. Mann (1985) compares computer aided surgery (CAS) to computer aided design (CAD). He gives an example of how CAS allows the surgeon to perform virtual surgery, to judge the outcome visually, to simulate the results in an animation based on a musculo-skeletal model and start afresh until he approves of the procedure to be carried out on the real patient. Mann compares this to the way in which CAD allows the engineer to redesign an artifact until he approves of it for production.

There are, however, some significant differences between 3D systems for medical use and those used in engineering and design. Medical voxel-based models have an inside, and information inside structures can be revealed by taking a cross-section. 3D CAD models on the other hand are usually limited to surface representations. Although surfaces of CAD models can be complex, they are only hollow shells, since they enclose a volume of homogeneous material which need not be modelled in detail. Medical models are organic in form, while CAD models can be both organic and geometric in form. Unlike CAD models a surgical simulator works with non-rigid bodies and requires tissue behaviour modelling. In medical simulation collision detection has a higher priority than in CAD. In 3D CAD the emphasis is on both the creation and manipulation of models, while in medical 3D systems the emphasis is on manipulation. Tools in medical simulators generally reflect the instruments in the operating theatre, while those in 3D CAD do not necessarily have a physical counterpart.

3D Displays

In this section I first argue the benefits of desktop VR over immersive for a medical 3D system. I then turn to desktop VR systems based on stereoscopy, movement parallax or both. For overviews of all 3D displays see Overbeeke and Stratmann (1988) and Jones and Wyatt (1994).

Immersive VR vs. Desktop VR for medical applications

The archetypal immersive VR systems makes use of a helmet with built-in displays. As a consequence the user often is tethered. Cables and helmet weight hamper the movements of the user. More importantly, as perception of the real world is blocked out the communication between VR participants and non-participants is impossible. With a view to using VR in medical practice these mobility and communication aspects of immersive VR are a drawback. Hinckley et al. (1994) say: "... 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 employ devices that will be difficult to put down and it must not have explicit modes that are easily forgotten".

The Desktop VR systems, on the other hand, require less headware. These are described in the following section

Desktop VR Systems

In the following discussion desktop VR systems are divided into three categories: those based on stereoscopy, those based on movement parallax and those which use a combination of these two methods.

Stereoscopy based desktop VR

Stereoscopy is a well established depth cue in desktop VR systems. It is based on the phenomenon that the two images cast on the retinas of the observer differ because they are set apart by a few centimetres. Wade (1987) points out that the important role assigned to stereoscopy in the study of depth perception is the result of the early development of the stereoscope (Wheatstone, 1838). The stereoscope allowed depth perception to be studied through manipulation of pictorial images rather than through solid objects. Attention in depth perception became focused on stereoscopy because, from a technological point of view, it had a head start compared to other depth cues. The technology to manipulate cues such as movement parallax and shading easily and in a controlled manner, did not become available until the advent of computer graphics.

The most important categorisation to be made from a user-friendliness point of view in terms of mobility and communication is that into so-called autostereoscopic and

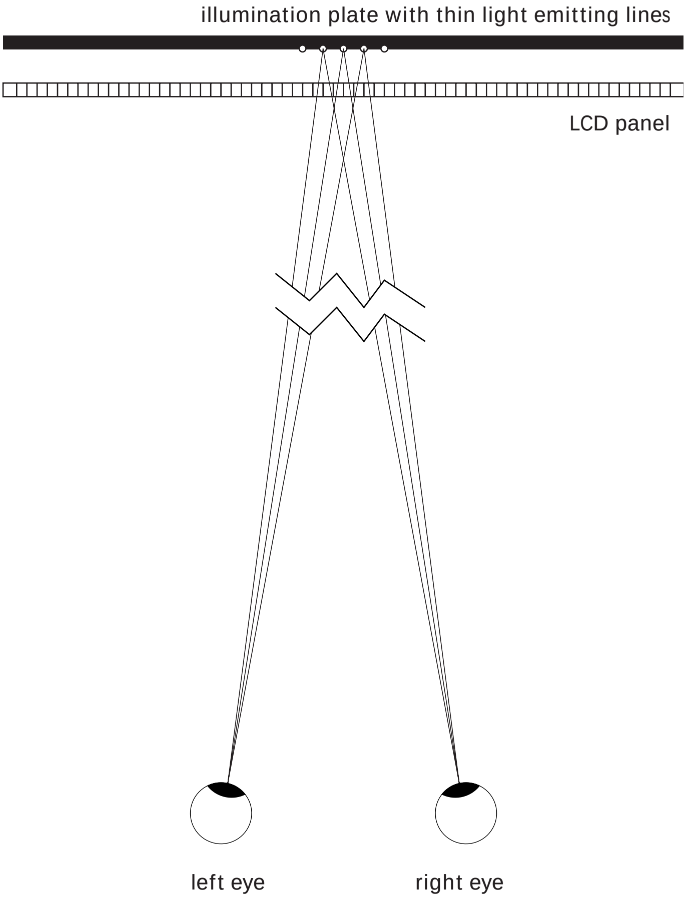
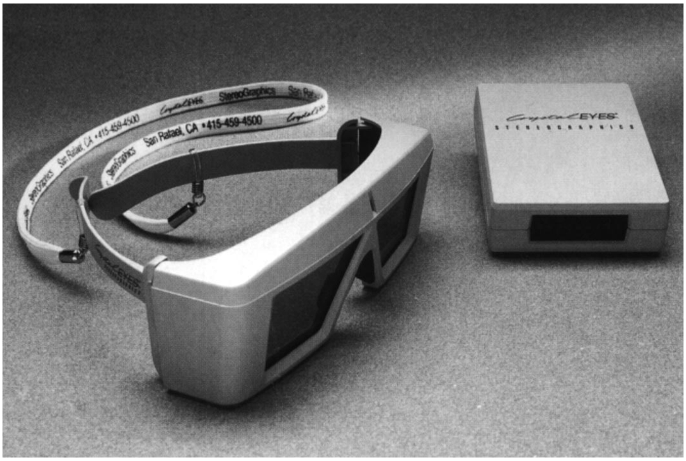


Figure 2.1
Schematic top view of the DTI-100M autostereoscopic display by Dimension Technologies. An ordinary LCD panel is placed several millimeters in front of an illumination plate with thin vertical light emitting lines. For the left eye the odd columns of the LCD panel are back lit while the even columns are not. For the right eye the opposite is true. The image intended for the left eye is therefore shown on the odd columns while the one for the right eye is shown on the even columns (Eichenlaub, 1990).

non-autostereoscopic systems. Autostereoscopic systems do not require the user to wear special viewing devices in order to see a stereoscopic image. For the moment autostereoscopic systems appear to stay inside the research laboratories and do not find their way on to the market. One reason for this may be that for some autostereoscopic systems to work the user must keep his head in one particular zone or in one of a number of different zones, as is the case with parallax barrier systems (Eichenlaub, 1990) (Figure 2.1). If the user’s left eye is in a zone intended for the right eye and vice versa a pseudostereoscopic image results in which far objects seem near and vice versa. Some implementations of autostereoscopic displays which try to prevent this situation by tracking the users head position

Figure 2.2
Active polarising glasses.



and swapping the images for the two eyes when required (Ichinose et al., 1989). Nevertheless the user's freedom of movement is limited to a zone at a certain distance from the screen and the stereoscopic impression collapses when the user moves too near or too far from the display. Another category of autostereoscopic displays based on multiplanar technology, such as the varifocal mirror (Cohen, 1979; McAllister, 1992) and the volume scanning LCD display (Kameyama et al., 1993a, 1993b), are mechanically complex due to their moving parts.

More popular are the non-autostereoscopic systems based on images presented as stereo pairs. These systems can be either of the time-parallel or the time-multiplexed variety. In a time-parallel system the monitor screen shows the images for both eyes simultaneously. In a time-multiplexed system the images are shown in rapid alternation, and only the eye for which the image is intended is presented the image. This is achieved either through active polarizing glasses (Figure 2.2) which block the eye which is not to receive the image, or through passive polarizing glasses which work together with a polarizing plate in front of the monitor which can switch the polarisation of the monitor image (Figure 2.3). From a user-friendliness point of view the passive glasses are preferable. Not containing any electronics they are lighter than active glasses and any flicker caused by switching polarisation is limited to the monitor and does not affect the real surroundings.

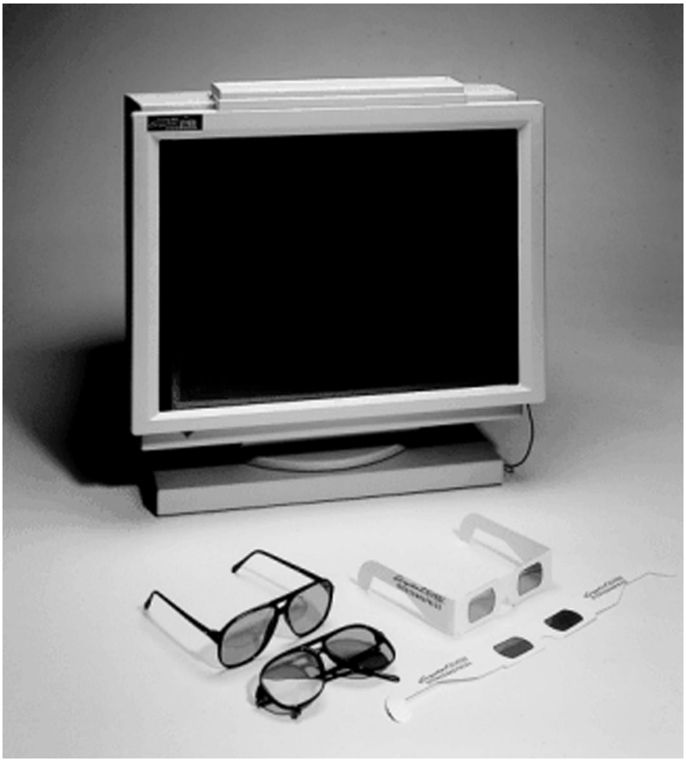


Figure 2.3
Polarising panel with passive glasses.

Based on thirteen years experience with various stereoscopic devices for molecular computer graphics, Lipscomb (1989) argues that for product acceptance in the high end markets neither price nor stereo image quality were important compared to unobtrusiveness. He illustrates this by means of contrast ratio, which increased from polarizing plate technology, via active polarizing glasses, to PLZT[1] glasses. This contrast ratio turns out to be a less important factor in product preference than unobtrusiveness as the order of increasing contrast ratio in fact reflected a decrease in user preference.

Stereoscopy is already being used in neurosurgery. Davey et al. (1994) describe how they extended an interactive stereotaxy system with active stereoscopic glasses to view images acquired through both MRI and Digital Sub-

---

[1] After the chemical elements Pb, La, Zr and Ti (Lead Lanthanum Zirconate Titanate) of which these shutter glasses are made.

traction Angiography (DSA). Worthington et al. (1985) conclude that such stereoscopic DSA images are considerably more useful than single view angiograms in surgical planning.

All pure stereoscopic systems in which the observer can move relative to the display suffer from unwanted effects under observer movement. Either the 3D impression collapses (for example in parallax barrier systems) or movement results in so-called pseudo-parallax. Pseudo-parallax is when the observer moves relative to a stereoscopic display and the virtual scene does not remain stationary but appears to move. For example, if the observer moves to the left, all elements in the scene that appear to leap out of the picture also move to the left at the same speed (Overbeeke and Stratmann, 1988). Pseudo-parallax distorts the virtual scene, and since it provides the observer with perceptual information which is in conflict with his experience, it can result in nausea, headache or eye strain. A pure stereoscopic image should thus be viewed from a position on the normal from the centre of the image.

Movement parallax based desktop VR

Many desktop VR systems make use of stereoscopy as it is often thought that stereoscopy is a necessity for depth perception. However, cue conflict studies revealed that stereopsis may be dominated by other cues, in particular motion and occlusion (Wickens, 1990). In Wickens' study motion does not only entail motion of objects within the scene but also movement parallax. Movement parallax is the phenomenon whereby the observer's head movements seem to cause objects in sight to shift relative to one another. While occlusion and motion of objects are common in commercial computer displays, movement parallax is not.

Two different, experimental movement parallax based desktop VR systems will now be discussed. The first is the Delft Virtual Window System (DVWS) (Smets et al., 1987; Overbeeke and Stratmann, 1988; Pasman, 1997a), the second is often referred to as Fish Tank VR (Ware et al., 1993; McKenna, 1992). The main technical difference between the DVWS and Fish Tank systems is the coupling method between the observer's head movement and the virtual camera (Pasman, 1997c; Voorhorst, 1998). The coupling methods for the DVWS and for the Fish Tank projection are shown in Figure 2.4. This technical difference[2] has got two consequences from an application point of view. First,

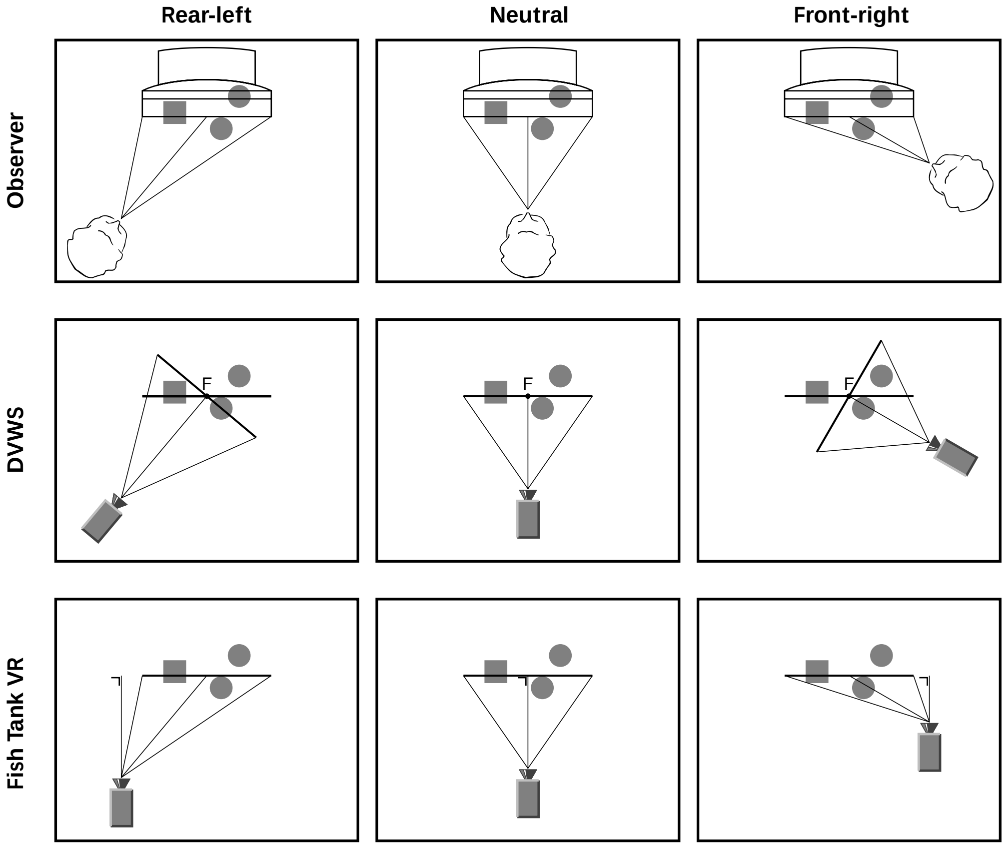


Figure 2.4
This figure shows the observer in three positions and the corresponding camera positions according to the DVWS and the Fish Tank VR coupling methods.

as the DVWS results in images in central perspective, regardless of the viewpoint of the observer, images do not appear distorted, even in case of viewpoint dislocation (See Chapter 3 for a description of a handheld computer which takes advantage of this characteristic of the DVWS). Viewpoint dislocation is when the position of the virtual camera does not accurately correspond to the head position of the observer. It results from inaccurate head position detection or from delay. Since with the DVWS the

---

[2] For details on how to implement DVWS and Fish Tank VR systems see Djajadiningrat and Gribnau (1998) and Gribnau and Djajadiningrat (1998).

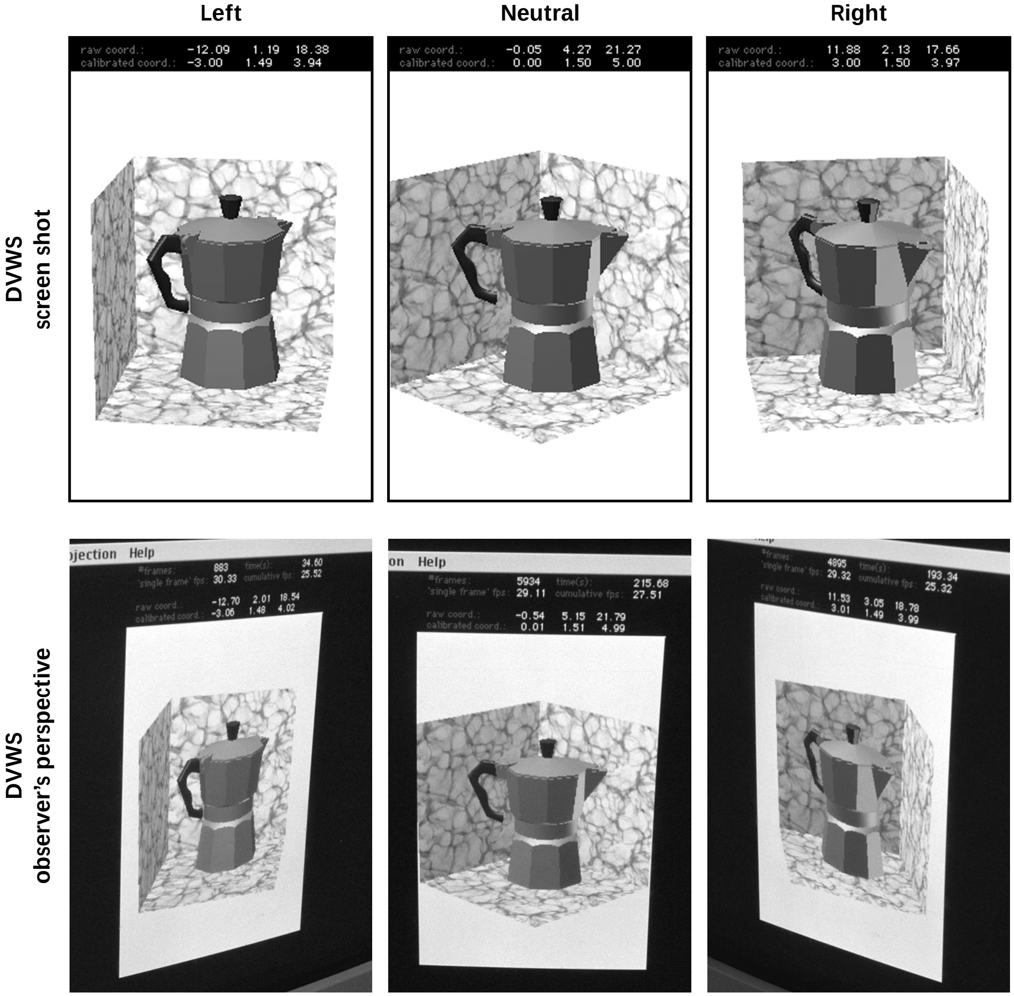
image does not appear distorted under viewpoint dislocation, it is possible to vary the couplingsfactor. The couplingsfactor is the ratio of the angle between the current camera axis and the neutral camera axis, and the angle between the line of sight and the screen normal. For example, with a x4 couplingsfactor it becomes possible for the user to look around the virtual scene completely (360°), whilst moving over 90° only. The fact that with the DVWS the images always appear in central perspective, also has consequences for passive observers. Passive observers are onlookers whose head movements do not influence the image shown on the monitor. They can be thought of as observers with large and varying viewpoint dislocation. With the DVWS, passive observers see a changing perspective of the virtual scene which does not appear distorted. With Fish Tank VR on the other hand, viewpoint dislocation does result in the active observer seeing a distorted perspective. The image appears undistorted only from the viewpoint which exactly corresponds to the virtual camera position. Passive observers thus see a distorted image. In summary, the DVWS can be thought of as less critical and more flexible than Fish Tank VR in terms of the coupling between head position and virtual camera.

The second difference between the DVWS and Fish Tank VR from an application point of view, is that with Fish Tank VR the frame of reference of the real world and that of the virtual world can coincide, even for large angles with respect to the viewing normal. What advantages does a virtual world of which the frame of reference coincides with the real world have to offer?

Unifying the display and manipulation spaces

An interesting property of a display method which makes the virtual world's frame of reference coincide with the real world, is that the display and manipulation spaces can be united. Currently, in desktop VR systems the display and manipulation spaces are separated. The users sees the virtual objects in 3D on the monitor but cannot manipulate them directly. Instead he has to make use of a separate input device. If the virtual and the real world coincide, the user could manipulate virtual objects with his hand or an instrument.

Note this cannot be achieved with a desktop VR system based on stereoscopy only. While the user's hand or instrument would be subject to parallax shifts under observer



```
raw coord.:          -12.09   1.19  18.38
calibrated coord.:   -3.00    1.49   3.94
```

```
raw coord.:          -0.05    4.27  21.27
calibrated coord.:    0.00    1.50   5.00
```

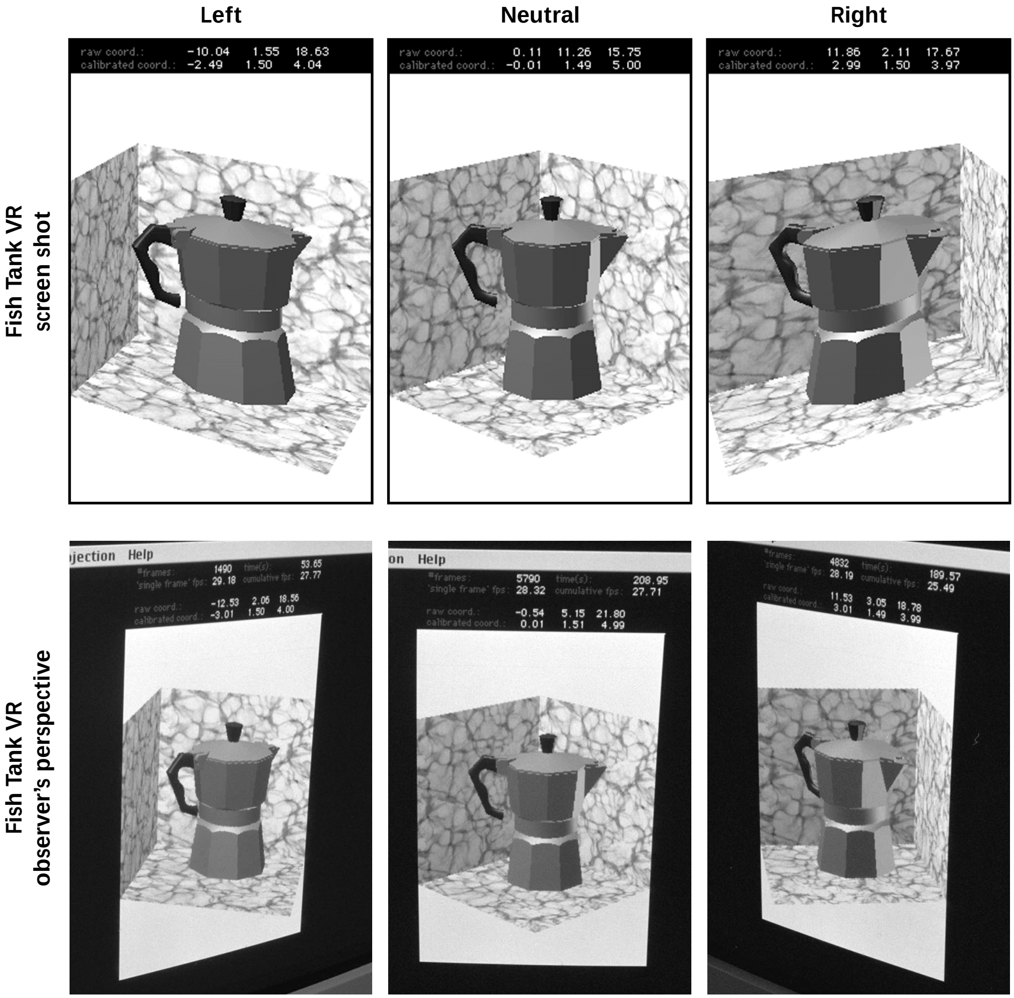
```
raw coord.:          11.88    2.13  17.66
calibrated coord.:    3.00    1.50   3.97
```

```
#frames:               883   time(s):          34.60
'single frame' fps:  30.33   cumulative fps:   25.52

raw coord.:         -12.70    2.01  18.54
calibrated coord.:  -3.06     1.48   4.02
```

```
#frames:              5934   time(s):         215.68
'single frame' fps:  29.11   cumulative fps:   27.51

raw coord.:          -0.54    5.15  21.79
calibrated coord.:    0.01    1.51   4.99
```

```
#frames:              4895   time(s):         193.34
'single frame' fps:  29.32   cumulative fps:   25.32
raw coord.:          11.53    3.05  18.78
calibrated coord.:    3.01    1.49   3.99
```

Figure 2.5
The DVWS coupling method. The top row shows screen shots for three different observer positions. The bottom row shows how these images look from the point of view of the observer. Note how the screen shots look ‘normal’: they are in central perspective. Also note how the front two edges of the background planes move relative to the screen.



```
raw coord.:          -10.04    1.55   18.63
calibrated coord.:   -2.49     1.50    4.04
```

```
raw coord.:            0.11   11.26   15.75
calibrated coord.:   -0.01     1.49    5.00
```

```
raw coord.:           11.86    2.11   17.67
calibrated coord.:    2.99     1.50    3.97
```

```
#frames:               1490  time(s):          53.65
'single frame' fps:   29.18  cumulative fps:   27.77

raw coord.:          -12.53   2.06  18.56
calibrated coord.:   -3.01    1.50   4.00
```

```
#frames:               5790  time(s):          208.95
'single frame' fps:   28.32  cumulative fps:   27.71

raw coord.:           -0.54   5.15  21.80
calibrated coord.:     0.01   1.51   4.99
```

```
#frames:               4832  time(s):
'single frame' fps:   28.19  cumulative fps:   189.57
                                               25.49
raw coord.:
calibrated coord.:    11.53   3.05  18.78
                       3.01   1.49   3.99
```

Figure 2.6
The Fish Tank VR coupling method. The top row shows screen shots for three different observer positions. The bottom row shows how these images look from the point of view of the observer. Note how the screen shots look distorted. Only from the point of view of the active observer do the images look 'normal'. Also note how the front two edges of the background planes stay attached to the screen, regardless of the point of view of the observer. This is the characteristic of the Fish Tank VR method that makes unification of the display and manipulation spaces possible.

movement, for the virtual objects this would not be the case. Consequently, virtual objects will appear to move with respect to the user's stationary hand under head movement if the user keeps his hand stationary. I will further address this subject in Chapters 6 and further, and in Chapter 9 in particular.

Combining stereoscopy and movement parallax

In the preceding discussion of stereoscopy and movement parallax based displays the two were treated separately. In the literature most displays based on the DVWS are indeed based on movement parallax only, with Suetens et al. (1988) being an exception. They used the DVWS in combination with stereoscopy to view wireframe 3D models based on CT data. However, Fish Tank VR based displays often use stereoscopy in addition to movement parallax. Exceptions are Fisher (1982) and Diamond et al. (1982) whose systems are based on movement parallax only.

Choosing for movement parallax

In the preceding paragraphs I have argued why to use movement parallax rather than stereoscopy. But why not use both of them together? In the remainder of this thesis I will use movement parallax exclusively rather than use both stereoscopy and movement parallax. The reasons for this are threefold. Stereoscopy is computationally expensive. That is, perspectives for two eyes need to be calculated. As in 3D systems for medical use the data sets are complex this is a burden which is not to be underestimated. Second, it requires additional hardware. Not only the headware needs to be different but also the graphics hardware, as twice the update rate is required. Third, stereoscopy is of no use for approximately 12% of the male population who are stereoscopically blind.

In the next chapter...

To improve the user-friendliness of a medical 3D system it should become less of a computer and more of a product. By pushing the computer into the background the user need not be aware that he is operating a computer. In Chapter 3 product interface design is discussed. J.J. Gibson's theory of affordances is considered as a framework for making electronic consumer products intuitive to use.

3

Gibson's Theory of Affordances

A Framework for Design

Summary

To make a medical 3D system intuitive it should be less a computer and more a product. The computer needs to be hidden from the user and a product tailored to the user's natural behaviour needs to surface.

In keeping with this approach, in this chapter I concentrate on human-product interaction (HPI) rather than human-computer interaction. I argue the value of Gibson's theory of affordances for Industrial Design Engineering, with an emphasis on formgiving and interaction. The objective is to investigate whether, and if so how, industrial designers can make use of the theory of affordances on a practical level to improve the usability of products. I start with why human-product interaction has become a pressing issue in product design in recent years, illustrated along the hand of developments in photographic cameras. Next I explain affordances and how they fit within the theory of direct perception. The relevance of affordances for solving the aforementioned interface issues is described. I then give an overview of established 'good practice' in HPI in order to clarify what affordances have to offer in addition to existing practice in HPI. Particular attention is given to the field of product semantics, which I consider to be part of established practice in HPI. I then compare product semantics to affordances in terms of what they mean on a practical, designers level rather than by the theories from which they originated. Through this comparison I attempt to highlight the benefits and shortcomings of these two approaches for industrial designers and how they might complement each other. Finally, I try to show by means of an example how an affordance conscious design approach can differ from existing good practice and how it can improve human-product interaction.

Human-product interface design: a pressing issue

An important aspect of product design is to make clear to the user how to operate a product. No matter how well the product performs from a technical point of view, its technical functionality is limited to that which the user can actually access. For example, Van Nes and Van Itegem (1990) show how users are not aware of much of the functionality which earns a particular car radio the label 'advanced'. In the past the mechanical workings of a product largely dictated the overall form and the positioning of the controls. The advent of microelectronics and miniaturised mechatronic components, did not only enable designers and engineers to create products that were smaller than their mechanical counterparts, but also gave them much more freedom for the overall form and layout of components and controls.

Take for example the way in which photographic cameras developed. The form of a mechanical camera was largely dictated by the mechanical transport of the film, a lens which needed to be perpendicular to the film plane and the path from the lens to the viewfinder. The viewfinder needed to point in the same direction as the lens in the case of a 'view through' camera or needed to be connected with a minimum of prisms and mirrors in the case of a single lens reflex camera. With distance, aperture and shutter speed controls, early mechanical cameras featured only a minimum of functions (Figure 3.1).

Mechanical camera
(Leica M-series)

'Electronified' camera
(Minolta 7000)

Digital camera
(Sony DSC-F1)

Figure 3.1
Development of photographic cameras.

In the eighties these mechanical cameras became increasingly electronified and acquired more functions. However, the camera's overall size and layout did not change much as these remained dictated by mechanical components (Figure 3.1).

With the advent of modern digital cameras, of which the most important components were a recording CCD, a solid state storage device and a LCD screen as a view finder, constraints on the form became far less strict as connections between the components were electrical rather than mechanical (Figure 3.1). With the digital camera the number of functions increased once again, not only when compared to the mechanical camera but also when compared to the electronically controlled camera. Added features included in-camera image manipulation and computer, printer and television interfaces. Though early digital cameras were similar in size to their mechanically based ancestors, they paved the way for smaller products as their electronic components were easier to miniaturise than mechanical components such as film transport.

In the process from fully mechanical to fully digital the camera gave less and less auditive and haptic feedback. The introduction of electronic zoom and shutter release button impaired the feel of the camera. With the digital camera even the sound of film advancement after each exposure disappeared.

Many other consumer products went through or are going through the same process as the photographic camera described above. Though certain controls could be eliminated through automation, this did not weigh up against the controls which spawned from additional functionality. More functions needed to be packed in a smaller housing which offered less room for controls. Feedback was reduced as electronic parts - unlike mechanical components - do not provide meaningful auditive feedback and because electronic controls offer less haptic feedback than mechanical ones. Moreover, interface designs became increasingly modal: depending upon which mode the device was in, a control had a different function. As the 'functions to controls ratio' increased descriptive labelling of controls became difficult. The feedback which disappeared from the controls themselves was partly taken over by electronic displays. Thus human-product interaction became more like human-computer interaction.

Increased functionality and miniaturisation increased the importance of user interface design. Also, as the production quality of most current products is beyond doubt and products are difficult to distinguish on their technical merits, user interface design is becoming more of a sales argument.

HPI would benefit from a theoretical framework to tackle the aforementioned issues. The theory of affordances can act as such a framework. In the next section I describe how affordances fit within the theory of direct perception and how they are interpreted in this thesis.

Affordances and the theory of direct perception

Affordances form part of the theory of direct perception, also known as the ecological approach. This theory of perception was started by J.J. Gibson (1904-1979). Central to this perception theory is the reciprocal relationship between animal and environment. Having gone through evolution together animal and environment are thought of as inseparable, with one implying the other. As in industrial design we generally design for humans, in this chapter I will talk of man and his environment, rather than animal and environment.

Gibson thought up the neologism ‘affordance’ as a noun to complement the verb ‘afford’. “The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill” (Gibson, 1986). In keeping with the notion of the inseparability of the man-environment system affordances can only be thought of in terms of such a system as a whole, not just in terms of man only or the environment only. It may be clear that affordances cannot be thought of as separate from the environment. Yet it may be less obvious that neither can affordances be thought of as being separate from man. There are two main reasons why affordances of products cannot be seen as separate from man. First, affordances have to be specified in terms of information which a human individual’s perceptual system is capable of detecting. Second, properties which specify an affordance are not meaningful when measured in physical terms but only when specified relative to a human individual. What the environment affords depends on whether the human is a child or an adult (body scale), frail or in perfect health (body condition). To a child the railings of a fence may afford squeezing through, while to an adult they do not. To a fit person a puddle may afford jumping over while to a granny it may not. To capture the capabilities and limits of both someone's perceptual systems and his motor systems in one conception we can speak of someone's perceptual-motor skills.

However, affordances are not only inextricably linked to the behaviour a human is capable of, but also to his intentions. The total of these potential purposeful behaviours are called a human's effectivities (Shaw andand MacIntyre, 1974; after Von Neumann, 1966).

Another idea which is essential to the theory of affordances is the inseparability of perception and action. Perception is seen as written in the language of actions. Structured energy only forms information for a human if that human can act on it. In other words, perception is only of use if it leads to appropriate action. Likewise, action can only be successful if it is guided by appropriate perception.

Furthermore, affordances are more than just properties perceived which specify possible or permitted actions. Affordances also specify the details of those actions (Michaels and Carello, 1981). A playing ball affords throwing but different types of balls afford throwing in different ways.

What can the theory of affordances contribute to HPI? In order to answer that question, we first need to look at current 'good practice' in interface design. Only then can we say which parts of current HPI are consolidated by the theory of affordances and what news affordances could bring.

Current 'good practice' in interface design

What follows is a list of points which are generally considered 'good practice' in HPI. It should be viewed as points that interface designers are aware of. 'Good practice' does not mean that these points are always considered in practice. Many products on the market violate one or more of these points. During the design process these point of ‘good practice’ may get sacrificed as compromises are made.

Furthermore, this list focuses mainly on interface design of electronic consumer products. Therefore, much attention is given to controls and displays. This list does not address such issues as dimensions and layout of interior spaces, despite the fact that usability of consumer products (product design) is influenced by the way they are integrated into our environment (interior design and architecture).

This overview does not pretend completeness. I feel, however, that it is necessary to give some ideas about what is generally considered ‘good practice’ in interface design, as often designers discard affordances as ‘old hat under a new name’. Having composed a list of ingredients of ‘good practice’ I will use it to dispel the view that the benefits the theory of affordance can bring to interface design are already present in a product in which the current ideas about HPI are well implemented.

Anthropometrics

The field of anthropometrics is concerned with the measurements of the human body. Measurements which can be used in the optimisation of products and our environment. The field of anthropometrics therefore forms an essential ingredient for the improvement of product usability.

When designing a product for usability it is a prerequisite that the product’s size and the sizes and spacing of its controls are in keeping with the dimensions of the human body. An example of how neglect of anthropometrics may negatively effect usability are the miniature buttons of the calculator on a digital watch as their physical size is not tuned to the size of a human finger.

Expressing the purpose of a product

Expressing the purpose of a product as a whole is the first step to intuitive operation. Once it is clear to the user what kind of product he is dealing with, his pattern of expectation with regard to the purpose of the controls is adapted to the purpose of that product. For example, if a user does not recognise an escape hatch, or fire extinguisher, or any piece of safety equipment for what it is, the device may never invite intuitive use no matter how well designed its controls. A less life threatening but nevertheless still very annoying example is when the new

Figure 3.2
Highly similar looking audio components (Bang & Olufsen Beosystem 5500).

owner of a set of audio equipment tries to operate the CD player, only to find out that the component he is dealing with is not the CD player but the highly similar looking pre-amplifier (Figure 3.2). Often components are designed to look identical. According to Dondis (1973): “Repetition is the uninterrupted visual connections that are particularly important to any unitised visual statement”. From a usability point of view, repetition is an easy way of achieving an aesthetically pleasing whole, for which the expression of the individual components is sacrificed.

Grouping of controls

Thoughtful product design does not only consider anthropometric aspects of controls, but also the way controls are grouped. By clustering controls according to their functionality or to the way in which they are used, an interface can be made more clear-cut. Through the drawing of arrows on a control panel Mayall (1968) shows how a visually well ordered arrangement of controls turns out to be thoroughly disordered in use (Figure 3.3).

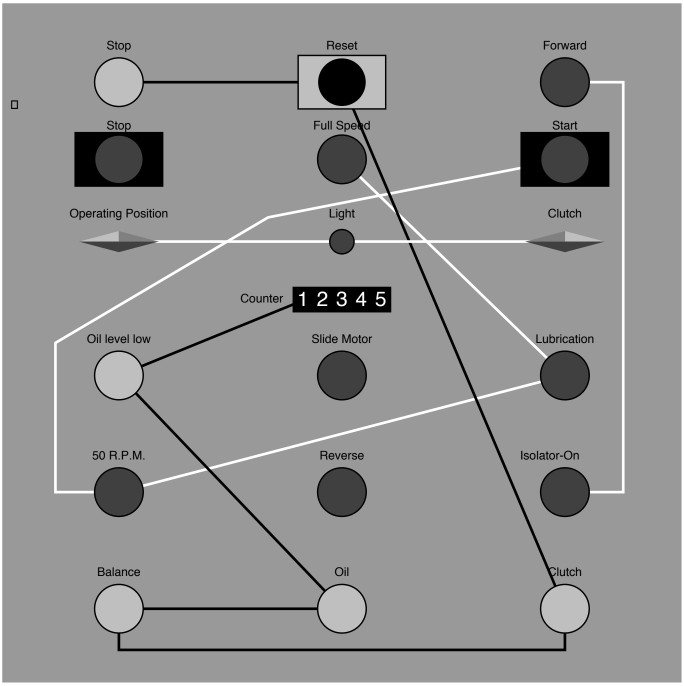


Figure 3.3
Visually well-ordered, disordered in use (after Mayall (1968), page 53).

Grouping should also take into consideration limitations of the human anatomy. Frequently used controls should demand a minimum of physical strain.

Importance of controls

Indicating a hierarchy of importance in controls can improve usability. The importance of a control may be expressed through form, colour, texture and material. Note that the importance of a control is not necessarily related to frequency of use. The recording button on audio/video equipment is given special status not because it is used the most often but because it may cause an irreversible process.

Differentiation between controls

Controls should distinguish themselves from each other. This can be achieved through form, colour, texture and material. For cost or aesthetic reasons controls often violate this principle, thus becoming difficult to tell apart except for through their labelling. As with the similar looking system components under ‘expressing the purpose of a product’, here too repetition is used to create unity. Products result which at first appear to be clear-cut through their rigorously organised control panels but which often turn out to be still confusing when it comes to activating a particular function within a particular group.

Controls which express what action operates them

Controls should express how they need to be operated. It should be clear that a particular control needs to be rotated and not pressed, or that it needs to be slid and not tumbled. If controls are formed the same, the user will expect that their mechanical behaviour is identical. An example of a device with such controls is that of an amplifier of which the front panel shows three identical looking rotary controls. Yet the volume control allows continuous rotation over 270 degrees increasing clockwise, the balance control allows continuous rotation over 180 degrees symmetrical to the neutral vertical position, while the input selector offers rotation in discrete, notching steps only. If the actual mechanical behaviour of the control is different from the control's behaviour as expected by the user, fluid interaction is hampered.

Controls which express how the action is to be carried out

Once a control expresses what action is needed to operate it, i.e. pushing, sliding, turning etc., for fluid interaction it is also necessary to consider the expression of how that action is to be carried out. For example, both the on/off button of a radio and the emergency stop of a lathe may afford pushing, yet the way in which they should be pushed are completely different. A control needs to express how much bodily interaction is to be involved: finger tip, one finger, multiple fingers, a whole hand etc. If a control expresses the wrong 'how' it may either not be activated or suffer excessive wear or even damage. If controls are optimised for a certain hand position then the device needs to express this, otherwise it may never be used in the intended, optimal fashion.

Fitting the control to the nature of the variable

A control should fit the nature of the variable which is to be adjusted. As a negative example, in many electronic products continuous variables (sound volume, time, temperature) are adjusted in discrete steps by means of up/down buttons. Such buttons reflect neither the continuous nature of the variable, nor the fact that there are limits to adjustment of the variable. Up/down buttons result either in waiting to achieve the right setting or in overshoot. To compensate for these shortcoming in comparison with, for example, a marked rotary control the up/down buttons need to be accompanied by some kind of electronic display.

Mapping

The way controls are laid out can have a great influence on how easy it is to operate a product. One form of mapping deals with how the controls map to what is being controlled. Norman (1988) gives the example of a four ring cooker and how the spatial relationship between the controls and the rings influences the cognitive burden on the user. While in all of the concepts presented the same controls and rings are used, because of the layout they differ significantly in how much knowledge the user needs to operate the correct control.

Another form of mapping considers how the movement of a control maps to movements on a display. When it is not possible to achieve natural mapping between control and display, considering their relative placement may help in minimizing error. An example is mapping of a rotary control to a pointer on a linear vertical scale. Brebner and Sandow (1976) and Petropoulos and Brebner (1981) show how the relative positions of the control and the display significantly influence consensus among subjects on how they expect the direction of rotation to influence pointer movement.

Feedback

In order to give users a sense of control a product should give feedback clearly indicating its current state or the execution of a process. Feedback gives the user information about the result of the user’s actions. Without such feedback the user remains in doubt whether the device is responding or not. As was noted in the example of the development of electronic cameras the use of more electronic components and less mechanical ones often leads to impoverished feedback. As Norman (1992) says: “Mechanical devices are often visible and audible, conveying considerable information about their operation, even to those who know nothing of mechanics. The designers do not have to provide feedback to the users. The very nature of the machine guarantees that.”. In addition to the visual and auditive feedback mentioned, mechanical objects offer natural haptic feedback. Natural haptic feedback offered by mechanical buttons is eliminated when electronic touch controls are used instead. Products which do not give any feedback regarding their current configuration or state are rare. There are, however, many products which give very little feedback. With video recorders and CD players for example it may take studying of the display to see whether a tape or disc is in the machine. It is not that it is impossible to see whether a tape or disk is present, it is just that it is made unnecessarily difficult as it is often indicated by tiny characters or icons on an electronic display.

Expressing the purpose of a control and making the result perceivable

Even though a product may satisfy the previous points of ‘good practice’ with regard to its controls, the controls may still not express what they are for and what operating them leads to. A control may have an anthropometrically correct size, be placed in the correct group, express its relevance relative to others controls, differentiate itself from other controls, express what action is required and how this action is to be carried out. It may also fit the nature of the variable it controls, be mapped correctly and offer feedback. Yet the control may fail to express what its purpose is. Clearly, this expression is essential for usability as it allows the user to pick out the control that is required to fulfil the task he has in mind. Expressing the purpose of a control is highly related to the previously mentioned ‘differentiation between controls’. In order to express its purpose, a control needs to differentiate itself from other controls. Still, merely being different is not sufficient, the control needs to express its specific purpose. In other words, it needs to communicate to the user what the result of operating it will be.

Product semantics

Product semantics is a design movement which attracted much interest in the second half of the eighties. In this section, I first turn to the theoretical background behind product semantics. Second, I give some concrete examples, and point out the shortcomings and potential of product semantics with a view to usability. Third, some design methodologies which resulted from product semantics are described.

Theoretical background of product semantics

Krippendorff and Butter (1984) say: “product semantics is the study of the symbolic qualities of man-made forms in the context of their use and the application of this knowledge to industrial design”. They present product semantics as a design theory which draws upon a mixture of semiotics and information theory.

In semiotics the central concept is the sign. Semiotics revolves around how meanings match or differ as these signs are created and exchanged between people. The sign is intended to represent something, though what is actually expressed to a user may be different. Different instead of matching meanings are not necessarily seen as undesirable 'miscommunication' but more as a potential 'richness', which shows important sociocultural differences between people (Byrne, 1990). Examples of semiotic analysis applicable to industrial design are Barthes' (1972) essays 'The New Citroën' and 'Toys'.

In accordance with information theory, in product semantics the designer is viewed as a communicator of a message in the form of a product and the user as the receiver of that message. The designer can only send the message once while the user reinterprets the message again and again through its use. The encoding of the meanings of a product is influenced by the designers competence and vocabulary, his socio-cultural background, his intentions and condition. The decoding of the meaning of a product on the side of the user is influenced in a similar fashion. Technical or economic compromises in the production process introduce 'noise' as the message is sent from designer to user.

The use of metaphor examined from a usability point of view

At a concrete level, product semantics often seems to work via metaphors and association. This is particularly well illustrated by conceptual work created at the Cranbrook School of Art and Design, where product semantics was the central theme in the eighties (Aldersey-Williams et al., 1990). Products resulted which tried to convey meaning through the use of metaphor. Metaphor can be used in various ways. First, it can be used to draw upon an existing language of forms from another field of interest, in order to position the product as being related to that field. Second, metaphor may go beyond a merely visual relationship, to express a relationship in functionality with an existing product or concept. Third, a product may be associated with an existing object or concept, natural or man-made, to convey that it has characteristics present in that object or concept. I will now give some examples of these forms of metaphor.

An example of a product which borrows from existing design languages is the audio receiver shown in Figure 3.4. It draws upon forms which are used for musical notation and for traditional musical instruments. This does not mean that the receiver can be used for musical notation or that it is an instrument. The design of the receiver draws upon existing visual languages to express that it fits within

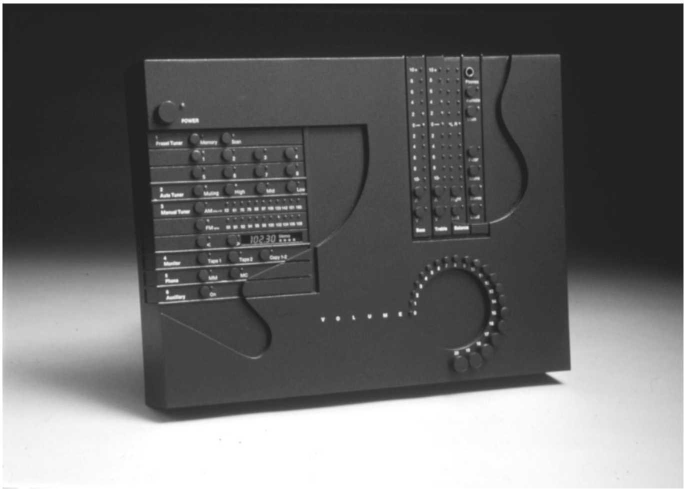


Figure 3.4
Concept for audio receiver drawing upon forms in musical notation and of traditional instruments (design: Robert Nakata, 1985).

a category of objects which have to do with music. While the metaphors of musical notation and traditional musical instruments have given rise to a visually interesting way of organizing the receiver's control surface, they have offered very little help in distinguishing the controls themselves. Apart from the volume control, which has been given a clear direction by the increasing size of the buttons, all the controls are buttons identical in form, texture and colour.

Some forms of product semantics may give rise to a new form language, yet do not help to clarify what the product is for. This may be the case when Uri Friedländer (1984) states: "... we now face the problem that metaphors are difficult to relate to shapes that are not sensed by most consumers: The abstraction of the electrons. Thus, we are forced to use the mound which houses these particles for the creation of a new metaphor: the chip. Sharp edges, dark boxes, sensitively detailed large flat areas interrupted by slim lines: These are the new abstract metaphors of advanced technology." While developing such a design language may be useful in breaking with a particular aesthetic tradition, it does not help the user in distinguishing different electronically based products and giving these an identity of their own.

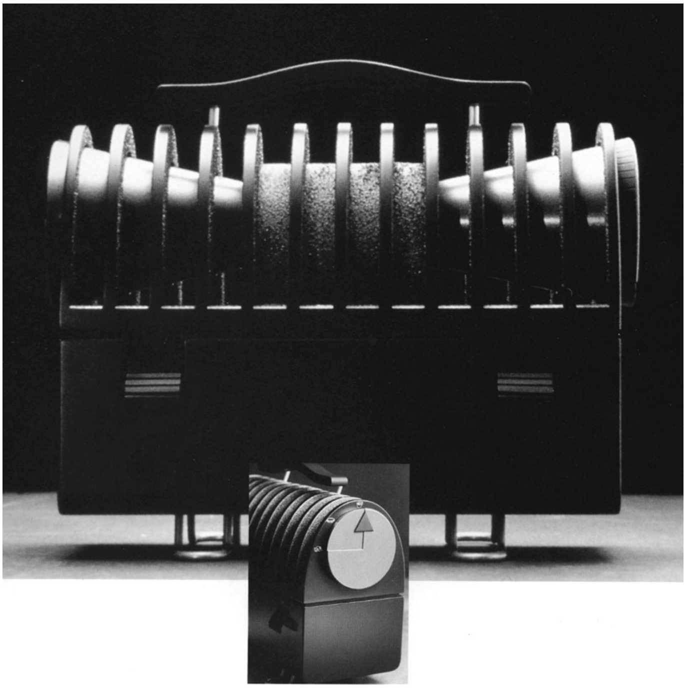


Figure 3.5
Concept for portable microwave oven reminiscing an American workman’s lunchpail (design: Paul Montgomery, 1986).

An example of the use of metaphor to express a relationship in functionality is shown in Figure 3.5. It is a conceptual model for a portable microwave oven which is form given as a traditional workman’s lunchpail. Styling cues hint that since there is a relationship in form, the user can expect functionality which is highly related. Both are food containers used during lunch, and both are portable.

I now turn to examples of products which use metaphor to express that they have characteristics in common with another object or concept, natural or man-made. For example, the curvi-linear forms of Art Nouveau lamps and fences draw upon those of flowers and trees. Likewise, the organic forms of some cars might be associated with the muscles of feline predators. However, generally this type of metaphor is used to shape the expressivity of the product as a whole, not to improve usability. Potentially though, this subtle application of metaphor could be used to improve usability. For example, a control could express that it needs to be treated with gentle care, like an egg. This requires much sensitivity on the part of the designer. It is not possible to use a literal copy of the egg as a control, since generally the control will need to express other characteristics than vulnerability. Therefore only the aspects of an egg which express vulnerability need to be copied to the control. What is it that makes an egg vulnerable? Is it its form, or perhaps its crackle texture? Or perhaps an egg does not express vulnerability at all, we may simply have learned that it is vulnerable, in which case another association is needed.

There are a number of drawbacks to the metaphoric approach of form semantics. Gentner and Nielsen (1996) name three problems with functional metaphors which compare a novel product to an existing product. They illustrate these problems by means of the ‘a word processor is like a type writer’ metaphor. The first problem is that at some point the metaphor will break down. The newly designed product may have more functions than the concept it is associated with has to offer. Based on the aforementioned metaphor the user of a word processor would never look for the replace command. The metaphor stresses the similarities of an electronic product with a traditional product, rather than its innovative qualities. The second problem is that the new product may not (yet) have the features of that which it is compared to. While a type writer allows you to mark up any piece of paper you get in the mail, a word processor by itself does not offer that functionality. The third problem is that some features exist in both the new product and the old one which it is compared to yet work completely differently. In the case of the word processor-type writer metaphor such features include tabs and line feeds.

Gentner and Nielsen also give an example of an interface which faithfully emulates the interface of an earlier technology. The Phelps tractor (Clymer, 1950) is a steam-engined vehicle from 1901, which is controlled by reins (Figure 3.6). It thus draws upon a metaphor with the interface for the familiar horse. Drawing the reins causes the vehicle to brake until it reaches a standstill. When the user continues to draw the reins at standstill, the vehicle will back up. Giving free rein will cause the vehicle to accelerate.

Summarising, if the metaphor is made highly concrete then its explanatory quality will be limited, as users will only expect the functionality of the object or concept it is compared to. Also, a highly concrete metaphor based on

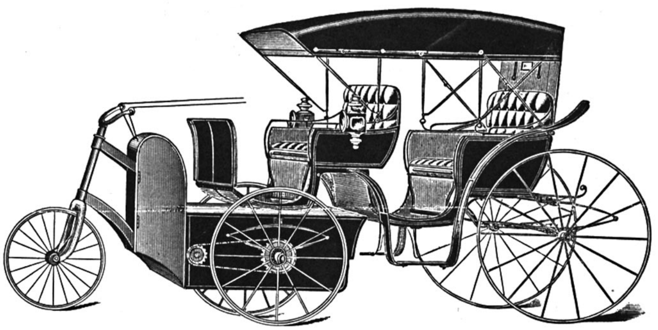


Figure 3.6
The Phelps tractor, a steam powered vehicle from 1901 which is controlled through reins.

earlier technology, as in the Phelps tractor, may not lead to the most appropriate interface for the new product. On the other hand, if the metaphor is very much abstracted, users may not catch the reference and cannot draw upon their experience with the obscurely referenced object.

Gaver (personal communication) points out that collapse of the metaphor is unavoidable and a natural part of the maturation of a new product. It is not until the metaphor dies that a thing gets meaning for itself.

Design methodologies based on product semantics

There are several examples of design methodologies based on product semantics. Friedländer (1989) illustrates the design process of an espresso-machine, which emphasizes the ceremonial values of coffee making. By studying existing objects with ritual or religious values he discovers recurring elements in those objects which he reuses in the design of the espresso-machine.

Lannoch and Lannoch (1987) developed a method which they call 'semantic transfer'. In this method the desirable characteristics of a product are first expressed in words. These words are then considered in all their possible contexts. Through this process the designer can build up a range of possible associations to arrive at a physical artifact.

Byrne (1990) gives an example of a working method based on semantics for the design of a brandmark. Through analysis of the facts about the company, its purpose and the goals of its brandmark redesign, a denotative word list is built. Visualisations of these words are put in a matrix in which they are narrowed down through combination after which a brandmark appears. The outcomes of several of the matrices are tested against a connotative word list forming a semantic differential scale.

The use of affordances to human-product interfacing

When comparing the description of affordances to the list of good practice in interface design it becomes evident that affordances form an elegant frame work to unify apparently disparate elements of HPI. The need to consider the interaction of the acting-perceiving human and his environment is reflected in the anthropometrist's concern with body dimensions and the ergonomist's interest in the relationship between human abilities and the design of the environment. Objects may be too large or too heavy to invite certain actions. Two vessels may be identical in form and colour, yet only one invites drinking from as the other is too large to hold, making it a vase.

Showing the possible behaviour the user can enter into is improved by clear expression of a product's functionality and what actions are required by its controls. A control which not only shows the required action but also how that action is to be carried out, is in keeping with Michaels' and Carello's (1981) assertion that affordances specify both the global aspects and the details of an action. The haptic feedback which is naturally present in mechanical components but so often lacking in electronic products was emphasised by Gibson (1986) in his example of the affordances of a pair of scissors in which "one can actually feel the cutting action of the blades".

Affordances may offer interface design another interesting opportunity for reflection in that Gibson's definition ends with 'for good or ill'. When thinking in terms of human-product interface design this leads to a distinction between affordances which invite effective action leading to the result desired by the user, and affordances which invite action leading to no results or different results. I will return to the first category in the example given at the end of this chapter. Examples of the second category are the volume and contrast controls of the Apple Emate (Figure 3.7). To a first time user these appear to be toggling push buttons with an increase and a decrease side. It turns out though that they are not push buttons but sliders. However, today there may be a third variety of products

Figure 3.7
The volume and contrast controls of the Apple Emate afford pushing while in fact they are sliders. In the photograph on the left, the right-hand slider has been slid to the left. In the photograph on the right, the right-hand slider has been slid to the right.

which afford very little: these products do not invite any action. Devoid from meaning they only cause puzzlement. Many videorecorders belong in this category. Groups of controls, identical in size, colour, form, texture and material, are used for a wide variety of different functions. While the controls in themselves may have certain degree of affordance, as a group they do not invite any action, because each control affords the same: the controls do not distinguish themselves in terms of affordance. Thus each control may have a clear ‘pushability’ affordance, yet this does not help the user to make a decision. Since the user cannot differentiate between the controls on the basis of their formgiving or positioning, he is forced to read the labelling (Figure 3.8). Note that while generally an affordance conscious designer will intentionally build-in affordances leading to desired behaviour and avoid affordances leading to non-desired behaviour, there may be some exceptions. For example, intentionally creating misleading affordances is everyday practice for designers of games and trick items.

Affordances vs. Product Semantics

In the list of points of established ‘good practice’ I included product semantics. While affordances form an elegant framework for inclusion of the other points in the list, the relationship between product semantics and affordances is more complex. Is product semantics really a subset of affordances or vice versa, do they partly overlap and partly differ, or are they one and the same thing? As may be clear from their respective explanations product

Figure 3.8
Many of the controls of this videorecorder (Blaupunkt RTV-910 Hi-Fi) are identical in size, form, colour, material and texture. It is impossible to tell them apart without reading their labelling. Since each control affords exactly the same (slidability), in a group the controls afford very little.

semantics and affordances stem from very different backgrounds. Product semantics draws upon linguistic and communication theories, while affordances has its roots in the ecological perception theory. Despite these different backgrounds, at the concrete level of designing artifacts it is quite difficult to point out the differences between product semantics and affordances.

As was pointed out in the description of product semantics various design methodologies have resulted from it. There are few examples of design methodologies which take affordances as their basis. Smets et al. (1994) describe three design exercises in which students of industrial design engineering have to design a walkman, a dessert packaging and a sculpture to fit a particular piece of music, the taste of a particular dessert and an artificial scent respectively. Matching experiments show that selected designs successfully reflect another sensory modality and can thus express higher order variables. The article mentions that students were discouraged from relying on cultural cliches by bypassing verbal descriptions and to rely exclusively on the experience of the music, taste or scent. This makes it difficult to identify the methods the students used in achieving their designs.

Table of comparison for Product Semantics and Affordances

The table of comparison shown below (Table 3.1) highlights the main differences and overlap between product semantics and affordances.

However, there is some cross-over in this respect. It would be an oversimplification to state that in product semantics meaningful actions are ignored. One of the dimensions in Lannoch and Lannoch (1987) semantic space is the 'possible actions' dimension (Händlungsmöglichkeitsdimension), which exists next to other dimensions such as a value/convention and relation dimension.

Table 3.1 Table of comparison for product semantics and affordances

Product semantics	Affordances
Behaviour is influenced by language and other learned signs and symbols.	Behaviour is immediately influenced by the environment.
Designers can draw upon metaphor. The meaning of the product is the total of all its contexts.	Designers should not draw upon metaphor or imagery through verbal association. Designers should 'trust' their senses.
The symbolic meanings and values of a product are central.	The perception of meaningful actions is central.

Likewise, elements typical of form semantics, such as product values and meanings, are indeed considered by exponents of affordances. As Smets (1995) puts it: "Form semantics have to do with the manner in which those affordances are expressed in form design. All chairs afford seating, but not all of them are thrones".

Affordances, product semantics and electronic products

Both proponents of product semantics (Lannoch and Lannoch, 1983; Scheuer, 1989) and affordances (Smets, 1995) introduce their articles with the problems posed by electronic products and how their particular theory may contribute to solving those problems. It is disappointing that the positive examples which are given, generally do not deal with expressing the purpose of controls of electronic products. The examples which are given are for what Norman calls surface artifacts rather than internal artifacts. With a surface artifact what is perceived is all there is. With an internal artifact part of the information cannot be perceived as it resides inside the artifact, often in a for humans non-readable form. Internal artifacts therefore need interfaces, so that the internal representations are transformed in a for humans accessible format.

Theorists from both camps, designers and users, will no doubt agree on the need to make controls express their purpose. When shown the examples of surface artifacts they are likely to agree that adopting a design philosophy based on products semantics or affordances is of benefit to those products. However, examples of the application of products semantics or affordances to the real problem of interfaces for electronic, internal artifacts are very hard to come by.

Yet the crux of the problem in human-product interaction is expressing the purpose of controls in electronic products, the need to make the *results* of an action perceivable. The overall formgiving of a product may specify the context and give the user an idea of the product's functionality. That product's controls may express perfectly well what kind of action they require and the details of that action. But still the product may fail to express to the user what activating a control leads to.

In accordance with this view, improvement of product usability may be seen as the implementation of expressivity — showing what the results of action will be — over the full hierarchy of user-product interaction. This hierarchy ranges from showing what the product will do, via what the major controls will do, through ever decreasing levels of importance to clarifying the effects the most minor controls will cause.

Product semantics became stuck after the first level: expressing the functionality of the product. Because the theory of affordances focuses on invited action it raises awareness of the expressive shortcomings of the rest of the hierarchy. It is in this raising of awareness that the value of affordances for industrial design lies. It does not provide the designer with a methodology to actually implement expressivity and to make the results of an action perceivable.

Figure 3.9
Control for car seat adjustment in a Mercedes (Norman, 1988).

A step towards expressivity: making the result of an action perceivable

How hard it is to implement this expressivity which allows a user to foresee the result of his actions depends upon the nature of the controlled variable. If the variable to be controlled is highly concrete, as for example car seat adjustment (Figure 3.9), the problem may be easily solved by using highly literal mapping. If the variable is of a highly abstract nature, as for example low quality VHS or high quality S-VHS, it may be much more difficult to express.

It is often said that the workings of electronic based products have become completely abstracted. This is only partly true. A digital camera still has a lens. Electronic video records still work with tape. Even digital video still works with tape. The workings of lenses and tapes are not abstract. There may be developments which will lead to the eventual elimination of lenses and tapes, but for the time being they form essential parts of the most modern equipment.

As the workings of these physical components are not of an abstract nature, variables related to them do have concrete physical manifestations. Only, current product design has a tendency to hide these physical manifestations, even those which are highly informative to a product's operation. A choice is made in favour of an alternative representation of the variable rather than showing its physical manifestation. Consider the following example. There are still many video recorders on the market of which the tape counter does not show absolute elapsed time in minutes and seconds. The tape counter may show elapsed time in meaningless units and the user does not know how many of those units fit to the length of the tape. Or the counter just runs on without considering which tape is in the recorder in which case the counter is also useless without rewinding the tape and zeroing it. The excuse for leaving the user in the dark is that "current tape technology does not allow us to encode time code on the tape". This focus on technology completely foregoes the fact that an approximate answer to the information the user desires, which is, "is there enough tape left to make the recording?" is in fact available. After all, through the window in the tape housing it can be seen how much tape is on one spool and how much on the other. Only, the tape has been hidden inside the recorder thus barring access to this perfectly useful information, and the tape does not express which spool starts as empty. If on the tape or on the tape compartment an approximate guide to elapsed time were shown, the user's question could be answered without having to wait for the latest time coding technology.

Example

Many practical books on interface design teach the desirable in interface design by giving examples of poor design. These negative examples often outnumber the positive examples. I too have given many negative examples in this chapter as positive examples are simply hard to come by. One of the problems with this approach is that it attempts to change the status quo by showing the status quo. Another is that in terms of affordances compromises have been made because of technical or cost issues. What I attempt here is to give an example of how product design could be made more expressive by means of affordances which show what results may be expected from a certain action.

The concept presented here is for a video deck, that is, a video recorder without programming features[1]. I realise that the interface for programming is one of the most difficult parts of a video recorder. The idea behind concentrating on the operation of the tape section is that, if an elegant solution to that part of the problem can be found, there will be less 'noise' surrounding the programming section. Furthermore, the video deck is not complete. There are still many features which have not yet been thought through from an affordances point of view and which remain unimplemented. One such a feature which has been omitted is a remote control. Remote controls are one of today's user interface consumer horrors together with digital watches, digital thermostats and microwave ovens, and therefore do in fact deserve close attention. However, I think that the main problem designers have with remote controls is that they have to create meaningful relationships between one meaningless box - the video recorder - and another - the remote control. I think that the design of an intuitive remote control may prove much easier when the design of the video recorder itself is meaningful.

The video deck as it stands should thus be regarded as a test bed for affordances. While not complete as a product it provides a context for affordances to be meaningful.

Video deck example

Overall formgiving

Figure 3.10 (left) shows a foam model of the video deck concept. For identification of the device as a piece of video equipment it relies on a tape compartment which shows the tape when present and echoes its form when not. Curving and converging lines indicate the insertion path for the video tape. Figure 3.10 (right) shows the same model with mains, video-in and video-out cables connected. Note how at each connector the outline of the video deck is broken as an indication that at these locations the device communicates with the outside world.

The tape compartment has a degree of direction to it, implying tape movement from the left tape reel to the right, which is indeed the case when the tape is played.

---

[1] The videodeck concept was entered for a competition organised by the Sekisui Design Corporation, Osaka, Japan, in October 1997

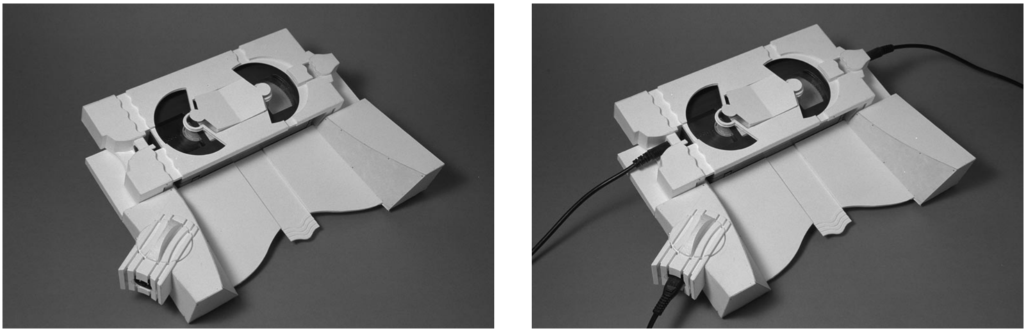


Figure 3.10
An overview of the video deck without (left) and with (right) cables.

Mains connector and on/off switch

Figure 3.11 shows the transformer unit with mains connector and on/off switch. The concavity of the surface surrounding the connector hints at the fact that the connector is an input. The ribs are meant to be reminiscent of power. On top of the transformer unit the ribs act as flow lines. The rotary switch diameter is larger than the width of the transformer unit to stress its ‘rotatability’. On the rotary switch the ribs are continued. By curving and tapering the ribs and heightening the middle rib on the rotary control ‘pinchability’ is expressed. The ribs in combination with the rotary control can express blocked flow, i.e. power off, and flow, i.e. power on (Figure 3.12).

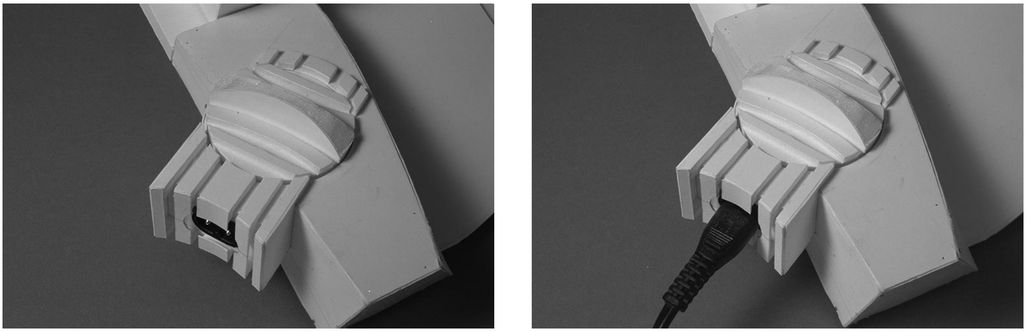
Fast forward and reverse

The fast forward and reverse control is a toggle (Figure 3.13), emphasizing its ‘either/or’ nature. Through the shape of the windows in the tape compartment which reveal the tape the toggle control assumes a two headed arrow like shape.

It is positioned in-between the tape reels with the finger sized button parts of the control acting right in the middle of the reels. In combination with the left to right direction of the tape compartment this cues the user that pressing the right part of the toggle will fast forward the tape while the left part will reverse it.

Eject

The eject is a ribbon coming out of the tape compartment (Figure 3.14). The ribs at the end of the ribbon suggest movement towards the user. Because of its ribbon-like nature only pulling is meaningful, while pushing is not.

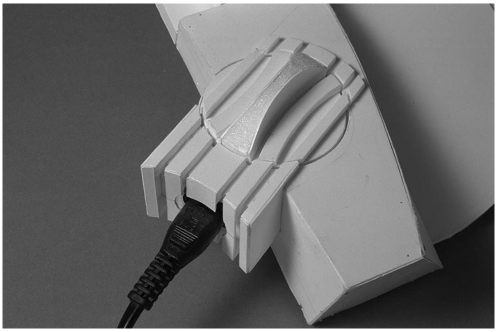


a: without mains cable and power is off

b: with mains cable and power is off

c: middle rib expresses 'pinchability'

d: powering up the video deck



e: power is on

Figure 3.11
Transformer unit with mains connector and rotary power switch.

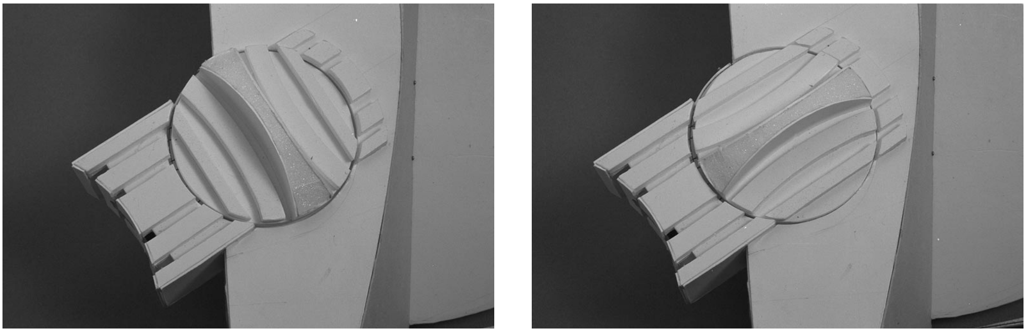
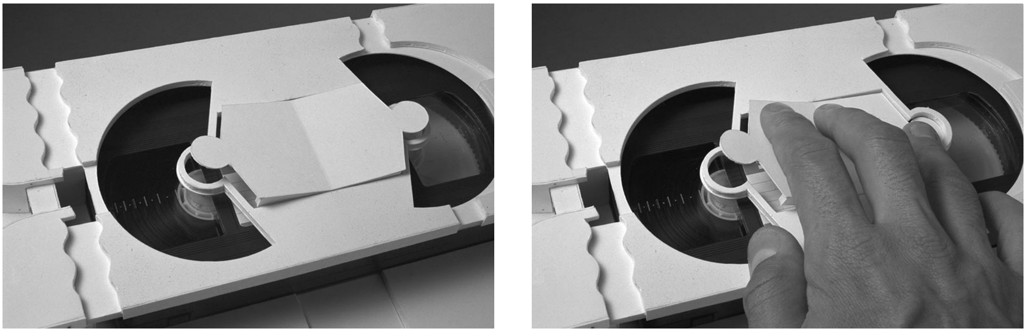
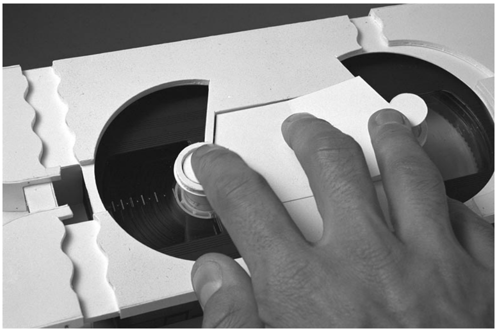


Figure 3.12
Top view of the power unit switched off (left) and on (right)



Neutral position

Fast forward



Reverse

Figure 3.13
The fast forward/reverse toggle control is positioned between the tape reels.

Figure 3.14
The eject ribbon.

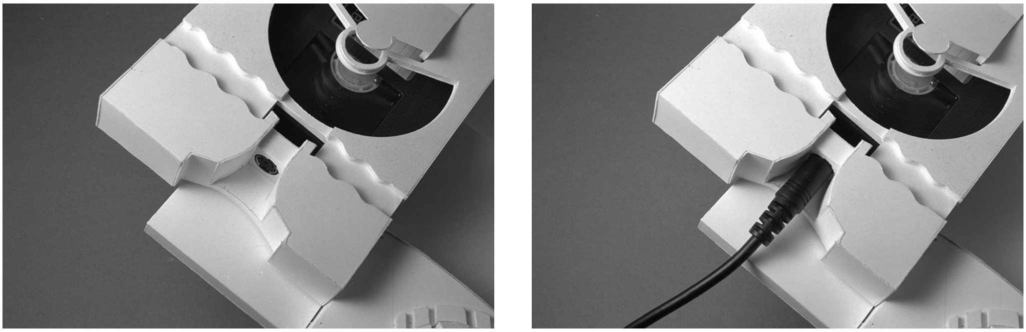


Figure 3.15
The video-in connector without cable (left) and with cable (right).

Video-out and video-in

With current video recorders it is often impossible to distinguish inputs from outputs except for through their labelling. Here an attempt is made to distinguish video-in and video-out connectors through formgiving. The video-in connector is shown in Figure 3.15 while the video-out connector is shown in Figure 3.16. In both cases the connectors are the same but the context in which they are placed is different and says something about their functionality.

Play and record

The play slider is situated to the right of the tape compartment. The wave form indicates that the play slider can mate with the central part of the tape compartment. By pushing the play slider inwards the play function is activated (Figure 3.17). The play slider houses the video-out

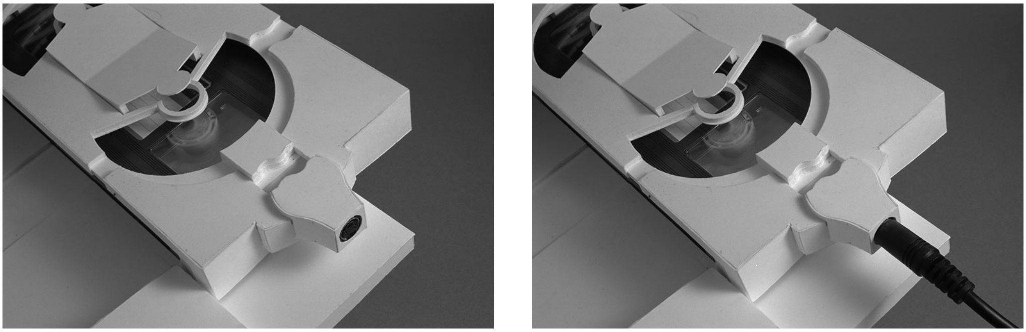


Figure 3.16
The video-out connector without cable (left) and with cable (right).

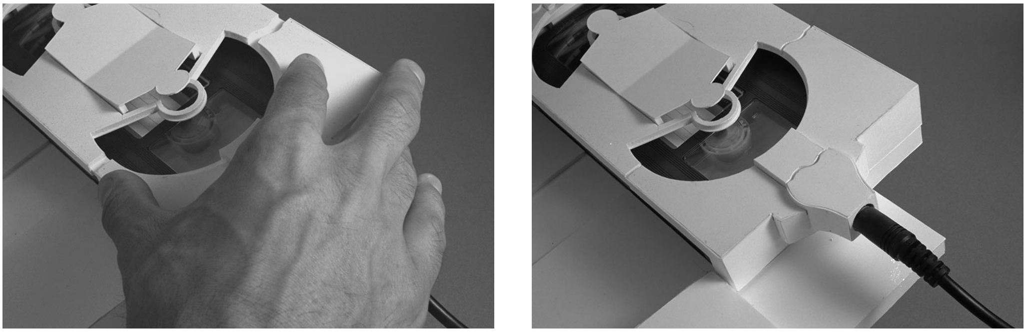


Figure 3.17
The play slider being activated (left) and the device in play mode (right).

socket. This is to emphasise that by sliding the play control inwards and thus activating the play function, information will flow out of the video-out socket to the television.

The record slider (Figure 3.18) is placed to the left of the tape compartment. Again a wave form indicates that the record slider can mate with the central part of the tape compartment. Pushing the record slider inwards activates the record stand-by mode. The record slider houses the video-in socket. In this way I try to stress that by sliding the record control inwards and thus activating record stand-by mode, information flows in through the video-in socket from another video deck.

The play slider envelops half of the right tape reel, something which the record slider does not do with the left tape reel. The forms of the control thus reflect that

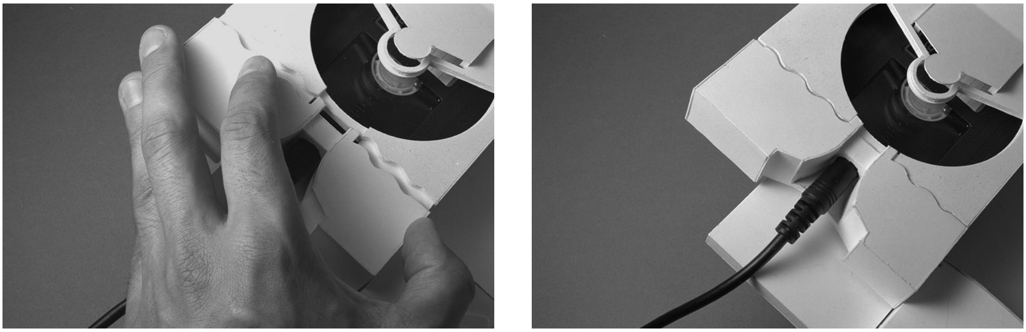


Figure 3.18
The record slider being activated (left) and the device in record-standby mode (right).

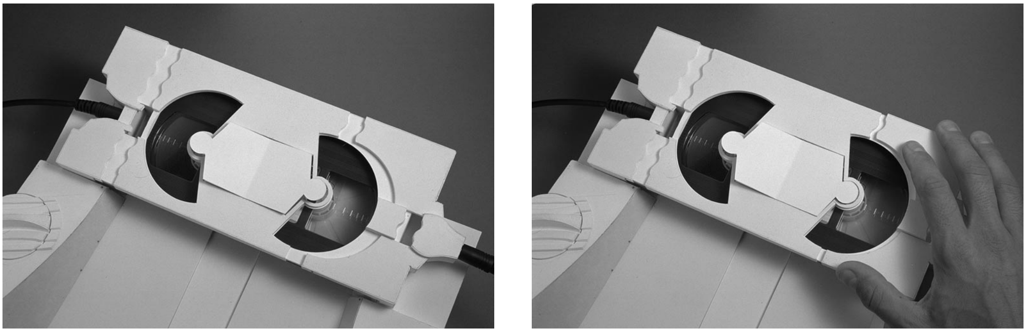
pushing in the play slider activates tape movement while pushing in the record slider does not. The latter does not cause the tape to run but merely activates record stand-by mode.

Note that both these controls violate some of the aspects of ‘good practice’ mentioned earlier. They do not express well their slidability and how they should be held by the user. The expression of slidability could be improved by adding a ribbed texture on the surfaces which are hidden and revealed by the sliders. To express better that they are controls the forms could be made less edgy and more organic, indicating where the user should place his fingers. On the positive side the forms of these sliding controls do express that they mate with the central part of the tape compartment and their relationships to video-in and video-out.

The different functions (stop, play, record stand-by and record) and how they are activated are summarised in Figure 3.19.

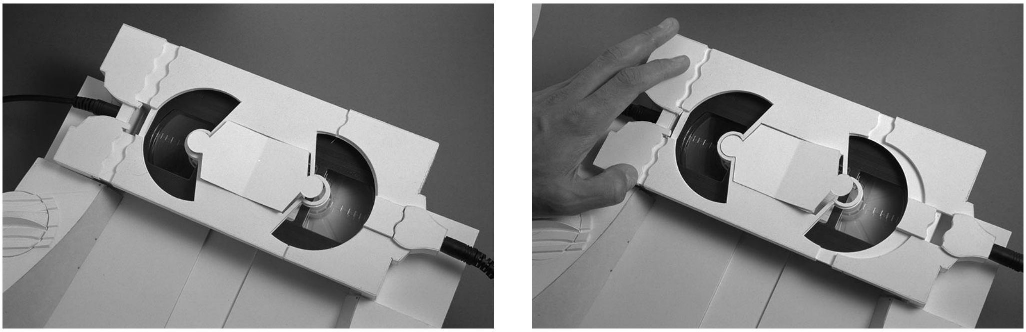
General remarks

An attempt was made to give all the controls of this video deck human dimensions. Sizes and spacing are generous to allow easy manipulation by human fingers. Note how the controls are both control and display at the same time. There is no need to look at a small electronic display to see whether a tape is present, whether the device is



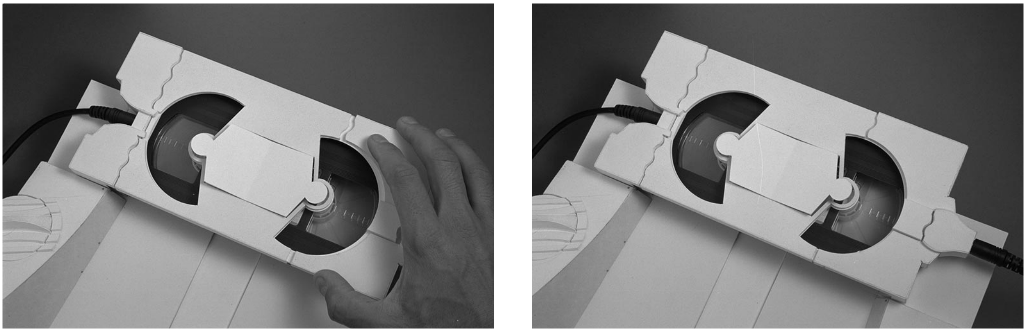
stop

activating play



play

activating record stand-by



activating record

record

Figure 3.19
Summary of the functions stop, play, record stand-by and record.

switched on and whether it is stopped, recording or playing. The device and its controls have become displays in themselves.

No symbols were printed on the controls. Similarly no colour signs were used for labelling functions. For example, red is often used for labelling the record button. Symbols and 'colour signage by convention' do not fall under affordances as they are not intuitively clear and rely on learning. Yet it may be clear that adding symbols and other agreed signs would eradicate any remaining doubts about the function of a control and would help in preventing the user resorting to a manual.

Conclusions

1. Both product semantics and affordances hold promises to address the expressivity problem in the interface of electronic products. However, as yet examples are limited to products with surface representations rather than electronic ones with internal representations.
2. Unlike product semantics affordances have not yet resulted in practical design methodologies. While a range of 'how to' examples exists for product semantics, such examples as yet do not exist for affordances. For affordances to make an impact in the design world it will be necessary to distil a design methodology from the theory of affordances. As long as this does not happen it will be difficult for designers to distinguish between affordances and product semantics, no matter how clear the distinction between the theories from which they originate.
3. Often, decreased usability in modern products is blamed upon the increasing use of electronics in product design. It is true that this trend impairs feedback about a product's internal states. However, the poor usability of products can not be blamed solely on the increased use of electronics. There is a tendency to hide physical manifestations of variables in favour of abstract representations of those variables. Somehow there is an unshakeable belief in the superiority of these abstract representations on electronic displays. The current trend in design is to create closed, meaning-

less boxes which prevent the user from interacting with the components hidden inside. When viewed in this light, the struggle of designers to create user-friendly interfaces, can be seen as an attempt to allow the user to communicate through the barrier devised by their own doing.

4. In books on interface design the negative examples often outweigh the positive ones by a wide margin. I argued that these negative examples need to be complemented with positive conceptual examples. An example of such a conceptual design which focused on usability of the controls of a video deck was given. As it is purely conceptual it does not bear the authority of a positive example which is in production. Yet it is exactly the conceptual nature which could make examples like this valuable. No compromises need to be made for technical, economic or aesthetic reasons. The lack of polish in these areas of product design allows full attention to be focused on usability. By absence of good examples in production, conceptual examples will have to make do. After all, the absence of good examples in production, is not simply the result of technical or economic pressures, or designer laziness. It is the result of a lack of a knowledge in the user-interface community on how usability is influenced by form giving.
5. In this thesis I will use the theory of affordances to imbue the medical 3D system with a more product-like and a less computer-like character. The interface will focus on conforming to and inviting everyday behaviour while hiding the computer as much as possible.

In the next chapter...

In chapter four design concepts are presented for both hand-held and desktop medical 3D systems. The interaction with a DVWS based hand-held computer named Wobbly is designed such that it fits closely with everyday behaviour. A number of concepts for desktop systems is introduced and an example is given of how formgiving can improve the interface of an existing medical 3D system.

4

Concepts for Hand-held and Desktop Computers using Movement Parallax

Summary

In this chapter a number of concepts based on the DVWS are proposed. Two lines of thought are followed through, one based on hand-held computers, the other on desktop computers. The first line of thought resulted in Wobbly, a hand-held computer designed by F.A. Voorhorst and myself. With this hand-held computer the parallax shifts on the screen are not coupled to the head position of the observer but instead to the orientation of the device itself. The benefits of a 3D display for a hand-held computer are argued. The advantages of our particular implementation relative to both other hand-coupled parallax systems and head coupled movement parallax systems are described.

Following through the second line of thought, and working from the brief for a medical work station as drawn up in Chapter 2, some concepts for desktop computers based on head-coupled movement parallax are discussed. These concepts explore ways of making the placement of a cutting plane through a virtual body more intuitive than is the case with existing medical work stations.

Wobbly, a hand-held computer

Hand-held computers, also known as Personal Digital Assistants

Hand-held computers, also known as personal digital assistants (PDAs), are a relatively new development. Although there are considerable differences between the currently available PDAs, as there is as yet no established hardware platform or operating system, they generally share the following characteristics. The size of a PDA is generally no larger than a pocket book, which makes them considerably smaller than a laptop computer. Instead of a keyboard and mouse or trackball a pen is used as an input device. Text and graphics are written directly on to the screen. PDAs are battery powered and make use of solid state memory storage instead of disk based storage. While they are often used as personal organisers (i.e. to manage contacts, appointments and notes), most PDAs are fully programmable and can therefore potentially be used for a wide variety of tasks. Figure 4.1 shows two currently available PDAs.

Figure 4.1
Two currently available PDAs: the PalmPilot by 3COM and the Newton by Apple Computer.

PDAs and the DVWS

The currently commercially available PDAs do not make use of 3D displays. There are two reasons why a PDA and the DVWS may form a symbiotic combination. The first reason is that a 3D display may alleviate a problem which is typical for hand-held computers, namely the limited amount of screen space. The second is that by using a PDA the problem of head tracking in the DVWS may be bypassed. In the next two sections I will cover these problems in more detail.

The problem of screen size in hand-held computers

Interfaces of the current generation of PDAs are much related to the graphical user interfaces (GUI) for desktop computers. Screen diagonals for desktop monitors are usually no smaller than 14" and even notebook computers now come with screens of 10" diagonal or more. PDA screens, however, are much smaller with screen diagonals of approximately five inches. Consequently, one of the main problems with hand-held computers is the lack of screen real-estate. The screen size does not only limit the output of visual information but also the space available for pen input. For this reason the majority of design efforts in the field of PDAs is aimed at space saving techniques for a 2D interface (Ahlberg and Schneidermann, 1994; Matias et al., 1994). However, there are a few papers on combining PDAs and 3D displays (Amselem, 1995; Fitzmaurice et al, 1993). If the user is allowed to explore a virtual space by moving his PDA, the space constraints imposed by the small screen can be alleviated.

Fitzmaurice et al. attached a six DOF tracker to a 4" colour LCD monitor which displayed images rendered in real-time by a graphics workstation, a set-up which they call Chameleon. One of the applications they proposed for Chameleon was a 3D spreadsheet. Moving the monitor horizontally allowed the user to explore the columns of the spreadsheet while moving it vertically allowed him to view the rows. Moving the monitor closer or further away gave the user access to older and newer versions of the spreadsheet. In this implementation the user always looked straight ahead into the workspace and the orientation information of the tracker was not used.

Fitzmaurice et al. noticed that if the user would be given control over his viewpoint through rotation of the PDA, he would need to change his position and viewing direction. In other words, the viewing direction would need to coincide with the normal of the screen. To allow the user to view the virtual space from a different angle without having to adjust his head they proposed that the angular information of the tracker would be exaggerated so that a 45° tilt would result in a full 90° smooth viewing change.

Apart from the 3D spreadsheet Fitzmaurice et al. suggested various other applications for Chameleon, some of which make use of the orientation information provided by the tracker. Examples included browsing of video data whereby the number of frames per second was proportional with the amount the PDA was tilted to the right or left, and scrolling through a text document by tilting the top or bottom edge of the PDA.

Fitzmaurice et al. also noted that tracking the user's head position along with the position of the PDA may provide greater depth sensation.

A similar set-up was used by Amselem (1995) with the difference that two users could be active simultaneously, each carrying a hand-held TV-monitor with a six DOF tracker. By walking around in a room the two users viewed the same virtual space from different vantage points. In contrast with Fitzmaurice et al. all six DOF freedom were used. A user could rotate the PDA to get a different perspective although this would of course require him to move his head back in line with the screen normal. Amselem tried out this set-up with a virtual model of the city of San Francisco, scaled down to fit in the area covered by the six DOF trackers. Although Amselem used a TV broadcast amplifier to achieve a wireless connection from the graphics workstation to the hand-held TV-monitor, the hand-held device remained tethered through the cable of the six DOF tracker.

The head tracking problem turned 'upside down'

A practical problem which impedes widespread use of the DVWS is the measuring of head position. There are various ways in which head position can be measured. They can be divided in head free and non-head free methods. Non-head free methods generally consist of some mechanical construction attached to the head of which the movement is measured through potentiometers or encoders. Head-free detection can be achieved through analysis of video, infra-red, electro-magnetic or ultrasound information.

The main disadvantage of non-head free tracking is that it forms a burden to the user. Although the helmet and the mechanical arm attached to it are much lighter than a helm in immersive VR and do not hamper communication as much, they remain encumbering devices. Head free trackers, on the other hand, are generally expensive.

This head tracking difficulty can be avoided while still obtaining depth impression through movement parallax by adopting a different approach. Instead of working with a stationary monitor and coupling the parallax shifts to the head position of the user, the user can be considered as stationary whereby the parallax shifts are coupled to the movements of the monitor. While it is possible to mount a desktop monitor in such a way that it can rotate about its screen middle, the weight of the monitor requires a large and expensive construction if it is to move smoothly (Groen, 1988). PDAs, on the contrary, lend themselves much better to being moved around because of their small size and light weight. While a conventional DVWS set-up with a monolithic monitor invites head movements, a PDA featuring the DVWS will invite hand movements.

Wobbly, a different approach

How Wobbly works

With Wobbly the virtual world is not connected to the space in which the user is present, but to Wobbly itself. The difference between Amselem's set-up and Wobbly can be explained as follows. In Amselem's set-up the screen of the PDA can be thought of as a plane of glass which offers a view onto a virtual world which exists in parallel with and is rigidly connected to the user's physical environment. Wobbly's screen can be thought of as the window of a diorama: the virtual scene is rigidly connected to the PDA itself (Figure 4.2). To allow the user to work with a virtual space which is larger than the screen, the virtual world can be moved relative to the hand-held computer by dragging it along the screen with a pen. Through rotation of the device, the user can look at the virtual world from different angles. By adopting this different approach pure orientation information suffices. In order to connect the virtual world to Wobbly, the parallax shifts on the screen need to be coupled to the orientation of Wobbly relative to gravity. Information on rotation about the normal of the screen is not needed and there-

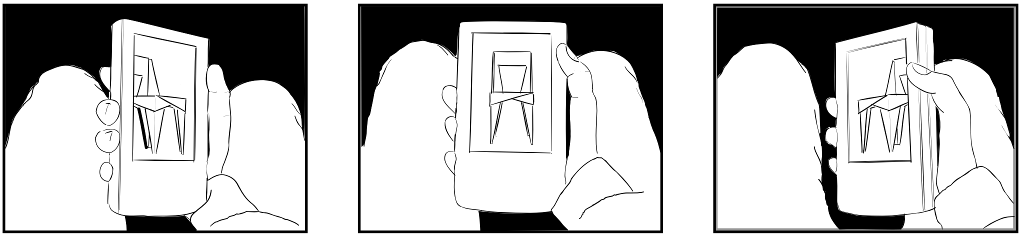
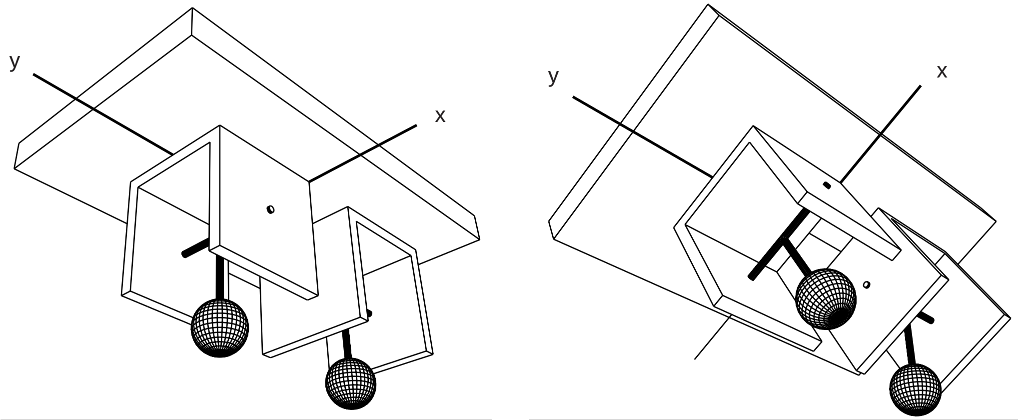


Figure 4.2
A user looking at a chair displayed on the PDA which he holds on his lap. Through rotation of the PDA, the user can look at the chair from different angles. Here the fixation point is placed at the front of the chair so that it appears as if the chair is ‘glued’ to the back of the screen by its front legs.

fore a two DOF rotation sensor suffices. In the first prototype orientation relative to gravity was measured by means of two pendulums. The working principle is illustrated in Figure 4.3. The two pendulums are suspended in U-profiles which are attached to the bottom of the PDA. When the PDA is angled, the weights of the pendulums adopt their lowest position. The rotation of the pendulums with respect to their U-profiles and thus to the main body of the PDA is measured by means of encoders taken from a mouse. It should be emphasised that the pendulums were chosen with a view to building a prototype quickly and at low cost (Figure 4.4). In the prototype the pendulums are attached to a Polaroid TFT Overhead Display with a light box underneath (Figure 4.5). The main disadvantage of the pendulums is that they react not only to the PDA’s rotation but also to its translation, resulting in undesired oscillations and changes in perspective. A production version of a DVWS-based PDA could be equipped with inclinometers or miniature gyroscopes instead of the pendulums. A version of Wobbly which uses gyroscopes has the advantage that the screen can also be kept vertical. Summarising, the two main characteristics of Wobbly are the illusion that the virtual world is attached to it, and the sensing of orientation with two DOFs, whereby the sensing system is contained entirely within Wobbly, and thus does not require external equipment. The sensing of orientation through pendulums is not central to the Wobbly concept and can be replaced by another method.

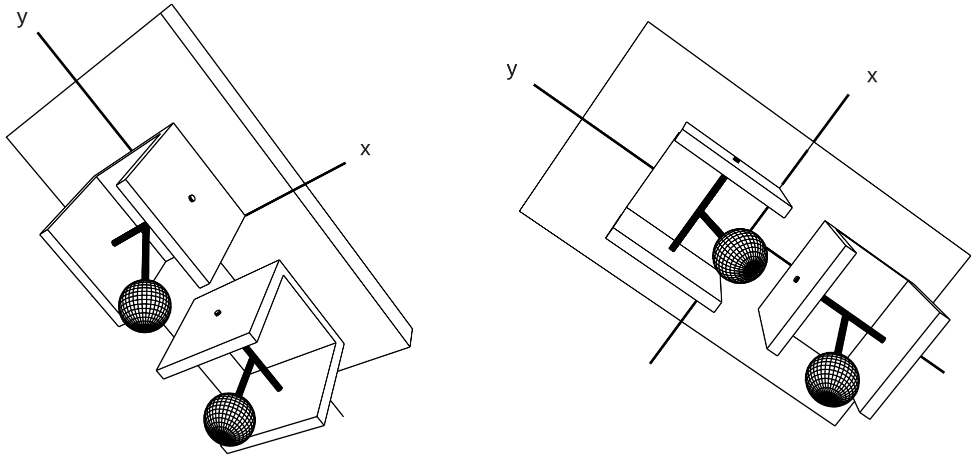
Wobbly's pros and cons

Wobbly's main advantage is that as it measures its orientation relative to gravity it does not require any external equipment. This is in contrast with the PDAs of Amselem and Fitmaurice et al. which use six DOF electro-magnetic trackers. These trackers need to have a source present which is stationary with respect to the environment and need to be within reach of this source. PDAs based on these electro-magnetic trackers could only be used in places where an infrastructure catering to its use would be present. This requirement limits the mobility of such a PDA which is exactly its raison d'être.



The PDA in horizontal position.

The PDA is rotated over the y-axis only. The left pendulum is not rotated with respect to its U-profile, whereas the right one is.



The PDA is rotated over the x-axis only. While the right pendulum is not rotated with respect to its U-profile, whereas the left one is.

When the PDA is rotated over both the x-axis and the y-axis, both pendulums are rotated with respect to their U-profiles.

Figure 4.3
Four bottom views of the DVWS-based PDA which show the working principle of the pendulums. Two pendulums, which have orthogonal directions of movement, are suspended in U-shaped profiles. These profiles are attached to the bottom of the PDA. When the PDA is angled the weights of the pendulums adopt their lowest positions. Thus the orientation of the PDA relative to gravity can be measured.

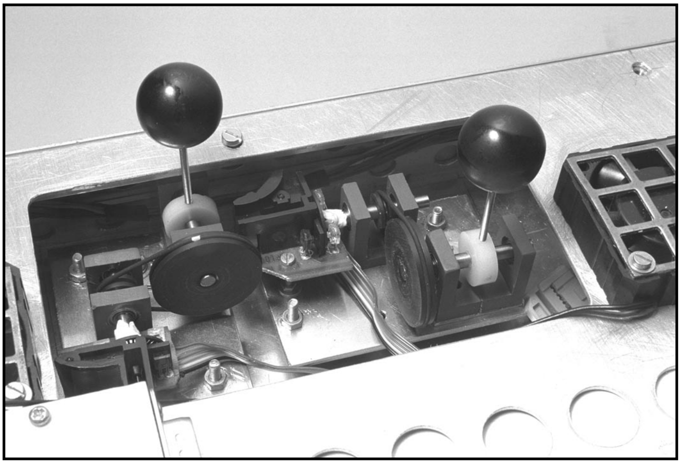


Figure 4.4
Bottom view of Wobbly showing two pendulums which measure tilt with two DOFs.

A second advantage of Wobbly is that it is relatively low-tech and costs less than hand-held computers equipped with six DOF trackers. By moving from six DOFs to two DOFs, the complexity of the overall set-up is reduced.

Third, since the virtual world is coupled to Wobbly rather than the space in which the user is present, it is possible for the user to obtain depth information about the virtual world without walking around.

Fourth, while we concentrated on using Wobbly with the DVWS, it would be possible to have a mode in which Wobbly is used as a panoramic viewing tool. This may be

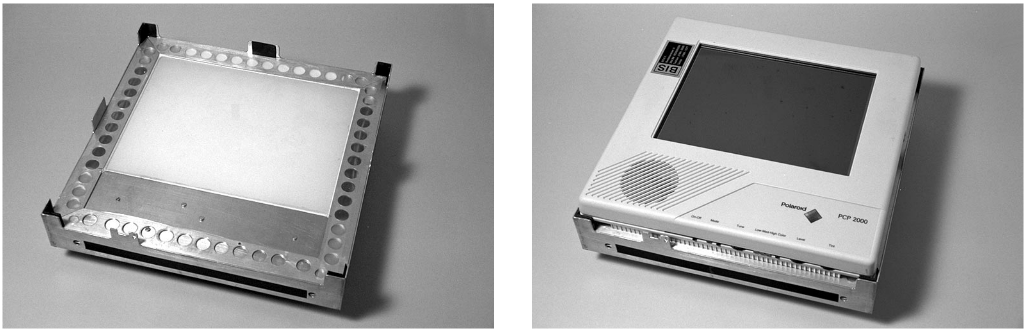


Figure 4.5
Wobbly prototype. Light box only (left) and light box with TFT overhead projection screen (right).

clarified by picturing a virtual camera attached to the back of Wobbly, with the line of sight of the camera coinciding with the screen normal.

A limitation of Wobbly is that while it knows its orientation with two DOF, it does not know its position in the world. With Wobbly, unlike with the aforementioned set-ups, the virtual environment as seen through the PDAs screen cannot be stationary with respect to the real environment. Parallel virtual and physical worlds cannot be realised. Also, since Wobbly senses two DOFs only, gestural input is limited compared to six DOFs PDAs. However, it should be noted that, depending on the method of orientation sensing, gestural input need not be limited to tilting. For example, while during prototyping we considered movements of the pendulums caused by lateral movements of the PDA as noise, these movements could also be looked at as being useful for richer gestural input.

Future developments

In this thesis the design of a PDA featuring movement parallax will not be pursued any further. However, research on this subject was continued in the form of a master degree project. Molenaar (1997) designed a PDA which facilitates information retrieval and transportation in a hospital environment. A user of this PDA can explore a 3D virtual human body by tilting the PDA and by moving it towards and away from him.

Concepts for a desktop work station

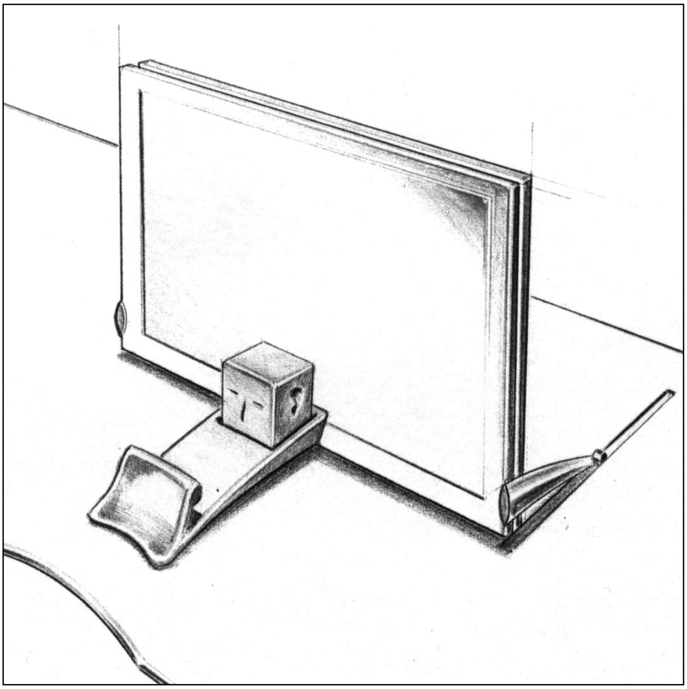
The starting point: An existing medical work station

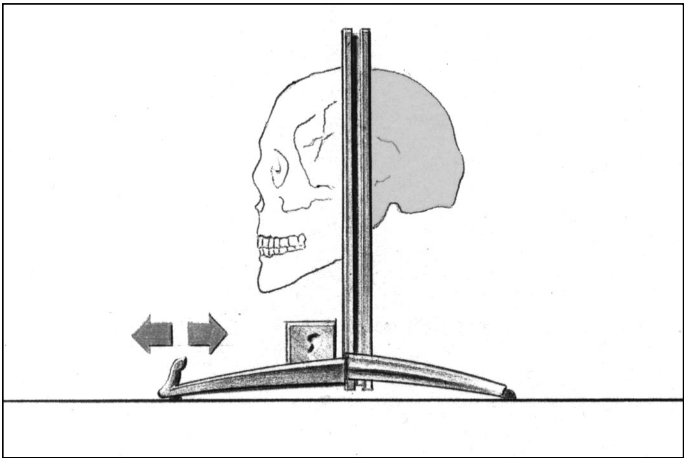
In the next section a number of concepts for desktop work stations for medical 3D work are described. As a starting point I have taken a commercially available set-up by ISG. This system is used in the Dijkzigt Academic Hospital in Rotterdam for preparation of craniofacial surgery. By initially considering the functionality of the system as fixed all attention can be focused on alternative ways in which the user can access this functionality.

In the Dijkzigt Academic Hospital the ISG system was used mainly as a visualisation tool. A laboratory assistant chose certain sections and views on those sections. These images were recorded onto transparent film and put up in the operating room for reference. Cutting planes could be placed only at right angles to the main axes of the skull. The position of these cutting planes on an axis was adjusted through on-screen sliders by means of a mouse. The orientation of the head was changed by means of mouse controlled scroll bars at the bottom and the side of the window displaying the head.

The base functionality of the aforementioned system can be summarised as the positioning of a cutting plane along one of the three main axes of a virtual head, whereby the cutting plane is always perpendicular to that axis. It appears unlikely that the orientation of the cutting plane was restricted for computational reasons, since it is possible to slice away any part of a voxel model. It seems more likely that the decision to restrict the orientation of the cutting plane was made to complicate the interface no further. I therefore assume that the desired functionality is to place a cutting plane through a virtual body with six simultaneous DOFs: three DOFs for rotation and three DOFs for translation. In addition the user needs to be able to rotate the total scene — which consists of the virtual body, the cutting plane, and possibly some markers — with three DOFs to achieve the desired view on the scene. However, the number of necessary degrees of freedom can be reduced for three reasons. First, when the user places a cutting plane through the virtual body, it seems reasonable to assume that the resulting cross-section is the part of the virtual body which he is most interested in. The most accurate way of inspecting the cross-section is when the cross-section is not perspectively foreshortened. This is the case when the cutting plane is formed by the monitor screen or a plane parallel to the monitor screen. Second, the dimensions of the largest possible cutting plane through the virtual head are within the size of a monitor screen. As a result any cutting plane can be achieved through rotation of the virtual head relative to the cutting plane with three DOFs, and translation of the virtual head relative to the screen with one DOF. Third, through the use of movement parallax the user can look around the scene to judge the positioning of the cutting plane within the virtual head without having to rotate the whole scene. When a Fish Tank VR display is used, the illusion can be created that the cross-section coincides with the screen or with a plane parallel to the screen. Of the six concepts which are presented in the remainder of this chapter only in the first one the orientation of the cutting plane is restricted to being perpendicular to the main axes of the head. In the other five the cutting plane can be placed in any orientation.

Figure 4.6
The user chooses one of six possible orthogonal views by means of the cubic head prop while the sliding screen acts as a cutting plane with one DOF.





Six alternative concepts

The idea behind the first concept (Figure 4.6) is that the virtual body, say a head, is stationary relative to the real environment, while the screen displaying the head can be moved and acts as a cutting plane. To allow the user a clear view of the cutting plane it is necessary to display the part of the head in front of the screen either in wireframe mode or as a transparent shell (In Figure 4.6 and the remaining illustrations in this chapter, the wireframe or transparent part of the head which is in front of the cutting plane is tinted light grey, whereas the part behind the cutting plane is tinted dark grey). The advantage of displaying the part of the head in front of the screen in one of these two manners over not displaying it at all, is that the user will always have a frame of reference. This reduces problems of disorientation. A small cube which represents the head is used as an input device. On this cube a stylised nose, eyes and ears can be seen and felt. When the cube is picked up and put down on one of its other faces, the orientation of the head which is displayed on the screen changes accordingly. This ‘cubic head’ thus functions as a kind of physical prop (Hinckley et al., 1994; Stoakley et al., 1995).The user can move the screen towards or away from him by means of the handle protruding from the screen. This will move the cutting plane through the head. Through the combination of the six faces of the ‘cubic head’ and movement parallax the user can view the head from different points of view. This set-up limits the possible orientations of the cutting plane to those at right angles to the main axes of the body. However, when a sphere is used instead of a cube the virtual head can be adjusted to any angle. As a consequence any cutting plane through the head can be chosen (Figure 4.7). Note that the spherical head prop has three DOFs unlike a conventional trackball which is limited to two DOFs. Furthermore, the coupling of the spherical prop and the virtual head is absolute. The abstracted ears, nose and eyes offer the user tactile feedback on the orientation of the spherical head prop and thus the virtual head. Since the sphere and the screen can be manipulated simultaneously the system allows the user to perform two handed input. Buxton (1986) considers two handed interaction as an important step in achieving better performance during input. In Guiard's terms (1987) this system offers bi-manual, asymmetric manipulation as the two hands play dif-

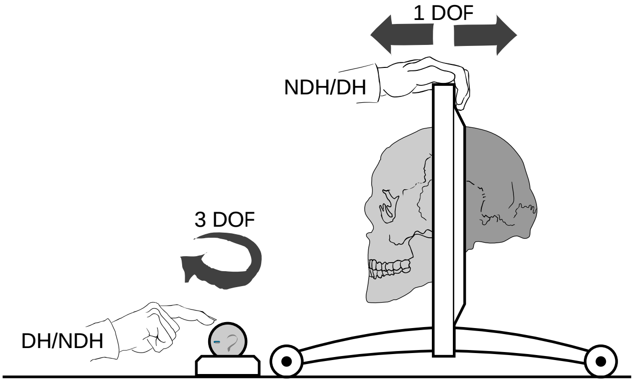


Figure 4.7
The virtual skull can be rotated by means of the spherical head prop with three DOFs while the screen acts as a cutting plane and can be translated with one DOF (Note: DH is *d*ominant *h*and, NDH is *n*on-*d*ominant *h*and)

ferent roles. One hand does rotation, while the other performs a translation. Guiard claims that in bi-manual manipulation the task executed by the non dominant hand is characterised by a low temporal and a low spatial frequency. The non-dominant hand sets the spatial reference frame for the dominant hand which operates with finer temporal and spatial resolution. The non-dominant hand can thus be labelled macrometric while the dominant hand is labelled micrometric. It may be noted that in the concept presented here there is no clear distinction as to which of the two tasks - rotation of the three DOF trackball and translation of the screen - is macrometric and which is micrometric. Lateral preference may differ among users and with this in mind the set-up was made symmetrical.

Whilst these set-ups may offer more intuitive positioning of the cutting plane than conventional configurations which feature a mouse or a trackball, there is still a split between the display space and the physical prop. One way to eliminate this split would be to rotate the virtual head directly with the one hand, while choosing the cutting plane through sliding the screen with the other hand (Figure 4.8).

Another set-up which avoids having a separate physical prop is shown in Figure 4.9. In this set-up the screen does not move but it still acts as a cutting plane. The user can both rotate and translate the virtual head and find the desired cross-section by pushing the head through the screen. In these set-ups, in which the screen acts as a cut-

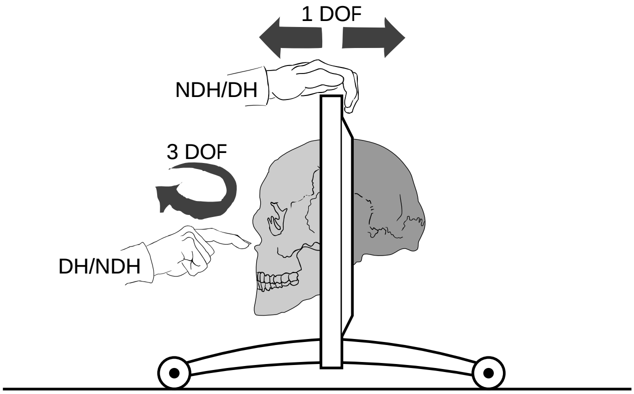


Figure 4.8
The virtual skull can be rotated by means of the spherical head prop with three DOFs while the screen acts as a cutting plane and can be translated with one DOF

ting plane and one hand directly manipulates the virtual body, a problem is that the virtual body can get ‘lost’ behind the screen. Although the virtual body remains visible, it can become difficult to seize when only very little of it extends in front of the screen.

One solution to this would be to have a stationary screen with a moving frame around it (Figure 4.10). Moving the frame will move the cutting plane while leaving the screen standing. Even when the frame, and thus the cutting plane, is moved nearest to the user, the user can still reach through the frame and rotate the virtual body without the monitor screen forming an obstruction

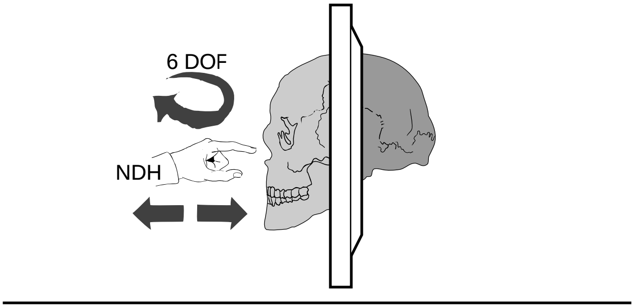


Figure 4.9
The user can rotate and translate the virtual head with six DOFs while the stationary screen acts as a cutting plane

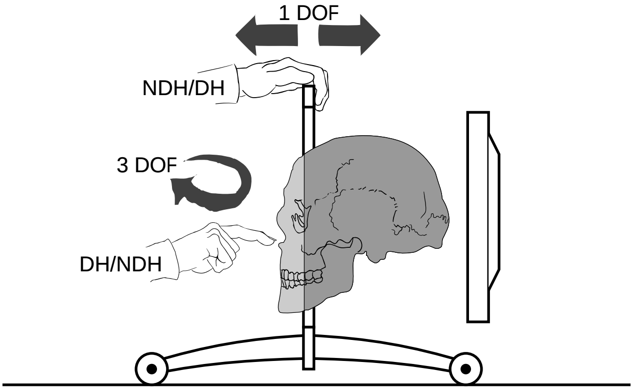


Figure 4.10
The user can rotate the virtual head directly with the one hand, while with the other he can move the frame which surrounds the stationary screen and which acts as a cutting plane.

Another concept in which the virtual body does not become unreachable is shown in Figure 4.11. Here the screen is stationary and has a stationary frame in front of it, which acts as both a cutting plane and a reduction screen. A reduction screen is a frame which is put in front of a monitor and which hides the monitor edges so that depth impression is enhanced by reducing depth cue conflicts between the image displayed on the monitor and the monitor itself. The user can push the virtual head all the way through this cutting plane and pick it up again without the screen forming a barrier.

Cardboard mock-ups were made of all the concepts to allow some informal ergonomics testing. This was done to make sure that the user could comfortably reach all parts

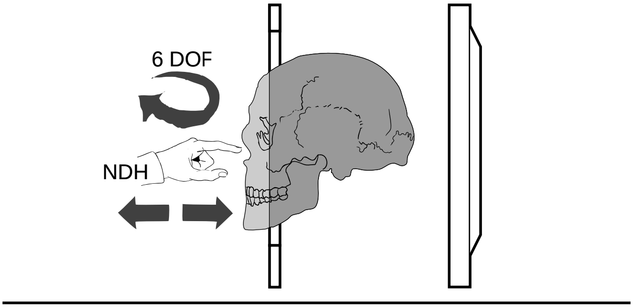
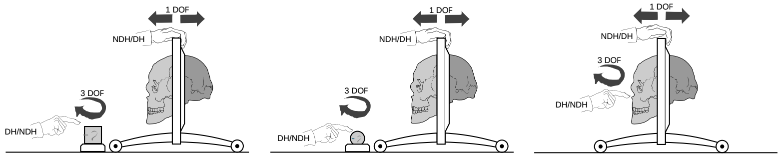


Figure 4.11
The user can rotate and translate the virtual head directly while the stationary frame acts as a cutting plane and stands in front of the stationary screen.



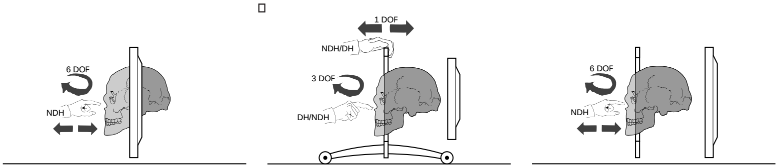
mech. moving parts	screen	screen	screen
virtual body movement	rotation (90º inc.)	rotation	rotation
uni/bi-manual	bi-manual	bi-manual	bi-manual
direct contact	no (cubic prop)	no (spherical prop)	yes: rotation
body can get lost?	no	no	no
max. dist. to screen	1 * body diameter	1 * body diameter	1 * body diameter

Figure 4.12
Desktop concepts compared

of the virtual head with the cutting plane, and that his hand did not obscure his view. None of the concepts were implemented. Summarising, the aforementioned concepts differ in terms of the following six characteristics:

1. whether there are mechanically moving parts.
2. whether the virtual body is stationary relative to the environment.
3. whether operation is uni-manual or bi-manual
   This of interest with regard to manipulation of the virtual body. For example, if the user needs to insert a probe and then decides that the positioning of the cutting plane is not correct, will he need to put down the probe in order to adjust the cutting plane bi-manually?
4. whether the virtual body is directly manually rotated and/or translated.
5. whether the virtual body can get lost behind the screen.
6. the maximum distance the virtual object extends in front of the screen. This is of influence on the amount of freedom the user has to move without the virtual body being clipped by the monitor edge. The further the virtual body extends in front of the screen the less the user can move.

   These characteristics are summarised in Figure 4.12.



mech. moving parts	none	frame	none
virtual body movement	rotation + translation	rotation	rotation + translation
uni/bi-manual	uni-manual	bi-manual	uni-manual
direct contact	rotation + translation	rotation	rotation + translation
body can get lost?	yes	no	no
max. dist. to screen	1 * body diameter	1 * body diameter	2 * body diameter

Figure 4.12
Desktop concepts compared (continued)

In desktop VR systems featuring direct manual rotation and/or translation, the user's hand is present in the graphics environment. Unlike in immersive VR, in desktop VR the hand which is seen rotating the virtual body is a physical object, not a virtual representation. Therefore the hand cannot be shrunk in size or replaced by an alternative representation. As a consequence the user cannot simply seize the virtual body in its centre as his hand would remain visible when it should be hidden, thus destroying the occlusion depth cue (Ware and Jessome, 1988, Ware 1990). Such occlusion anomalies do not occur with the first two concepts in which the virtual body is rotated by means of a cubic or spherical prop. Also, when the hand is to rotate the virtual body directly, it becomes necessary to detect contact of the hand and the virtual body. With an object as organically formed as the human body this is very computationally intensive. An acceptable solution may be to encase the virtual body in a translucent sphere, and to rotate the virtual body by means of this sphere. When rotation is done at the side of the sphere little occlusion occurs and the user can still see what he is doing. In addition, it is less computationally intensive to detect contact of the hand and the sphere than of the hand and the virtual body.

In the next chapter...

An experiment is described which tries to establish the smallest number of finger tips to be registered without neglecting performance and comfort issues in rotation of a transparent sphere which encases an object.

5

The Importance of Simultaneously Accessible Degrees of Freedom

Summary

> *“Es ist kein Zufall, daß das Wort dreidimensional das Wort räumlich fast verdrängt hat.”*
>
> *(Lannoch and Lannoch, 1987)*

In many 3D systems the display method receives more attention than the input method, resulting in sub-optimal manipulation. Currently available medical 3D work stations often use a conventional input device with two degrees of freedom such as a mouse or a trackball. Another popular input device is the box shown in Figure 5.1, which offers control over the six degrees of freedom of a virtual object, with one degree of freedom at a time.

A recurring theme in the concepts shown in Chapter 4 was the use of movement parallax to unify the display and manipulation space to enable direct manual manipulation. In the current chapter[1] it is argued that manipulation by means of the finger tips is a suitable approach to direct manipulation in desktop VR systems. An interface is pro-

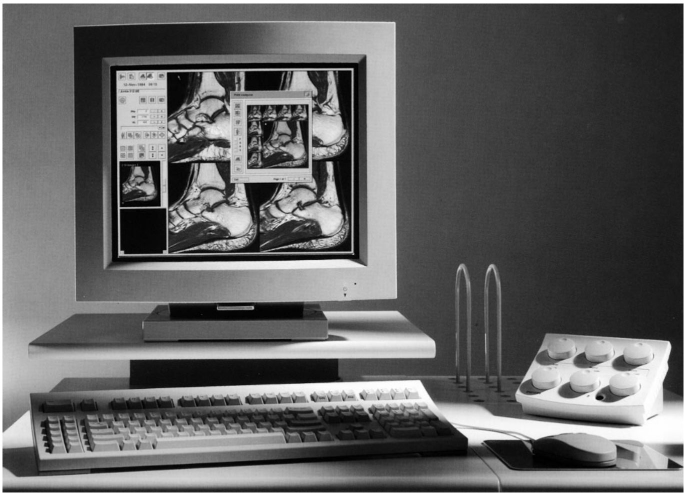


Figure 5.1
A medical workstation with a mouse with two degrees of freedom and a control box with one rotary control for each of the six degrees of freedom.

posed in which the non-dominant hand rotates a transparent virtual sphere which encases the virtual object to be rotated, while the dominant hand holds an instrument for manipulation. When designing an interface which makes use of finger tip control, for technical reasons the number of finger tips to be registered should be minimised. In order to predict performance and perceived comfort during rotation of virtual objects with different numbers of finger tips, an experiment was set up using real, physical objects. Since the different numbers of finger tips correspond to different numbers of degrees of freedom, the relevance of the results is not restricted to the finger tip controlled interface proposed here, but can also be used to evaluate existing approaches to rotation. An overview of those existing approaches to rotation is given.

Introduction

In 3D software rotation of virtual objects plays an important role. For example, to fully understand the geometry of a virtual object, the user has to be able to view all sides of that object. In current 3D software rotation often still is decomposed into three separate rotations around orthogonal axes. This implementation only allows one DOF to be accessed at a time. With most workstations running 3D software there is a dichotomy between the display space showing the virtual objects and the manipulation space containing the input device. The user cannot manipulate the object which exists “there” on the screen directly but has to do this by means of the input device.

In immersive virtual reality these two problems are solved by seizing virtual objects in the middle with a dataglove (Sturman and Zeltzer,1994). The hand manipulates an object at its centre, after which the object’s position and orientation are coupled to that of the glove. Unlike handling of objects in everyday life, the virtual object is not handled by its outside. Instead, the virtual hand moves into the object and obscures part of it. While in immersive VR this problem may be alleviated by replacing the virtual

---

[1] This chapter was published in the journal “Behaviour and Information Technology” (Djajadiningrat, 1997b)

hand with an alternative representation, in desktop VR much of the virtual object would be obscured by the physically present hand. Although this also occurs with physical objects in the real world, in most desktop VR systems this forms a problem since tactile and force feedback are lacking. Visual and possibly auditory feedback are then the only means through which the user can determine whether he is in contact with an object.

In the interface proposed here the orientation of a virtual object can be changed through finger tip control of an encasing virtual sphere. One approach would be to add virtual handles to the sphere, each handle allowing rotation around one axis, again resulting in decomposed rotation. A second approach would be to rotate the virtual sphere under direct manual control. A compromise then needs to be made between the number of fingers needed for comfortable object control and the technical desire to keep the number of finger tips to be registered to a minimum. In this article rotations by different number of finger tips, which differ in the number of DOFs which can be accessed simultaneously and which include a decomposed variant, are compared.

To make direct control of a virtual object possible, it is necessary that the user can judge distances between finger tip and object. The Delft Virtual Window System (Smets et al. 1987; Overbeeke et al. 1987; Overbeeke and Stratmann,1988) is a desktop VR set-up which enables depth perception through head-coupled movement parallax. With the DVWS the illusion can be created that virtual objects hover in front of the screen. Thus a virtual object can be reached at without the monitor screen forming an obstruction.

With the interface proposed here this advantage is used. It features two-handed input (Buxton et al.,1986) whereby the non-dominant hand changes the object’s orientation while the dominant hand holds a manipulator.

Through rotation of the object itself, the user would be able to view this object from any angle, rather than merely from those which can be comfortably achieved through head movements.

First the currently available methods for rotation of virtual objects are described. Secondly, the proposed interface is described in more detail and finally, the experiment is described.

Existing Approaches to Rotation

In a complete interface for 3D visualisation there are three DOFs for rotation and three for translation, bringing the total number of DOFs to six. As this is an overview of rotation of virtual objects, only the three rotational DOFs are discussed. With regard to rotation operations, Mountford et al. (1986) state that subjects mostly used the single axes x, y or z (Figure 5.2). Very few subjects were able to use coupled axes control, and users are particularly unfamiliar with the visual appearance and movement associated with rotating an object around xz or yz. During rotation of objects, users apparently relied on visual, exteroceptive information feedback and not so much on proprioceptive feedback.

The literature on rotation with a variety of input devices is described in the following paragraphs.

Rotation by Means of Input Devices with Two DOFs

Control of One Rotational DOF with an Input Device with two DOFs

A common method of controlling rotation around each axis is through graphically displayed x, y and z sliders as shown at the bottom of Figure 5.2. Note that, although a mouse or a trackball is a device which offers two DOFs, only one DOF is used to control the linear sliders. The motivation for choosing this type of interface appears to be ease of mathematical implementation rather than ease of use for the end user. For rotation about each axis there is a rotational matrix and each of these matrices is coupled to one slider.

However, usually the desired rotation cannot be achieved through rotation around a single axis, but requires rotation about all three axes. When rotation is decomposed into rotation about three axes, users are forced to perform the three operations sequentially. Figure 5.3 (sequence A, bottom picture) shows a sphere which has been rotated about all three axes. If this is the desired end result and the starting point is the top picture, in a system with decomposed rotation the user will have to go

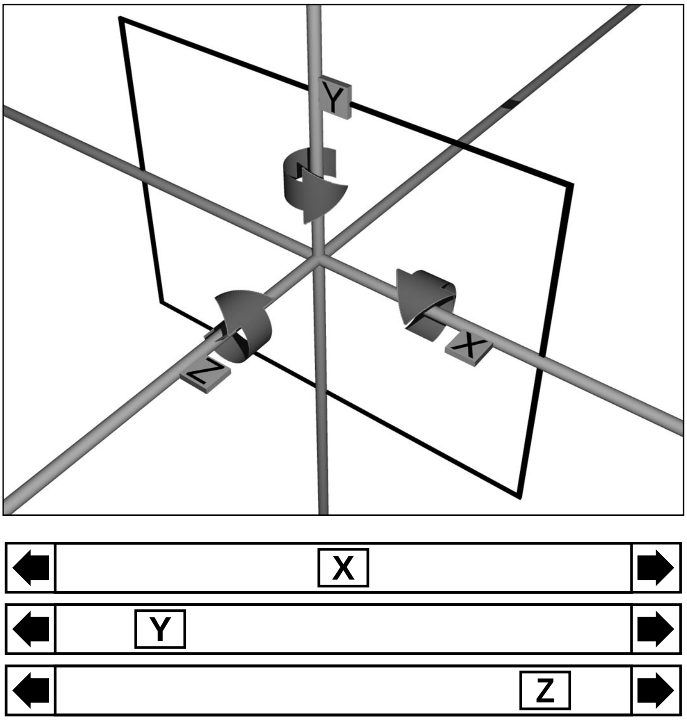


Figure 5.2
A right handed coordinate system. On screen sliders only allow decomposed rotation around one axis at a time.

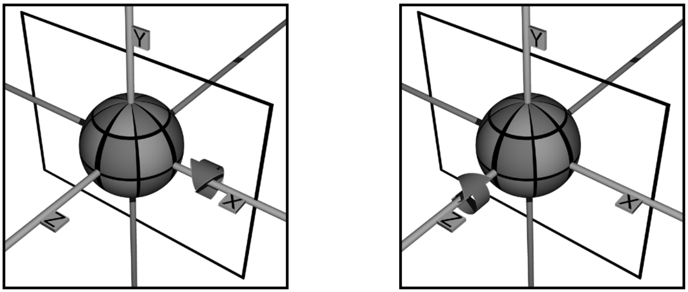
through steps 1-3 (Figure 5.3, sequence A). Note that rotations are not commutative: if the same rotations are performed in a different order, the resulting orientation will be different (Figure 5.3, sequence B).

Control of Two Rotational DOFs with an Input Device with Two DOFs

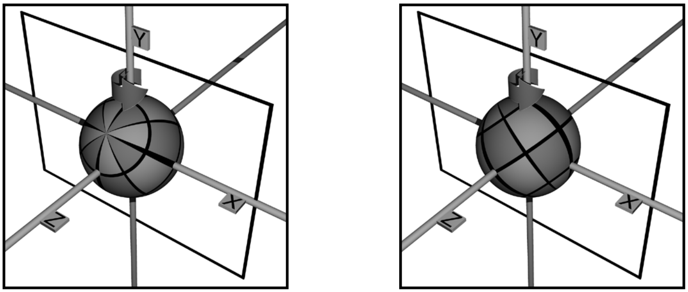
When using an input device with two DOFs it would make sense to allocate these to two of the three rotational DOFs, rather than just control one of the three rotational DOFs. Many programs make use of this feature. To access all three rotational DOFs it is necessary to swap one of the two rotational DOFs allocated to the two DOFs of the pointing device for another one. This is often done through modifier keys, mouse buttons or on-screen buttons.

Chen (1988, 1993) proposes an interface for pointing devices with two DOFs in which the virtual object is encased in a virtual sphere. A similar interface is used in the 'Scene Viewer' utility by Silicon Graphics. These interfaces use a 'glass ball' paradigm. On-screen movements are converted into movements over the surface of the sphere which allows the user to choose an arbitrary axis of

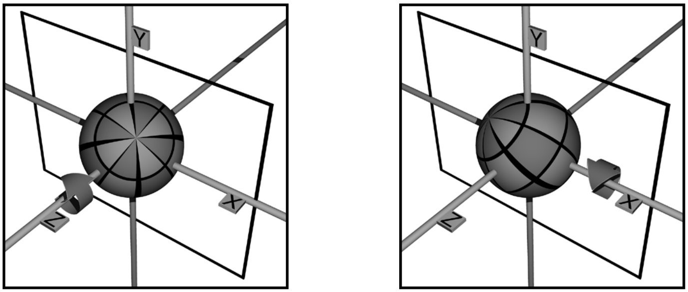
Sequence A **Sequence B**



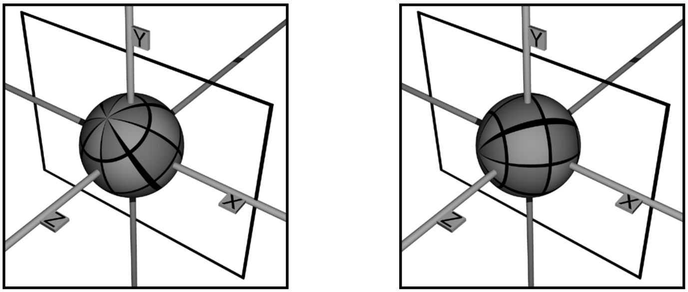
step 1: rotate x-axis step 1: rotate z-axis



step 2: rotate y-axis step 2: rotate y-axis



step 3: rotate z-axis step 3: rotate x-axis



end result end result

Figure 5.3
The same rotations in two different orders.

rotation through the centre of the sphere with an input device with two DOFs. While it is possible to choose an arbitrary axis of rotation, it is not possible to instantaneously switch to rotation about the axis connecting pointer and centre of the sphere without releasing and repositioning the pointing device.

Rotation by Means of Input Devices with Three DOFs

Trackballs are available which provide simultaneous three-axes control, thus offering three DOFs. Beaton et al. (1987) include such a trackball in one of their experiments, though it deals with translation in three dimensions rather than rotation. Buxton (1986) describes how difficult it is to physically twist a 3D trackball while rolling it, since - for constructional reasons - only part of the top hemisphere is exposed. Not all degrees of freedom can be accessed at the same time, for which reason Buxton uses the term 2+1D.

Rotation by Means of Input Devices with Six DOFs

Zhai and Milgram (1993a, 1993b, 1994) propose a taxonomy for six DOFs input devices. They conclude that for a positioning task, as opposed to a pursuit task, an isotonic[2] position[3] device offers the best performance, followed by, in decreasing order of performance, isometric[4] rate[5], isotonic rate and isometric position. The superiority

---

[2] An input device is referred to as isotonic when its use requires the shortening of muscle fibres without significant increase in muscle tone whilst not offering any significant resistance. An example of an isotonic input device is a position sensor. When the user moves such a device around, he needs to contract his muscles to create the movement, but the position sensor does not offer any great resistance and therefore the user needs to exert only little force.

[3] Refers to the mapping relationship between the user’s limb and the resulting movement of an object being manipulated. With position control, the transfer function is a pure gain: the position of the object is directly coupled to the position of the user’s limb.

[4] An input device is referred to as isometric when its use requires a marked increase in muscle tension due to resistance, without a significant shortening of muscle fibres. An example of an isometric input device is a force stick. When the user pushes against the stick it hardly moves and as a result muscle length hardly changes. However, the resistance does cause a distinct increase in muscle tension.

[5] Refers to the mapping relationship between the user’s limb and the resulting movement of an object being manipulated. With rate control, the transfer function is a first order time integration: the velocity of the object is coupled to the position of the user’s limb.

of isotonic position over isometric rate disappeared after 20 minutes of practice. Ideal position control is superior to ideal rate control (Kim et al., 1987) and is recommended for small-workspace telemanipulation tasks. McKinnon and King (1988) note that the absence of kinaesthetic feedback in isometric control results in a tendency to over-control, particularly in stressful situations. Zhai and Milgram (1993a, 1993b, 1994) also note that some of the comparative disadvantages of isometric controllers which they found in the earlier phases of their experiments may recur, either under stress or in the presence of secondary tasks. These results are important since of all the six DOFs input devices available, the force-stick, which is an isometric rate controller, appears to be the most popular for use with desktop VR systems. The force stick may be technically elegant, but it does not offer the best possible performance to the end user.

Proposed interface

A system is proposed in which the virtual object is encased in a translucent virtual sphere. Rotation of the virtual object is through rotation of this virtual sphere by means of the finger tips.

One advantage of rotating the sphere rather than the object is that the user can judge more easily whether he is in contact with the sphere than with the object. This is important since often tactile and force feedback are lacking[6]. A second advantage is that it is computationally less intensive to register contact between fingers and a sphere than between fingers and an irregular object. The sphere could be made visible only during rotation so as not to unnecessarily obstruct the view of the virtual object.

The proposed interface features a form of isotonic position control, but differs from glove-based virtual reality systems in that rotation is controlled through contact with the encasing sphere rather than through grabbing the object in the middle. It differs from most other isotonic position set-ups, for example the one by Hinckley et al.

---

[6] To improve feedback the spots where the fingers are in contact with the virtual sphere could be highlighted through, for example, a change in colour.

(1994), in that the display and manipulation space are unified. Although there are other isotonic position set-ups in which this is the case (for example Schmandt,1983; Ishii et al., 1994) the proposed set-up differs in that virtual objects are manipulated through the use of the encasing virtual sphere rather than in their centres, and in that the input device has become transparent to the user. To him it appears as if he rotates the virtual sphere directly with his fingers.

The higher the number of fingers of which the position needs to be detected, the more complex, the less robust and the slower the system. However, it seems likely that a larger number of fingers will also improve user performance and comfort during rotation. An observation experiment was set up to watch subjects perform rotations of real objects. The aim of this experiment was to arrive at a well thought-out decision on the number of finger tip positions to be registered. Explanation of the possible means of hand and finger tip registration (Bröckl-Fox et al., 1994) is beyond the scope of this article.

Experiment

Experiment Design

To predict user behaviour during rotation of a virtual sphere encasing a virtual object, an experiment was set up with a real sphere encasing a real object. The object to be rotated was a transparent plastic sphere with a teddy bear inside. The teddy bear was chosen to keep the subjects from thinking in terms of axes. With a more geometric object, the axes would have been more pronounced. The subjects were presented with two spheres. On the right was the sphere to be matched, on the left the sphere to be rotated (Figure 5.4). Both transparent spheres were supported at three points by small spheres (Figure 5.5).

Subjects

Six non-paid voluntary subjects participated in the observation experiment. All subjects were right-handed, two were female, four were male. Three were members of our laboratory, the others were novices. Only one of the subjects had some experience with a 3D modelling software package.

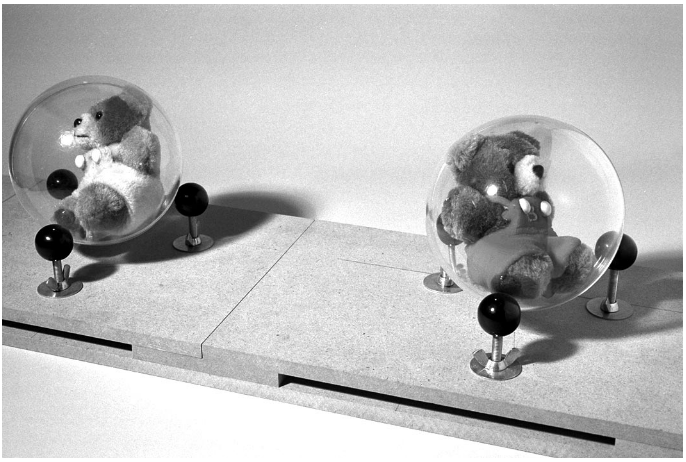


Figure 5.4
The sphere to be rotated (left) and the sphere to be matched (right).

Figure 5.5
Teddy bear in transparent sphere.

Procedure

The set-up was based around two-handed interaction. It is envisaged that an object is rotated by the non-dominant hand, while the dominant hand holds a tool for fine manipulation of, for example, the surface of the object. In accordance, subjects were asked to rotate the sphere with their non-dominant hand while holding a pencil in their dominant hand. They were asked to point this pencil at the nose of the bear to be rotated at the start of each trial. During rotation subjects were not allowed to pick up the sphere and rotate it in their hand. Subjects were asked to match the orientation of the sphere on the right as accurately as possible. There were five different conditions:

- free.
- with three fingers.
- with two fingers.
- with one finger[7].
- orthogonally restricted.

These conditions are explained below.

Free condition

This condition was included to establish how subjects would rotate a spherical object in everyday life. There were no rules as to how many fingers subjects were allowed to use. The free condition always came first, since the other conditions were expected to influence the spontaneity with which subjects would operate in this condition.

Three finger / two finger / one finger conditions

In these conditions subjects were explicitly told to rotate the sphere with the specified number of fingers only. However, the subjects were left free as to which fingers were used. These three conditions were randomised, resulting in six different orders of conditions.

Orthogonally restricted condition

This condition was included to allow a comparison with systems which feature decomposed rotation. A slotted, transparent cover (Figure 5.6) allowed rotation about three orthogonal axes only. Subjects were asked to rotate with one finger only and were able to rotate the sphere only about one axis at a time. As this condition forced subjects to think in terms of decomposed, orthogonal rota-

---

[7] One-finger rotation of a virtual object allows for interaction which is not possible with a physical object. When rotating virtual objects with one finger, it is possible to rotate the object about the finger-axis if both finger tip position and orientation are measured. With the set-up used for the real spheres this is difficult, since the friction between the finger tip and the sphere relative to the friction between the sphere and its suspension is too small to achieve rotation about the finger-axis. The one-finger condition is therefore comparable to position only detection of one finger.

Figure 5.6
Orthogonally restricted sphere.

tions, it always came last so as not to influence behaviour in the other conditions in which rotation was not orthogonally restricted.

Rotations Offered

All rotations were complex insofar as the sphere to be matched had been rotated about three orthogonal axes.

There were two classes of rotation, small and large, which differed in the amount the subjects had to rotate the sphere in order to match the sample sphere. The reason for creating these two different classes of rotation was that subjects may choose a strategy for coarse rotary adjustment different from that for fine rotary adjustment.

Small rotations were created by putting the teddybear to be matched head-up and facing the subject, and using 45° rotations about all three axes (Figure 5.7 left). Large rotations were created by putting the teddy bear to be matched head-down with its back facing the subject, and using 45° rotations about all three axes (Figure 5.7 right).

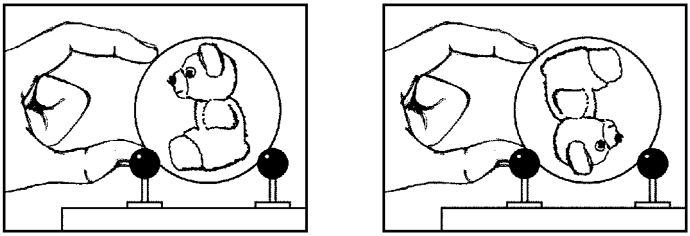


Figure 5.7
Starting point for small (left) and large rotations (right).

Two possible directions of rotation per axis for three axes results in eight possible total rotations. These existed as both small and large rotations. It was assumed that all small rotations were equally difficult and that all large rotations were equally difficult. Seven out of eight possible small rotations occurred twice, while the remaining one occurred only once per experiment, resulting in 15 small rotations (2x7+1). The same approach resulted in 15 large rotations. This total of 30 trials was spread out randomly over the five conditions, resulting in six trials per condition whereby care was taken to have three small and three large rotations per condition.

A cover prevented the subjects from seeing how the rotations were set up. A trial started as soon as the cover revealed the spheres and ended when subjects declared to have matched the spheres as well as they could[8].

Recorded Information

The experiment was recorded on video-tape for later analysis. Two cameras were used, a head mounted camera and an overview camera. The head mounted camera showed what the subjects looked at, while the overview camera gave a more stable image. By means of a mixing console the two camera images were put side by side and recorded on a single tape.

---

[8] A totally accurate match was not possible to achieve since the two teddybears were slightly different and the two spheres, being situated side by side, were viewed from different perspectives.

After a session subjects were asked to complete a questionnaire on the perceived comfort of the five conditions. Subjects were to give each condition a comfort rating varying from one (uncomfortable) to five (comfortable).

Dependent Variables

The following parameters were looked at:

1. Time spent per trial. This is a measure of how easy it is to achieve the desired rotation. Time measurement started when the subjects withdrew their dominant hand after touching the sphere with the pencil, and ended when they declared to have matched the spheres as well as possible.
2. The number of times the sphere is touched and released. This is a number of errors measure of how easy it is to achieve the desired rotation.
3. Comfort. As how comfortable do subjects experience a certain condition?

Hypotheses

Time

The three and two finger conditions were not supposed to take significantly more time than the free condition since they both offered three rotational DOFs at a time. It was expected that subjects would need significantly more time in the one finger condition because it offered only two rotational DOFs at a time. It was also expected that the orthogonally restricted condition would take significantly more time than the free condition since it offered only one rotational DOF at a time and for reasons explained in the paragraph “Rotation by means of 2D input devices” (Table 5.1).

Table 5.1 Hypotheses for Time

H_0	μ_{T3f}	$=$	μ_{Tfr}	H_1	μ_{T3f}	$\neq$	μ_{Tfr}
	μ_{T2f}	$=$	μ_{Tfr}		μ_{T2f}	$\neq$	μ_{Tfr}
	μ_{T1f}	$\leq$	μ_{Tfr}		μ_{T1f}	$>$	μ_{Tfr}
	μ_{Tor}	$\leq$	μ_{Tfr}		μ_{Tor}	$>$	μ_{Tfr}

Note: fr = free; 1f = one finger; 2f = two finger; 3f = three finger; or = orthogonally restricted; T = time.

Number of times touched and released

The three and two finger conditions were not supposed to differ significantly from the free condition in terms of the number of times the sphere was touched and released, since they both offer three rotational DOFs at a time. It was expected that subjects would touch the sphere significantly more often in the one finger condition, since it offers only two rotational DOFs at a time and therefore the subjects need to reposition their hand more often. A higher number of touch/release were expected in the orthogonally restricted condition than in the free condition since it offers only one rotational DOF at a time and for reasons explained in the paragraph “Rotation by means of 2D input devices” (Table 5.2).

Table 5.2 Hypotheses for Number of Times Touched and Released

H_0	μ_{N3f}	$=$	μ_{Nfr}	H_1	μ_{N3f}	$\neq$	μ_{Nfr}
	μ_{N2f}	$=$	μ_{Nfr}		μ_{N2f}	$\neq$	μ_{Nfr}
	μ_{N1f}	$\leq$	μ_{Nfr}		μ_{N1f}	$>$	μ_{Nfr}
	μ_{Nor}	$\leq$	μ_{Nfr}		μ_{Nor}	$>$	μ_{Nfr}

Note: N = number of times touched and released.

Comfort rating

Subjects were expected to rate the three finger condition about as comfortable as the free condition since it offers three rotational DOFs simultaneously and gives a feel of stability. Subjects were expected to rate the two finger, one finger and orthogonally restricted conditions as less comfortable than the free condition. The two finger condition - although offering three rotational DOFs simultaneously - for lack of stability, the one finger condition because it allows only two DOFs simultaneously, and the orthogonally restricted condition, since it offers only one rotational DOF at a time and for reasons explained in “Rotation by means of 2D input devices” (Table 5.3).

Table 5.3 Hypotheses for Comfort Rating

H_0	$\mu_{C3f} = \mu_{Cfr}$	H_1	$\mu_{C3f} \neq \mu_{Cfr}$
	$\mu_{C2f} = \mu_{Cfr}$		$\mu_{C2f} \neq \mu_{Cfr}$
	$\mu_{C1f} \leq \mu_{Cfr}$		$\mu_{C1f} > \mu_{Cfr}$
	$\mu_{Cor} \leq \mu_{Cfr}$		$\mu_{Cor} > \mu_{Cfr}$

Note: C = comfort.

Results

Quantitative Results

One of the subjects took a completely different approach from the other subjects. This subject - a member from our lab - would rotate for a very short period of time and then check the result for a very long period of time. It was felt that this behaviour did not represent well the behaviour people exhibit when rotating objects in daily life. The data for this subject were therefore excluded from further analysis. Among the five other subjects the pattern of behaviour was uniform.)

It was assumed that the free condition offered the most intuitive manner to rotate a spherical object. While orthogonally restricted rotation is common in 3D visualisation, an interface for rotation should be as intuitive as pos-

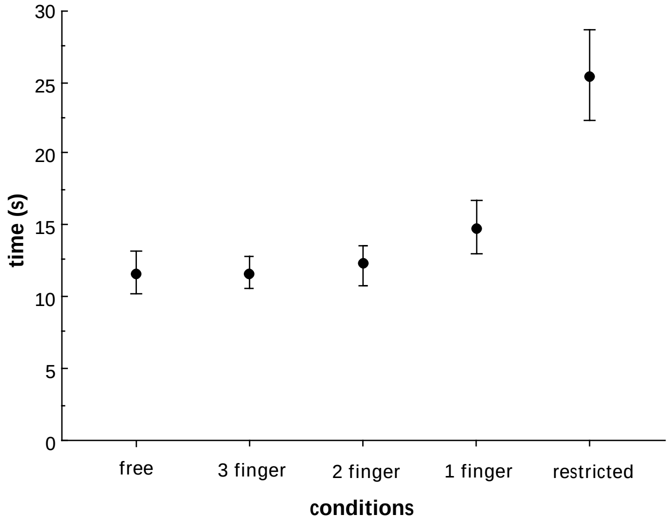


Figure 5.8
Time spent (Means, 95% confidence intervals).

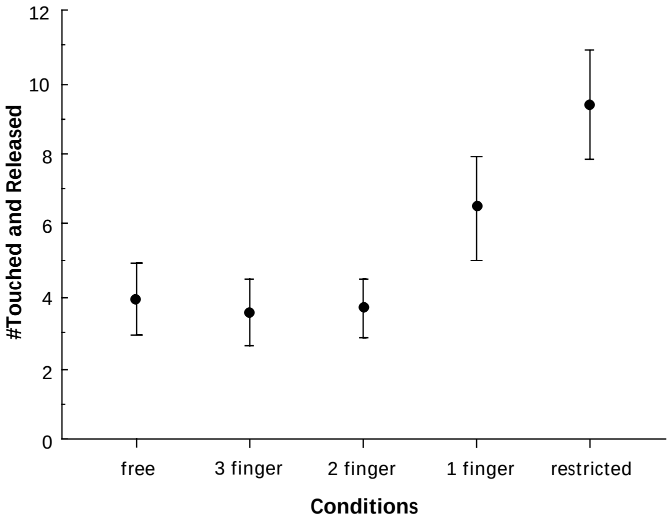


Figure 5.9
Number of times touched and released (Means, 95% confidence intervals).

sible and therefore the free condition was taken as the bottom line against which performance in all other conditions was compared.

Time — Both the time differences 'three fingers - free' and 'two fingers - free' were not significant. However, there was a significant difference in the amount of time which subjects use in the one finger condition as compared to the free condition (t-test, $p<0.05$), and in the orthogonally restricted condition as compared to the free condition (t-test, $p<0.001$) (Figure 5.8 and Table 5.4).

Number of times touched and released — Both the differences 'three fingers - free' and 'two fingers - free' with respect to how often the sphere was touched and released were not significant. However, there was a significant difference between the one finger condition and the free condition (t-test, $p<0.05$), and between the restricted condition and the free condition (t-test, $p<0.001$) (Figure 5.9 and Table 5.5).

Comfort rating — The three finger condition was the only condition which was rated not significantly different from the free condition. The two finger and the one finger condition were rated as significantly less comfortable than

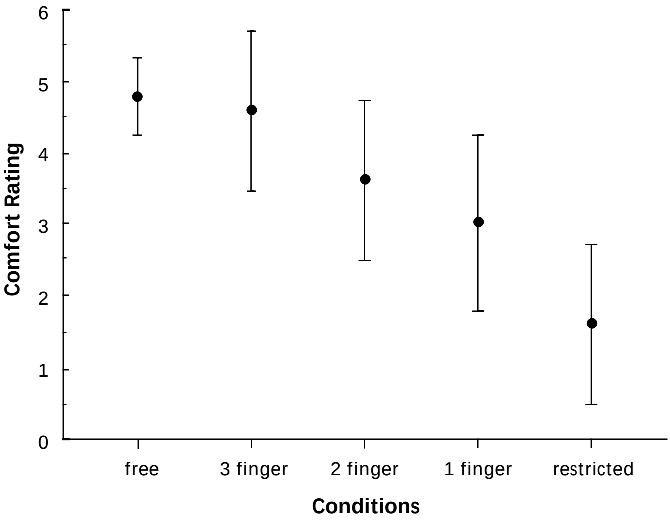


Figure 5.10 Comfort (Means, 95% confidence intervals).

Table 5.4 Extra time spent relative to free condition (t-test)

	free rotation
three finger	-0.02
two finger	-0.62
one finger	-3.14*
orthogonally restricted	-13.82**
*p<0.05, **p<0.001	

Table 5.5 Number of times touched and released relative to free condition (t-test)

	free rotation
three finger	0.37
two finger	0.23
one finger	-2.53*
orthogonally restricted	-5.47**
*p<0.05, **p<0.001	

the free condition (t-test, $p<0.05$). The restricted condition was also rated as significantly less comfortable than the free condition (t-test, $p<0.001$) (Figure 5.10 and Table 5.6).

Table 5.6 Comfort rating relative to free condition (t-test)

	free rotation
three finger	0.2
two finger	1.2*
one finger	1.8*
orthogonally restricted	3.2**
*p<0.05, **p<0.001	

Qualitative, Observation Results

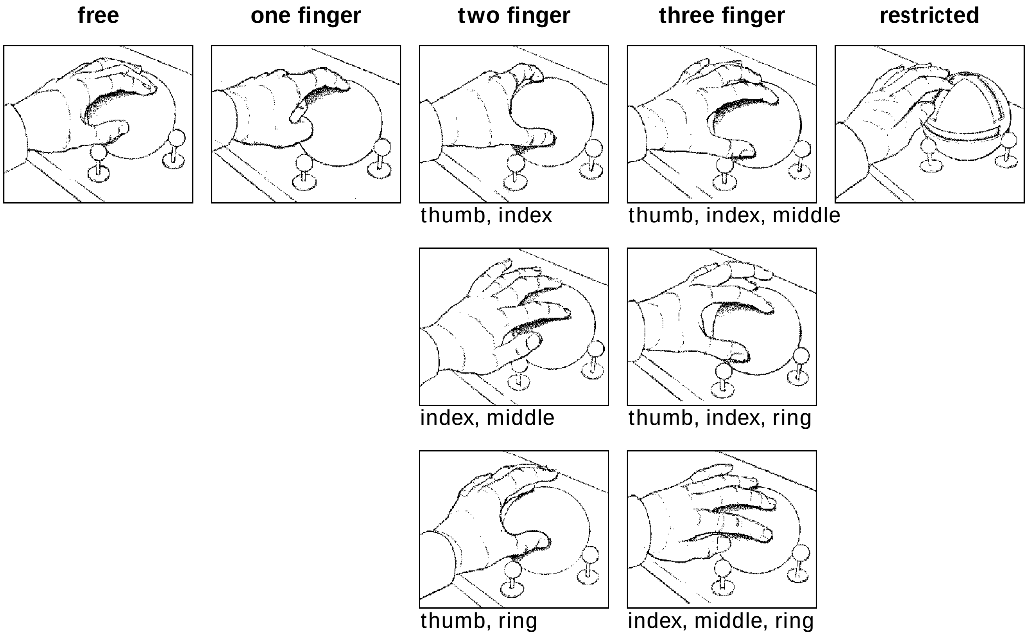
Free — In this condition there was a fair amount of interaction between the fingers, some of the fingers stayed in contact with the sphere, while others were released: a kind of “walking with the fingers” action.

Three finger — Subjects were free to use any three fingers for rotation. Not all subjects used the same fingers. For rotation with three fingers, one subject used thumb, index finger and ring finger, one used index finger, middle finger and ring finger whilst the remaining three used thumb, index finger and middle finger (Figure 5.11).

In this condition there was very little interaction between the fingers: the position of fingers relative to one another stayed almost the same. This may be the result of explicitly asking the subjects to rotate the sphere with three fingers.

Two finger — Again, subjects were free to use the fingers they preferred and not all subjects used the same fingers. For rotation with two fingers, one subject used thumb and ring finger, two used index finger and middle finger and the remaining two used thumb and index finger (Figure 5.11).

The subjects who used index and middle finger exhibited more interaction between fingers - a kind of “walking” action - than the other subjects.



free

one finger

two finger

three finger

restricted

thumb, index

thumb, index, middle

index, middle

thumb, index, ring

thumb, ring

index, middle, ring

Figure 5.11
Finger combinations which were used in the five conditions.

One finger — In this condition all subjects used their index finger (Figure 5.11). The amount of rotation between starting and stopping appeared to be quite small. Subjects rotated a small amount, judged the outcome, continued along the same lines when satisfied or took another approach when unhappy with the outcome.

They appeared to spend much time in contact with the sphere without rotating.

Orthogonally restricted — As with one finger rotation, subjects appeared to make very small adjustments rather than one large adjustment in one go. It also appeared that subjects would hesitate more often than in the other conditions. They would move their hand towards a slot as if thinking: “I now have to rotate around that axis”, and then suddenly stop, as if changing their minds, and instead move towards one of the other slots.

They appeared to spend much time in contact with the sphere without rotating.

Subjects Comments

During completion of the comfort questionnaire, two subjects commented on how they were tempted to 'cheat' in the orthogonally restricted condition. Although they were explicitly told not to rotate the sphere through a slot perpendicular to that slot, they were highly tempted to do just that.

Discussion

When designing an interface in which virtual objects can be rotated with the finger tips, for technical reasons the number of finger tips to be measured should be minimised. This experiment was set up to investigate how much time and how much effort it would take to rotate a sphere with different numbers of fingers, and how comfortable subjects would rate these conditions. The results show that the number of simultaneously accessible DOFs is an important factor for quick and intuitive rotation. Both from the time spent rotating and the number of touch/release it can be seen that the number of rotational DOFs should be three. This can be realised by two and by three fingers. Types of control which offer less than three rotational DOFs simultaneously fall behind in performance compared to those which do offer three DOFs simultaneously.

It might be dangerous to generalise over five subjects only. However, inspection of the intra-subject variances showed the results of the subjects to be similar.

Of all conditions, orthogonally restricted rotation offered the least performance in terms of time, touch/release and comfort. Three characteristics of subject behaviour were observed which suggest that this type of interface is not intuitive. Firstly, the subjects commented on being tempted to cheat, i.e. to attempt to circumvent the restrictions. Subjects appeared hampered by the restrictions which this set-up imposed upon them. When rotating objects on a flat screen these restrictions are not as apparent as in the experimental set-up in which a 3D object was physically present. The physical object clearly showed the desired way of rotation while the slotted cover inhibited those movements. Secondly, subject hesitation was more pronounced in the orthogonally restricted condition than in the other conditions. Thirdly, the stroke of rotation was small compared to the other conditions. All three of these observations may be attributed to the decomposed nature of a one rotational DOF interface: subjects appeared to be aware that in the sequence of rotations about three orthogonal axes, all three rotations influence one another. Subjects also appeared to realise the importance of “how they started off”: since they had problems visualizing the cumulative results of the individual rotations they started to hesitate, and to minimise the risk of a big overshoot they rotated the sphere by small amounts.

In contrast with the objective performance measures, for experienced comfort it was not merely the number of simultaneous DOFs which counted. While two finger control and three finger control both offer three DOFs simultaneously, two finger control was rated as less comfortable than free rotation while three finger control was not. This is likely to be the result of two finger control being perceived as less stable than three finger control.

In the two finger and three finger condition, the preferred fingers differed from user to user. Ideally, the finger measuring system should be able to cope with user preference.

Conclusions

Although all 3D visualisation packages offer rotation with three degrees of freedom, it is the number of rotational DOFs which can be controlled simultaneously which counts for intuitive rotation of virtual objects.

Decomposed rotation about orthogonal axes, which offers one DOF at a time, is a popular way of rotating 3D objects. The experiment showed that, while decomposed rotation may be an indispensable feature since it allows rotation about a single axis without influencing rotation about the other axes, it results in cumbersome interaction.

For intuitive interaction the subject’s proprioceptive information should be consistent with the visual information about the rotation of the object. In the experiment this was the case in all conditions, regardless of the number of DOFs. However, with computer input devices with less than three DOFs this is often not the case. When the same input device movement is remapped to virtual object rotation about a different axis time and again, the user can never benefit from proprioceptive information which matches his visual information. Therefore, in an actual human-computer interface, switching from rotational control with one or two simultaneous DOFs to three, could result in a more pronounced improvement in performance than in this experimental simulation.

Furthermore, even when using input devices with three rotational DOFs there remains a separation between the display space with the virtual objects, and the manipulation space with the input device. By allowing the user to rotate objects with his finger tips, a feel of direct manipulation can be achieved whereby the proprioceptive and visual information about the user's hand coincide with the visual information on the rotation of the virtual object.

Refining the brief

In order to achieve intuitive input for a medical 3D system the input devices need to be carefully considered. In the experiment described above it was shown that decomposed rotation, which is often used in current 3D systems, results in poor performance and perceived comfort. The user should not be forced to think in terms of a Cartesian coordinate system. Of course the number of degrees of freedom is not the only consideration. At least as important is a form of input which is in tune with the skills of the surgeon and radiologist. The following points are therefore added to the brief:

- if possible decomposed rotation should be avoided
- the method of input must be in tune with the skills of its user

In the next chapter...

A movement parallax based system named Cubby will be introduced in which the display and manipulation space are unified. Through the use of multiple, orthogonal screens the user is given a much larger viewing angle than is the case with single screen set-ups which offer a united display and manipulation space. Cubby offers the possibility for direct instrumental manipulation of virtual objects.

6

Introducing Cubby

Summary

In this chapter[1] a desktop VR system named Cubby will be introduced. Starting from the concepts presented in chapter four and the conclusions about input in chapter five, the reasons why Cubby came to be are explained. A technical description of Cubby is given. Cubby is compared to CAVE, a technically related immersive VR system. Finally, the possibilities of Cubby as a medical virtual environment are described.

Introduction

In the previous chapter it was concluded that the number of accessible degrees of freedom during input has a significant influence on both performance and perceived comfort when rotating virtual objects. From the literature on input devices for 3D it was found that isotonic, position devices give better performance than isometric, rate devices. The method of input must be in tune with the skills of the user.

With systems which make use of isotonic, position input devices a 1:1 coupling between the manipulation space and the display space can be realised. In other words, a virtual object coupled to an input device will follow both the position and the orientation of that input device. However, even with such systems (Hinckley et al., 1994; Suetens et al., 1988) the display and manipulation spaces remain physically separated. As the user looks at the display which shows both the virtual object he wishes to manipulate and a cursor representing the input device, he does not see his own manipulating hand at the same time. In order to con-

---

[1] Two articles based on this introduction to Cubby were published, one in the proceedings of Medicine Meets Virtual Reality (Djajadiningrat et al., 1997a), the other in the journal Displays (Djajadiningrat et al., 1997c). The former is targeted specifically at medical applications, while the latter gives a general overview.

trol the cursor and thereby the virtual object, the user has to convert the desired cursor movements into hand movements. If the virtual objects could be manipulated directly, a more intuitive interface could be achieved. This requires the virtual objects to be accessible to the hand or to an instrument. Therefore, the virtual objects need to be depicted in front of the screen. With movement parallax, it is in fact possible to achieve this. Still, with systems which make use of movement parallax the virtual scene is often displayed totally or at least partly behind the monitor screen and it thus becomes inaccessible. The reason for depicting the virtual objects at least partly behind the monitor screen is to make the best use of the other main benefit which movement parallax offers: the possibility to explore. The user can move around in front of the monitor screen to view the virtual scene from various points of view. However, when the virtual scene is put completely in front of the monitor, the user's freedom of movement is considerably reduced. This is illustrated in Figure 6.1 which

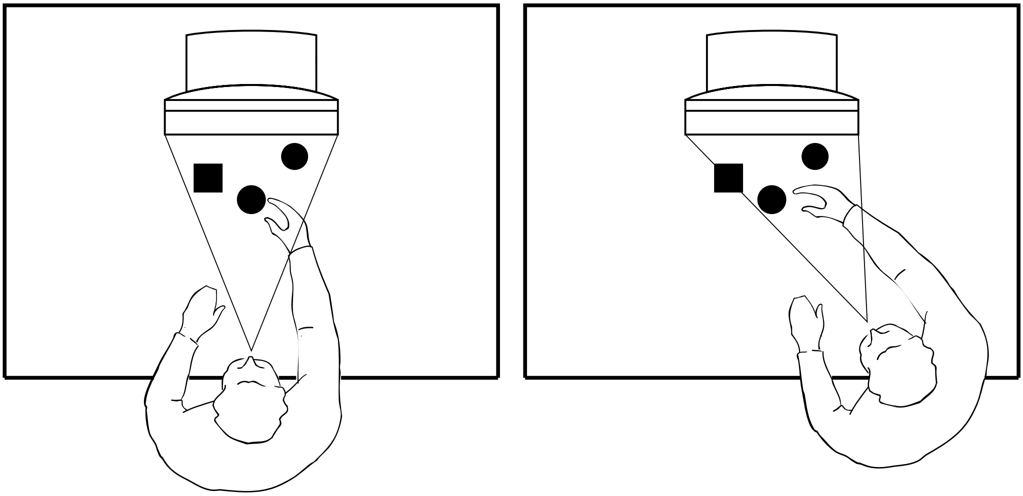


Figure 6.1
As the user moves to the right, the left-most virtual object disappears from the display area of the monitor.

shows a top view of an observer looking at a monitor with a virtual scene completely in front of it. When the observer moves from the neutral position to the right, the left most virtual object falls off the display area of the monitor and the 3D impression collapses.

The Cubby Concept

Although it is possible to alleviate the aforementioned clipping problem by choosing a monitor with a larger display area, a more effective way is to add a second monitor perpendicular to the first one. This is shown in Figure 6.2. Now as the observer moves to the right, the left most virtual object disappears from the original monitor but reappears on the second monitor. The angle over which the observer can move before clipping occurs is considerably increased.

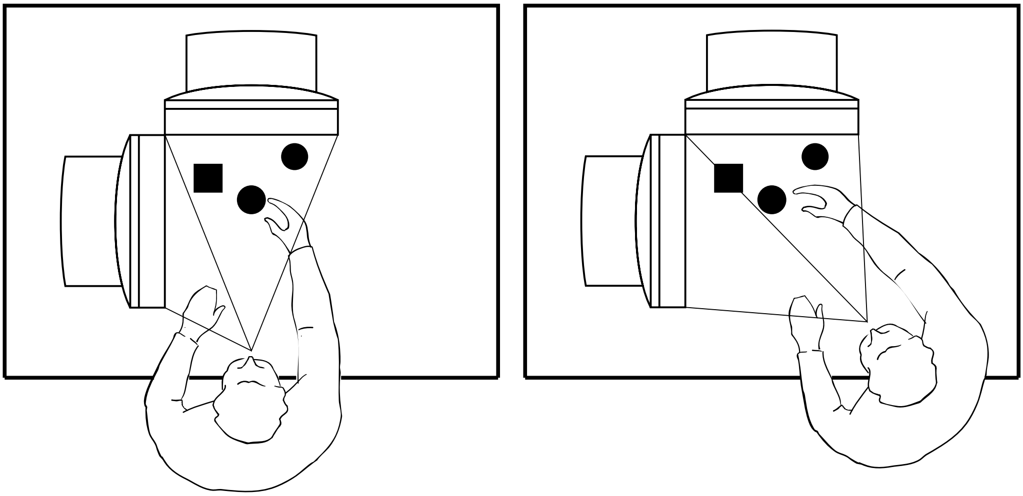


Figure 6.2
As the user moves to the right, the left-most virtual object disappears from one monitor but reappears on the other.

When a similar approach is taken to clipping under vertical observer movement, a set-up with three monitors is the result. Clipping under vertical movement is less a problem than clipping under lateral movement, since the

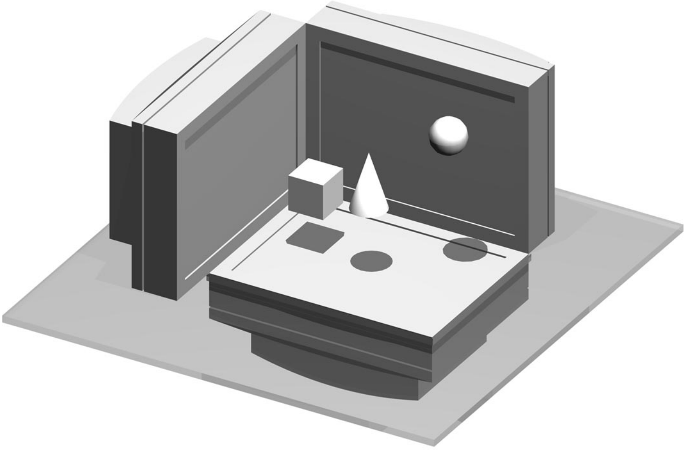


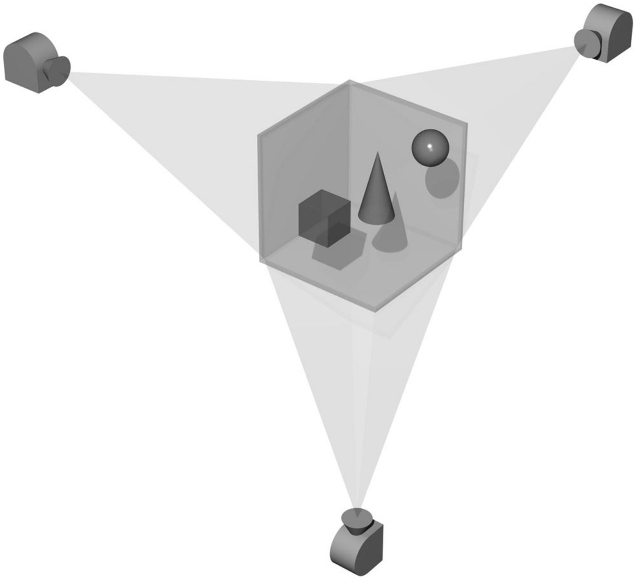
Figure 6.3
By adding a third, horizontal monitor, a ground plane is created on which virtual objects can stand and cast shadows.

observer movement is more limited in the vertical direction, especially when seated. However, an advantage of a three screen set-up over a two screen one is that the former offers a ground plane on which objects can stand and cast shadows. Shadows cast onto a horizontal plane resemble lighting conditions in everyday life more closely than shadows cast onto vertical planes.

The problem with conventional monitors is that it is not possible to make their display areas match seamlessly as they always have a bezel. This results in a non-imageable area (Figure 6.3).To overcome this problem back projection screens can be used instead (Figure 6.4).

By coupling the parallax shifts on all three back projection screens to the head movements of the observer it should be possible to create the illusion that the virtual objects stand within the concave, cubic display space. If the virtual objects are displayed in front of the screen they become accessible for direct manipulation, without the screens forming a barrier.

Figure 6.4
Three back projection screens form a concave cube in which the virtual objects appear



Cubby Prototypes

A quick, cardboard prototype with three perpendicular screens was made, using pre-rendered images. A diagram of the set-up is shown in Figure 6.5. The observer's head-position was determined by means of an infra-red headtracker (DynaSight by Origin Instruments). On the basis of this information a computer (Apple PowerMacintosh 6100/60 AV) selected the nearest available pre-rendered image. The video output of this computer was split into three identical signals using a video mixer (Panasonic WJ-MX10). Each of these three video signals was sent to a video projector (Sony CPJ-100e). The original set-up built out of cardboard is shown in Figure 6.6 and Figure 6.7. Cubby uses the Fish Tank VR projection method (Ware et al., 1993) as described in chapter 2 for all three screens. Figure 6.8 shows a top and front view of Cubby for a single virtual camera location. There is a virtual camera for each projection screen image. All three virtual cameras share the same location but are oriented perpendicular to each other. The orientations of the cameras do not change with

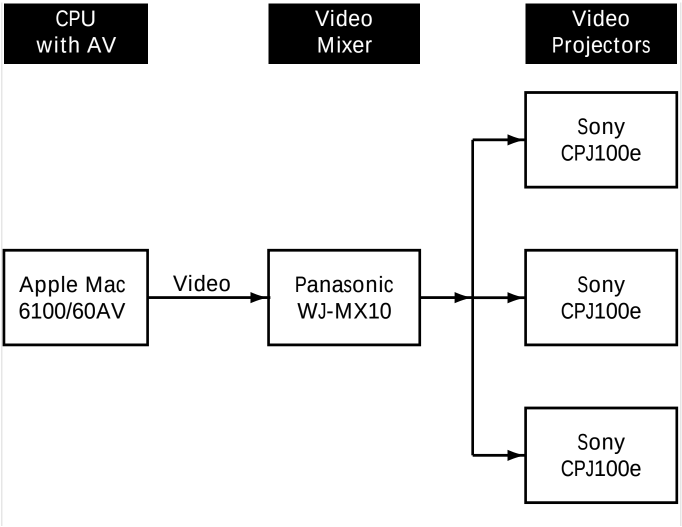


Figure 6.5
The hardware components of the first prototype using pre-rendered images.

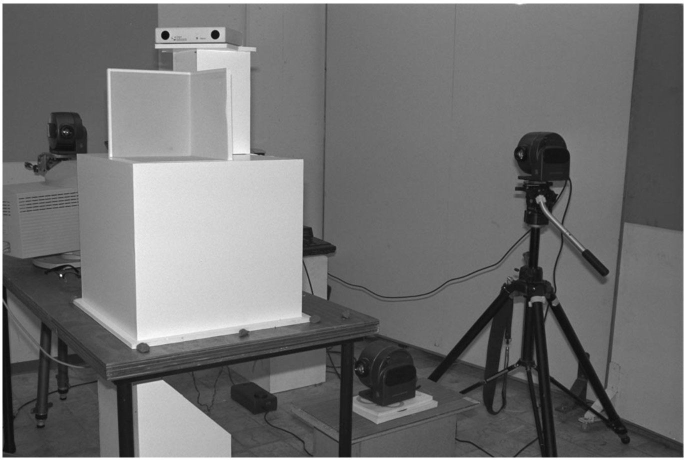


Figure 6.6
First, cardboard prototype of Cubby. The Dynasight head tracker is positioned above the display area. The projectors stand on tripods.

observer movement. Instead their lines of sight always remain perpendicular to their respective projection screen. The images (Figure 6.9) were rendered for 31 viewpoints, which lay on a single horizontal line. For each viewpoint one image had to be rendered per screen, resulting in three images per viewpoint, and 93 images in total. For each viewpoint three images were composed in an L-shape of 480x480, which fit on a 640x480 screen. The virtual scene consisted of 27 cubes arranged in a matrix of

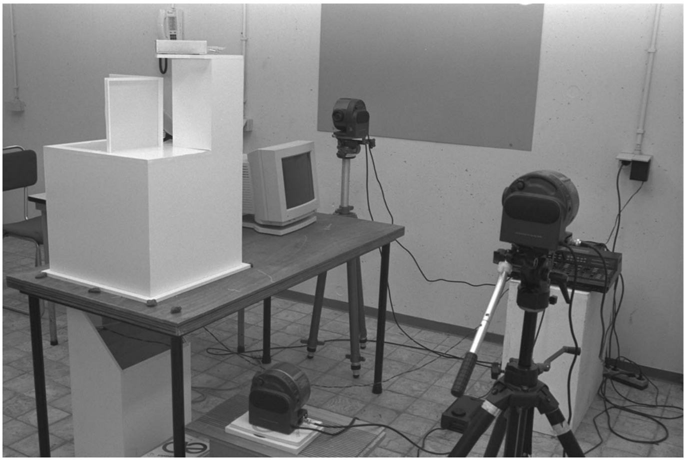


Figure 6.7
Another view of the first Cubby. Note how the horizontal projection screen is projected on via a mirror.

3x3x3. The cubes were rotated by different angles and each one had a unique texture to maximise the effect of the parallax shifts under head movements. Because there were only 31 viewpoints and because the head movements of the observer were not limited in any way, deviation from the line with these viewpoints introduced considerable distortions which would make the 3D impression collapse. However, as long as the observer kept his head close to the line for which the images were rendered, the results looked promising. Figure 6.10 shows the matrix of virtual cubes as seen by an observer.

A new set-up was built with which the projected images could be updated in real-time to the head position of the observer. A diagram of the set-up is shown in Figure 6.11. This set-up was built around a faster computer (Apple PowerMacintosh 9500/132), three graphics boards (ATI XClaim) and two 3D graphics accelerator boards (Apple QuickDraw3D accelerator boards). A 3D graphics library (Apple's QuickDraw3D) was used to facilitate programming. The monitor signal of each video board was converted to S-Video by means of a scan converter (Display Research Laboratory Televisor Zoom). Each of these videosignals was sent to a video projector (Sony CPJ-100e). To reduce the distance between the projector lens and the back projection screen the projectors were equipped with wide-angle lenses.

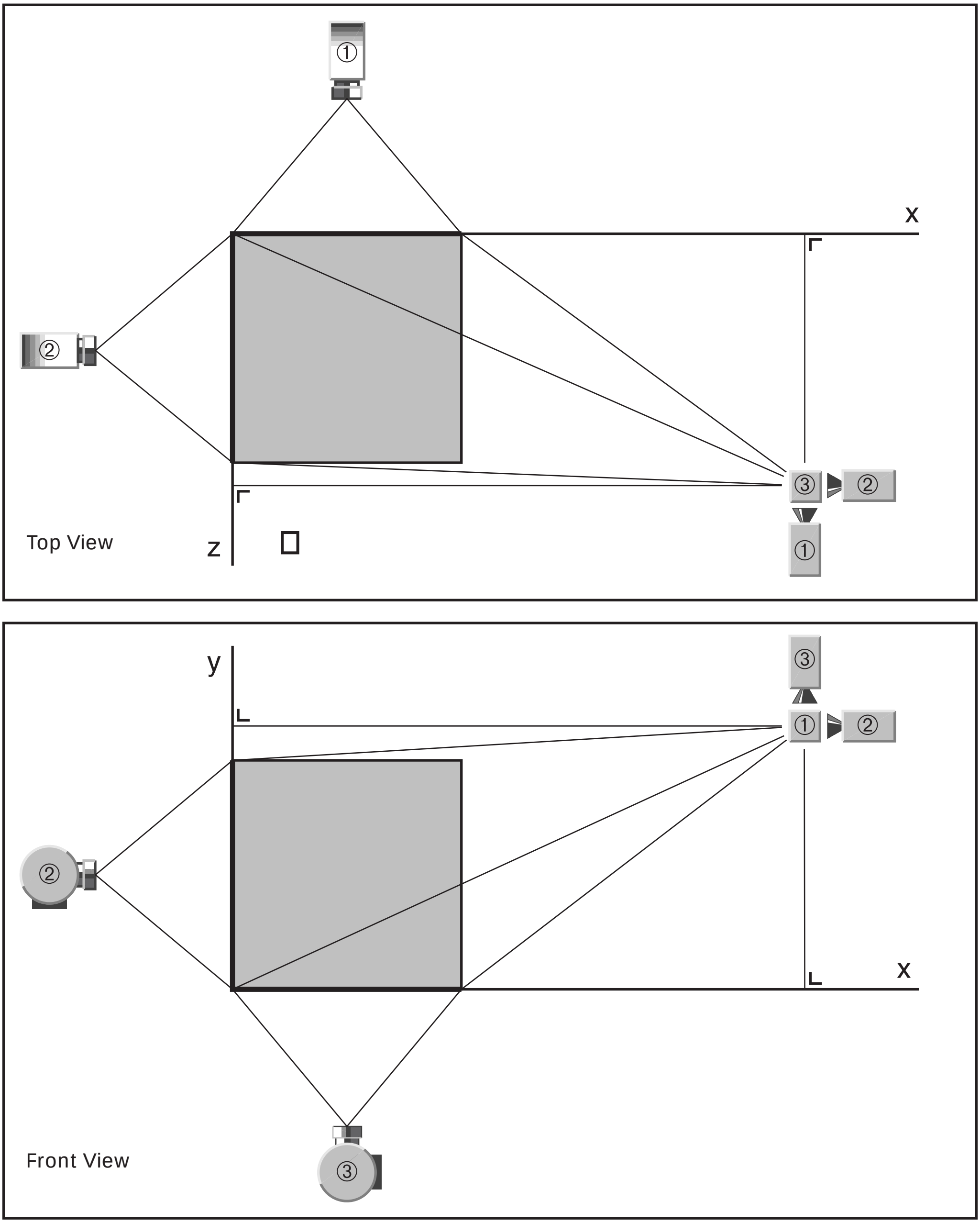


Figure 6.8
Top and front view of Cubby’s three virtual cameras.

Figure 6.9a
Viewed from the left.

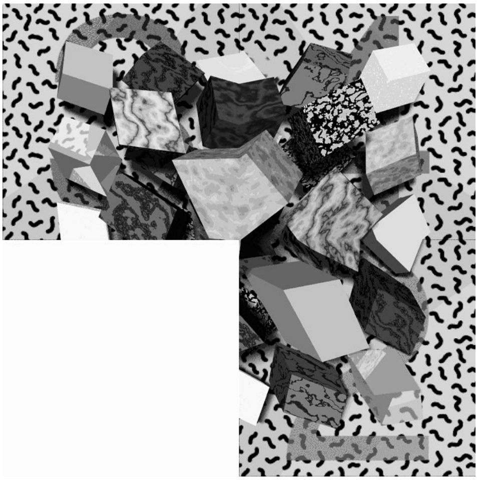


Figure 6.9b
Viewed over the diagonal.

Figure 6.9c
Viewed from the right.

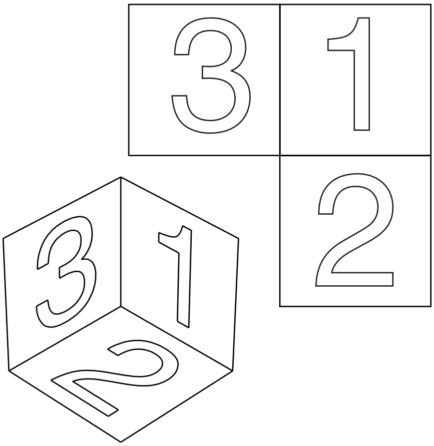


Figure 6.9d
How the three images in a set are ‘folded’ to form the back projection screens.

Figure 6.9
Figure 6.9a, b and c each show one of the 31 sets with three pre-rendered images. Each set corresponds to a viewpoint and each of the three images in a set belongs to back projection screen. Figure 6.9d shows how the images in a set are folded to form the projection screens.

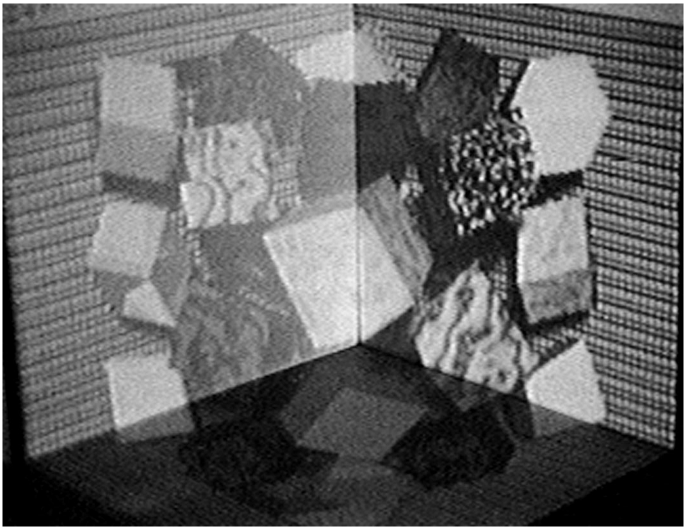


Figure 6.10
The matrix of virtual cubes as seen from the point of view of the observer.

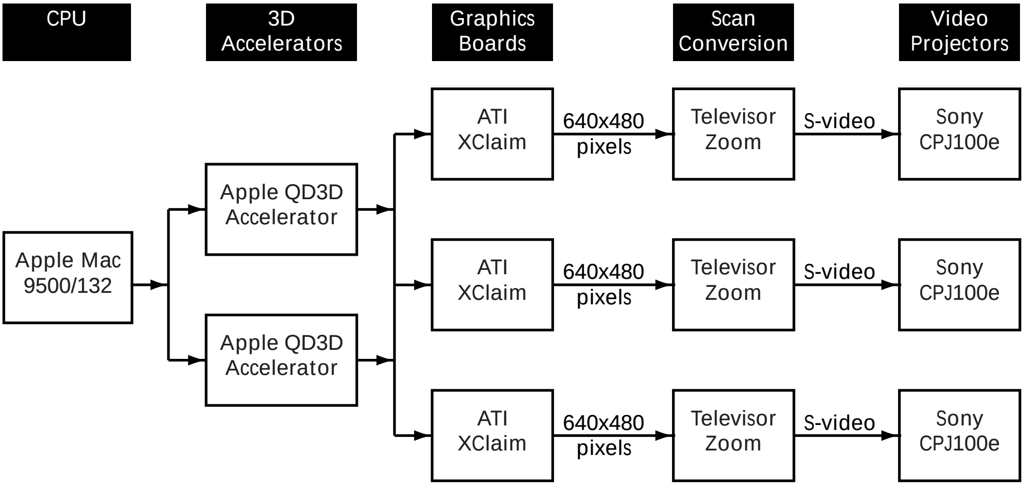


Figure 6.11
The second prototype which worked with images generated in real-time

Visualisation

Figure 6.12 shows four views of an office chair in Cubby, corresponding to different head positions. The chair is rendered at such a height that it appears to stand on the

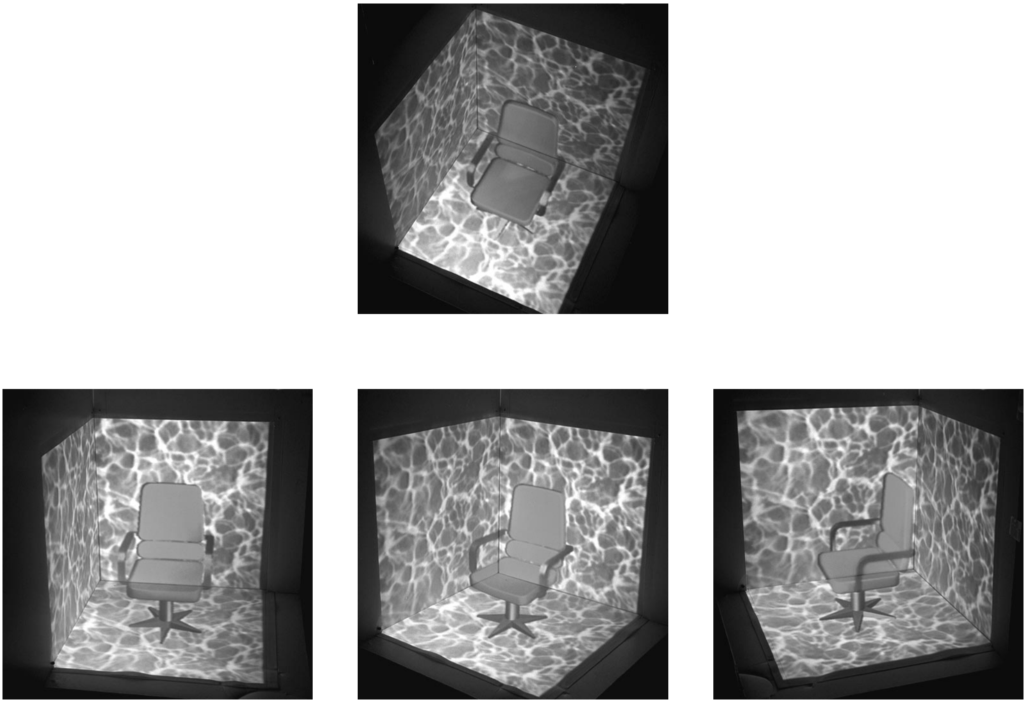


Figure 6.12
A virtual desk chair inside Cubby seen from different points of view, corresponding to different head positions.

ground plane.

Another scene which we tried in Cubby was a 3D reconstruction of the bones in a human wrist. These so-called carpal bones form a particularly complex arrangement. The geometries and spatial relationships are difficult to express adequately in two dimensional drawings or X-ray photos (Figure 6.13). The 3D meshes were generated from contours in a set of MRI scans[2]. Cubby allows the user to explore the virtual wrist by viewing it from various points of view (Figure 6.14).

---

[2] I gratefully acknowledge J.A. Snel and C.A. Grimbergen of the Amsterdam Medical Centre for supplying the contours of the carpal bones.

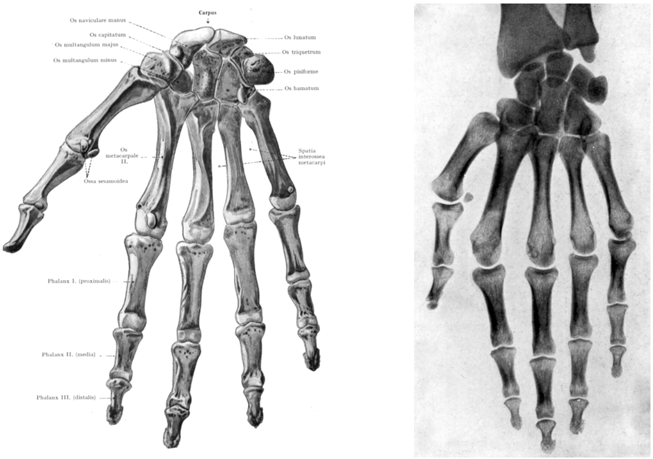


Figure 6.13
Drawing of a hand from a traditional anatomical atlas (left) and an X-ray photograph of a hand (right). The carpal bones are in the wrist shown at the top of both pictures (Spalteholz, 1954).

Manipulation

While this early version of Cubby offered only visualisation, it was felt that the open cubic space with the virtual objects lent itself well to manipulation. Although it would be possible to use Cubby in combination with glove based input, instrumental manipulation is better suited to Cubby for two reasons. Firstly, a glove negatively influences dexterity and is difficult to put on and off. Secondly, it is not possible to put a hand in between a virtual object and its projection screen as then the virtual object partly disappears.

If both position and orientation of the instrument are measured, six degrees of freedom are accessible in an intuitive manner. Consequently, the desired movement need not be decomposed in orthogonal translations and rotations. as is the case with many existing interface.

By rendering a virtual tip in line with the barrel of a physical instrument it is possible to achieve natural looking occlusion. As such a virtual tip is rendered concurrently with the scene, visual feedback on the position of the virtual tip can help the user to compensate for differences between the instrument's measured position and its true position caused by delays and errors on position measurement. An example of a medical task to which Cubby may be suited is a stereotactic task. Figure 6.15 shows a tree-like structure of blood vessels in which a tumour is present.

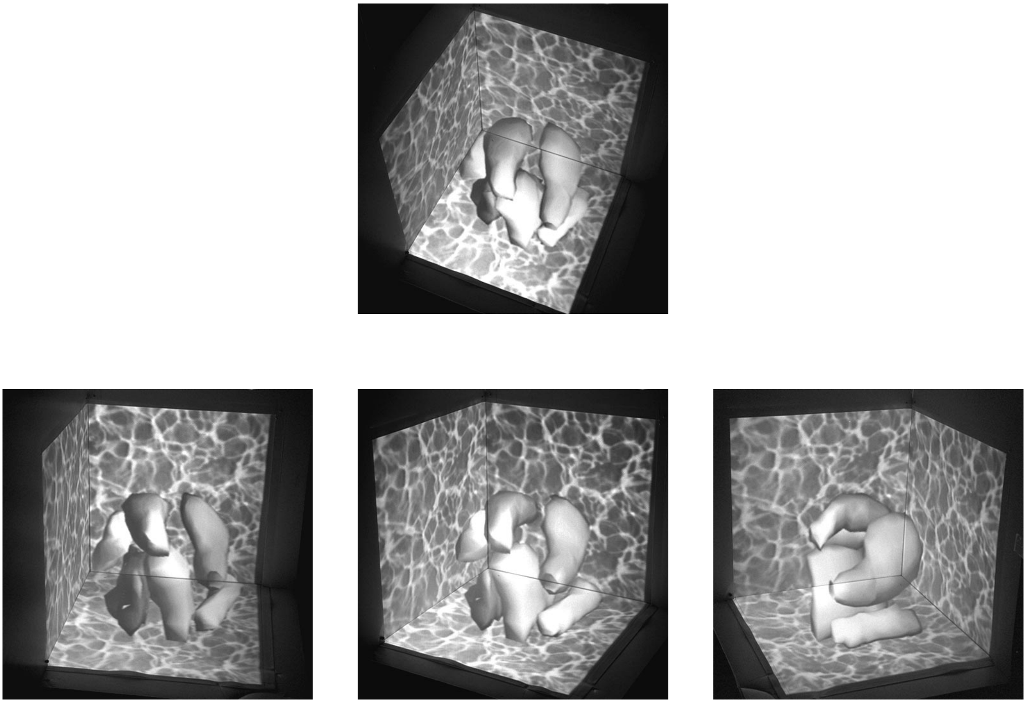


Figure 6.14
A 3D reconstruction of the carpal bones seen from different points of view, corresponding to different head positions.

Fiducial markers (Figure 1.5) can be displayed as reference points. Cubby can establish the position and orientation of the instrument relative to the stereotactic markers. Such information can be used for tracker or robot assisted surgery, or for radiation treatment. The task is to find a path to the tumour without damaging any critical structures. By first touching the tumour the user can fix one end-point of the path with the other end-point becoming attached to the instrument's tip. The user can then move the path around while looking at the scene from various points of view, thus exploring the various possibilities.

The manipulation aspects of Cubby will not be covered in detail till chapter 9. However, in the intermediate chapters, the name Cubby implies that the denoted system allows manipulation. The name Cubby is thus used to describe a desktop-sized virtual environment with the fol-

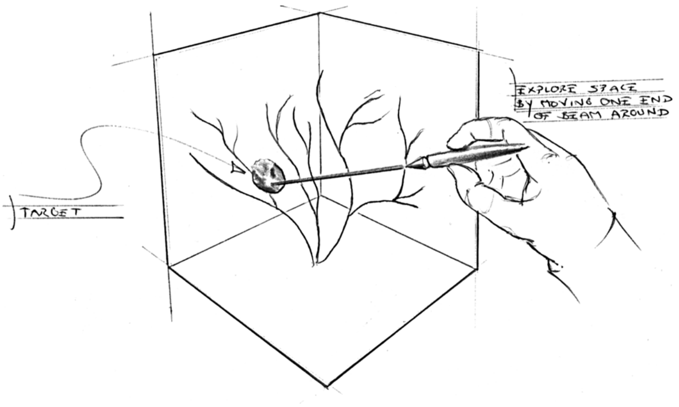


Figure 6.15
Using an instrument to probe a tree-like structure containing a tumour.

lowing characteristics. The illusion is created that virtual objects stand within a cubic display space through the use of head-coupled movement parallax, and these virtual objects can be directly manipulated with six DOFs by means of a physical instrument which has a virtual tip.

Cubby vs. CAVE

From a technical point of view Cubby is quite similar to CAVE. CAVE is a room-sized, immersive virtual environment in which, depending on the implementation, walls, ceiling and floor are formed by projection screens (Figure 6.16). CAVE was developed at the Electronic Visualisation Laboratory at the University of Illinois (DeFanti et al., 1993; Cruz-Neira et al., 1993).

Cubby is thought to be an interesting supplement to CAVE for the following reasons. Cubby differs from CAVE in that it is much smaller and non-immersive. Because of this, Cubby and CAVE invite different behaviour. CAVE offers a panoramic view and thus the observer will look around him, rotating about his axis, rather than look around an object (Figure 6.17). Cubby offers a much smaller scene and therefore the user will look around the objects in that scene. With regard to size, for many medical tasks Cubby's workspace is sufficient. Because of its smaller size Cubby is less expensive than CAVE. Not only

Figure 6.16
A CAVE

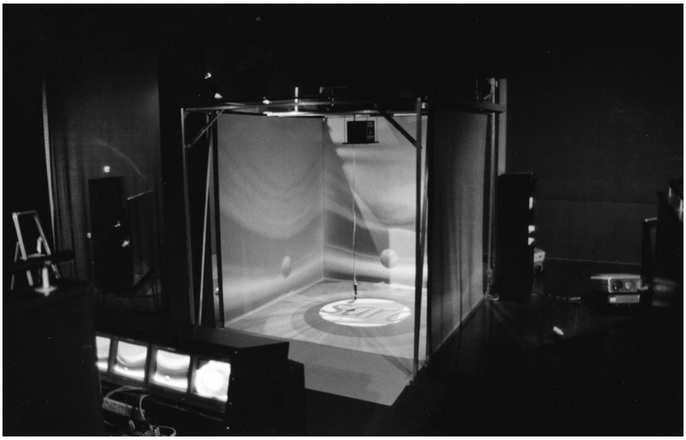


Figure 6.17
An observer inside CAVE

does the smaller size decrease the cost of physical components such as projection screens, it also decreases the cost of the required computer hardware. To achieve similar resolutions a computer driving Cubby needs to calculate fewer pixels than a computer for a CAVE. As the projection areas are much smaller the projectors can be of lesser resolution and less bright. For these size and cost reasons it may be easier to integrate a Cubby within a radiology room or operating theatre than a CAVE.

In the next chapters...

A problem with Cubby was that to a moving observer virtual objects seemed to deform. In chapter 7 the possible causes of this problem are investigated and solutions proposed.

Chapter 8 covers an experiment which tests the veridicality of depth perception in Cubby. Finally, in chapter 9 direct instrumental manipulation in Cubby is assessed.

7

Searching a Cure for Perceptual Instability

Summary

One of the first things to be noticed when the prototype of Cubby with real-time graphics was up and running was that virtual objects seemed to deform. Users described the behaviour of virtual objects inside Cubby as being ‘rubbery'. To be able to manipulate inside Cubby it is necessary for objects to appear stable to the user. This chapter describes the possible causes of the deformation of virtual objects inside Cubby, the way in which their effects were investigated, and the remedies which were taken.

Introduction

To allow a systematic approach the possible causes of deformation are divided into three categories:

1. flatness cues
2. static distortion causes
3. dynamic distortion causes

With flatness cues characteristics are meant which indicate that the virtual scene is in fact not 3D but the result of projection on a 2D screen, and which therefore counteract the 3D impression. For example, in desktop VR systems screen glare acts as a flatness cue which makes the observer aware of the presence of a monitor screen.

Under ‘static distortion' those causes are ranked which result in distortion of virtual objects when the observer is stationary. An example of such a cause is inaccuracy in the head tracking system which results in the placement of the virtual camera not corresponding to the eye position of the observer.

Dynamic distortion causes distortion of virtual objects during observer movement. Delays caused by various system components fall into this category, as they lead to the display of images which belong to a certain head position at a time when the head has already moved to a different position.

In this chapter I will report in what form these three causes of deformation manifest themselves in Cubby, how they were investigated and which steps were taken to counteract them.

Possible causes of deformation

Flatness cues

The first Cubby prototype was constructed out of foam board with the projectors mounted on tripods. While this set-up was quick to build and allowed for easy modification, it resulted in the projectors being difficult to position accurately. This resulted in various flatness cues. First, the projected images could be slightly different in size which meant that they would not quite match each other. Secondly, when the projected images did not quite meet up pronounced dark lines running along the seems could be seen, or the images could overlap resulting in a lighter region. To the user the image would seem distorted along the edges of the screens.

Another undesired effect was the cockling of the projection screens which were made of out of draughting film. This cockling resulted in distortion of the projected images.

Another flatness cues is the colour difference between the projected images. One cause for this difference is the light of the projectors being reflected in different and unpredictable ways. Another cause is that each of the LCD projectors used had a different colour bias. Yet another is that the brightness of the projected images changes with the angle between the line of sight and the screen normal.

The aforementioned flatness cues are typical for Cubby’s arrangement of three backprojection screens. In addition, there are a number of other flatness cues from which all desktop VR systems suffer. The pixels on the screens are visible and the angle a pixel takes up in the visual field changes with the angle between the line of sight and the screen normal. Because raster graphics are used the images themselves are also built out of pixels. The screen is brighter than its surroundings. Finally, since all of the virtual scene is in focus on the 2D screen, there is no depth blurring. All parts of the scene are equally in focus, no matter what their distance to the observer.

Static distortion

Causes of static distortion

As Cubby is used with monoscopic images, it is best viewed with one eye. If it is viewed with two eyes, stereoscopic information on the physical surroundings of the display space will conflict with the monoscopic virtual objects, and this conflict will act as a flatness cue. Therefore the observer is provided with a pair of glasses which blocks one eye.

One established (Pasman, 1997c) cause of deformation in fish tank VR systems is when the virtual camera position does not match the eye position of the observer. This miscorrespondence can be the result of three factors. The first is an error in the position given by the tracking system. The second is miscalibration through an inaccurate transformation of the tracker coordinate system to the display space coordinate system of Cubby. The third factor is that the position which is actually tracked is not the optical centre of the eye. For Cubby a three DOFs position tracker (DynaSight by Origin Instruments) was used which has an advantage over most currently available six DOFs trackers in that it is wireless. The Dynasight tracks the position of a reflective disc which is usually attached to the frame of a pair of glasses. Because this tracker does not supply orientation information it was not possible to accurately take into account the reflective disc-eye offset. As a compromise the eye position was estimated through use of a constant vertical offset of 30mm. As the observer moves around his/her head does not remain upright which makes this simple eye position estimate inaccurate.

Investigating static distortion

To investigate the possible misalignment of virtual camera position and eye position a somewhat unconventional approach was taken. Normally, in a calibration procedure measurements would be taken at points of which the position in the display space coordinate system is known. By combining the raw positional data from the tracking system and the desired positions in the display space coordinates the transformation matrix can be found. This transformation matrix is then used to calculate the eye

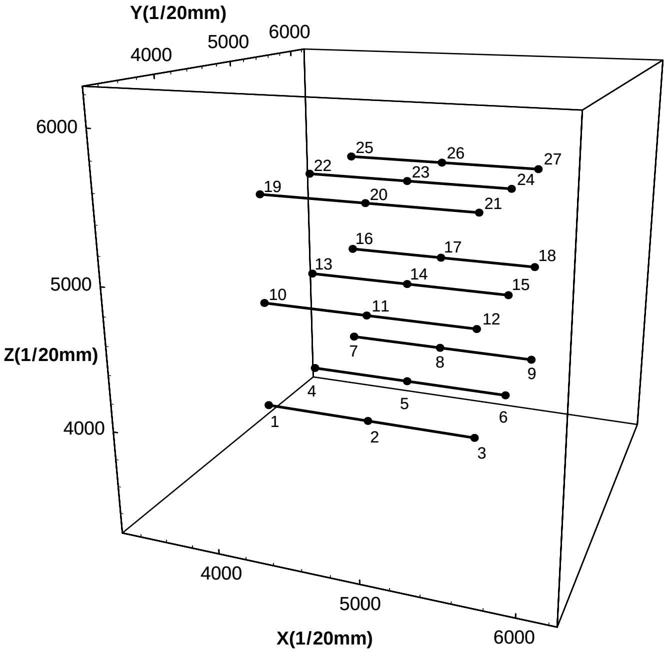


Figure 7.1
The spatial matrix of 27 virtual camera positions.

position and thus the virtual camera position in display space coordinates. The calibrated virtual camera is then used to generate the perspective images.

To calibrate Cubby, we took the opposite approach. A virtual cubic frame was built by means of a computer modelling program. Sets of three images, one for each back-projection screen, were generated for 27 different views of this virtual cubic frame. These views were arranged in a spatial matrix of 3x3x3 (Figure 7.1). It should be noted that in Figure 7.1 and the other diagrams related to the calibration procedure the axes are in Dynasight units (1 Dynasight unit = 1/20mm). When shown in Cubby the images did not vary with the head position of the observer. They were static and corresponded to only a single observer position. To act as a physical reference geometry a cubic frame was built out of wooden sticks. The physical cubic frame and its virtual counterpart were displayed simultaneously inside Cubby. For each of the 27 perspectives the observer now had to move his head to find a viewing position for which the virtual and the phys-

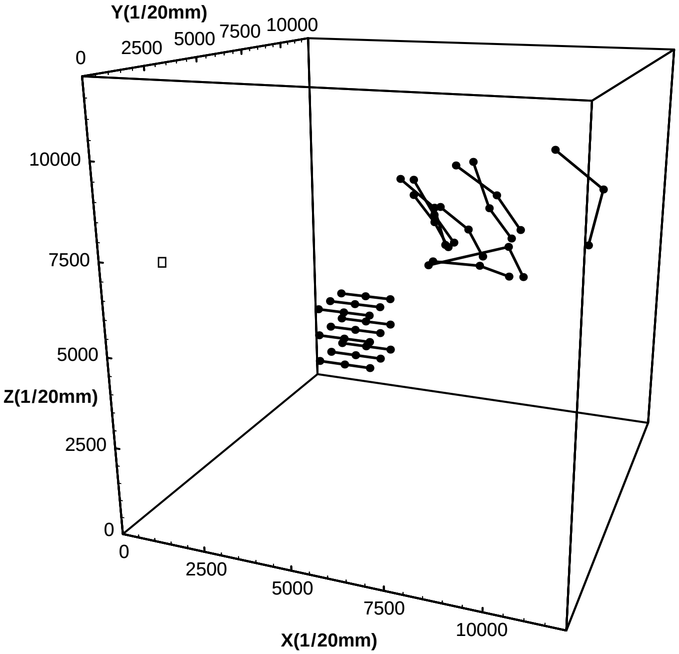


Figure 7.2
Overview showing both the virtual camera positions (cluster bottom left) and the observer's eye positions (cluster top right).

ical cubic frames coincided[1]. When the observer indicated that he had found this position, his head position as measured by the Dynasight tracker was recorded.

These eye positions, which were measured in raw Dynasight units (i.e. without a transformation applied), were compared to the virtual camera positions for which the perspectives were generated. As these virtual camera positions were arranged in nine straight, parallel lines with three positions on each line, the measured eye positions should also be arranged in straight, parallel lines. Figure 7.2 is an overview showing both the virtual camera positions for which images were generated, and the measured eye position. The two sets are offset from each other as the virtual camera positions are in the Cubby coordinate system while the eye positions are in the untransformed Dynasight coordinate system. Orthogonal views of the virtual camera positions and the measured eye positions are

---

[1] A hole was punched in the reflective disc which acted as a target for the Dynasight position tracker. By looking at the virtual objects through this disc, the difference between the tracked position and the eye position was minimised.

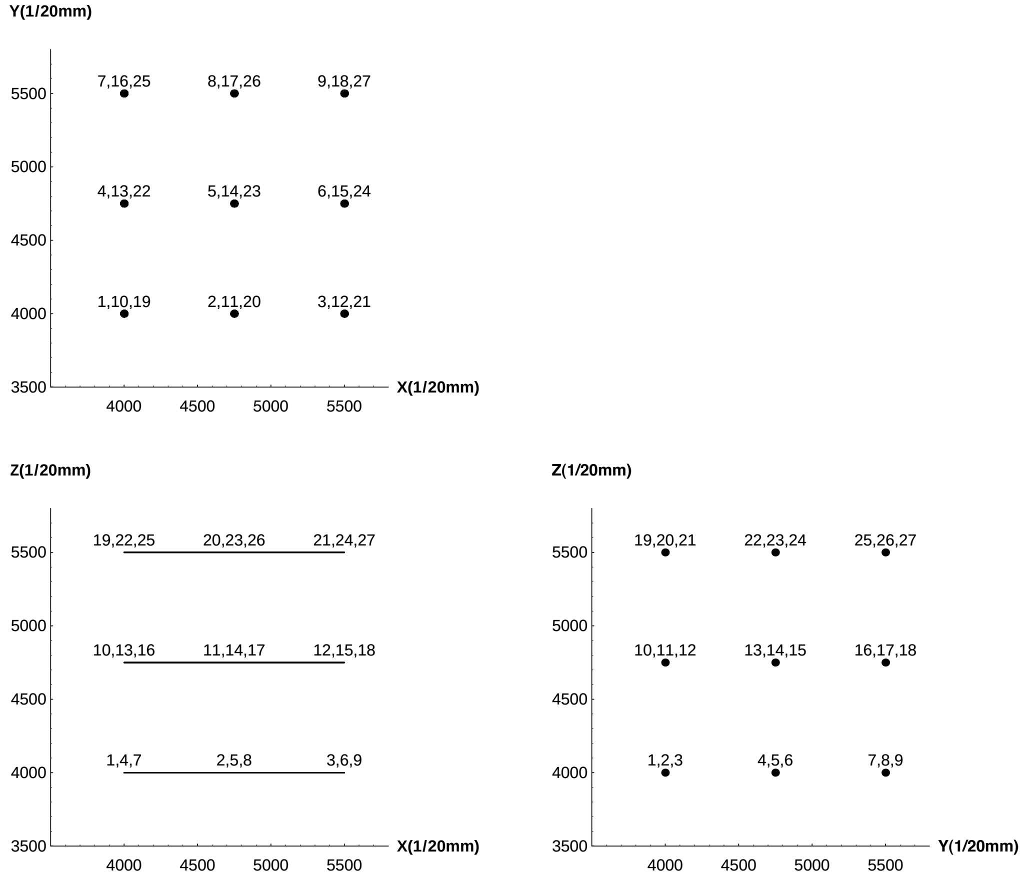


Figure 7.3
Orthogonal views of the virtual camera positions.

shown in Figure 7.3 and Figure 7.4 respectively. By rotating the straight lines which were fit to the measured eye positions to coincide with the straight lines formed by the virtual camera positions, the transformation from the Dynasight coordinate system to the Cubby coordinate system was found. This transformation was then used instead of the transformation based on measurements of the physical set-up.

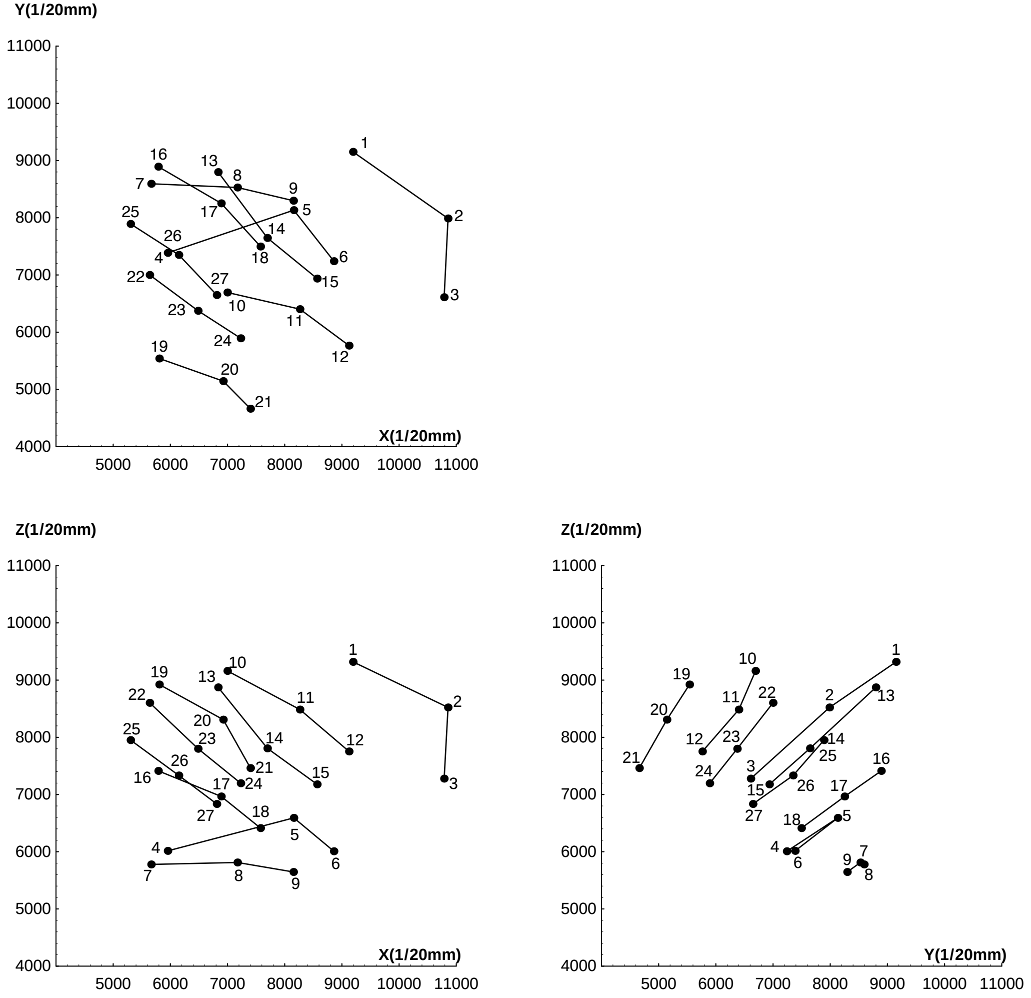


Figure 7.4
Orthogonal views of the observer’s eye positions.

The advantage of this calibration method over the previously described traditional method is that it involves a perceptual test. It would be possible that virtual objects are deformed in such a way that there is no corresponding eye position. The applied calibration procedure tests whether there is an eye position at all for which the virtual and physical cubic frames coincide.

However, both with and without the changes to the transformation matrix, the differences between the real and virtual cube when observed from various static viewpoints, were too small to fully explain the elastic band like deformation which users described. Figure 7.5 shows the real, the virtual and the real and virtual cubic frames combined for five different points of view.

Dynamic distortion causes

The difference between the real and virtual cube only was small, however, when the observer remained stationary or moved slowly. The virtual object became quite noticeably distorted under quick observer movement, clearly lagging behind the real cube. This is shown in Figure 7.6. These images were recorded by applying the Dynasight reflective disc to a video camera. This sequence makes visible what happens as the observer moves from one stationary viewpoint to another stationary viewpoint. At the starting position and the end position where the observer is stationary the physical and the virtual frame coincide. However, as the observer starts moving the virtual frame lags behind the real frame. This is most pronounced in Figure 7.6 e, f and g where the observer is moving his head with the highest velocity. As the observer slows down the difference between the real and the virtual frame diminishes until it is approximately zero when the observer is stationary. This lagging of the virtual object with respect to the physical object appeared to be the result of the delay between the moment the head position of the observer was detected and the moment the projections for that particular head position appeared on screen.

Another cause of distortion has to do with the distance of the point on a virtual object to the plane on which it was projected. This can be clearly illustrated by looking at the wire frame cube standing on the floor of Cubby. When looking at a vertical rib of this cube, the lowest point, which lies in the bottom projection screen is not affected by delay because its projection on the screen remains stationary. All other points on this rib do move with observer movements. The further away on the rib from the projection screen, the more a point must move, and the bigger the distance it appears to lag. This too can be seen in Figure 7.6. The result is a shearing effect. For example, a rectangle is distorted into a parallelogram.

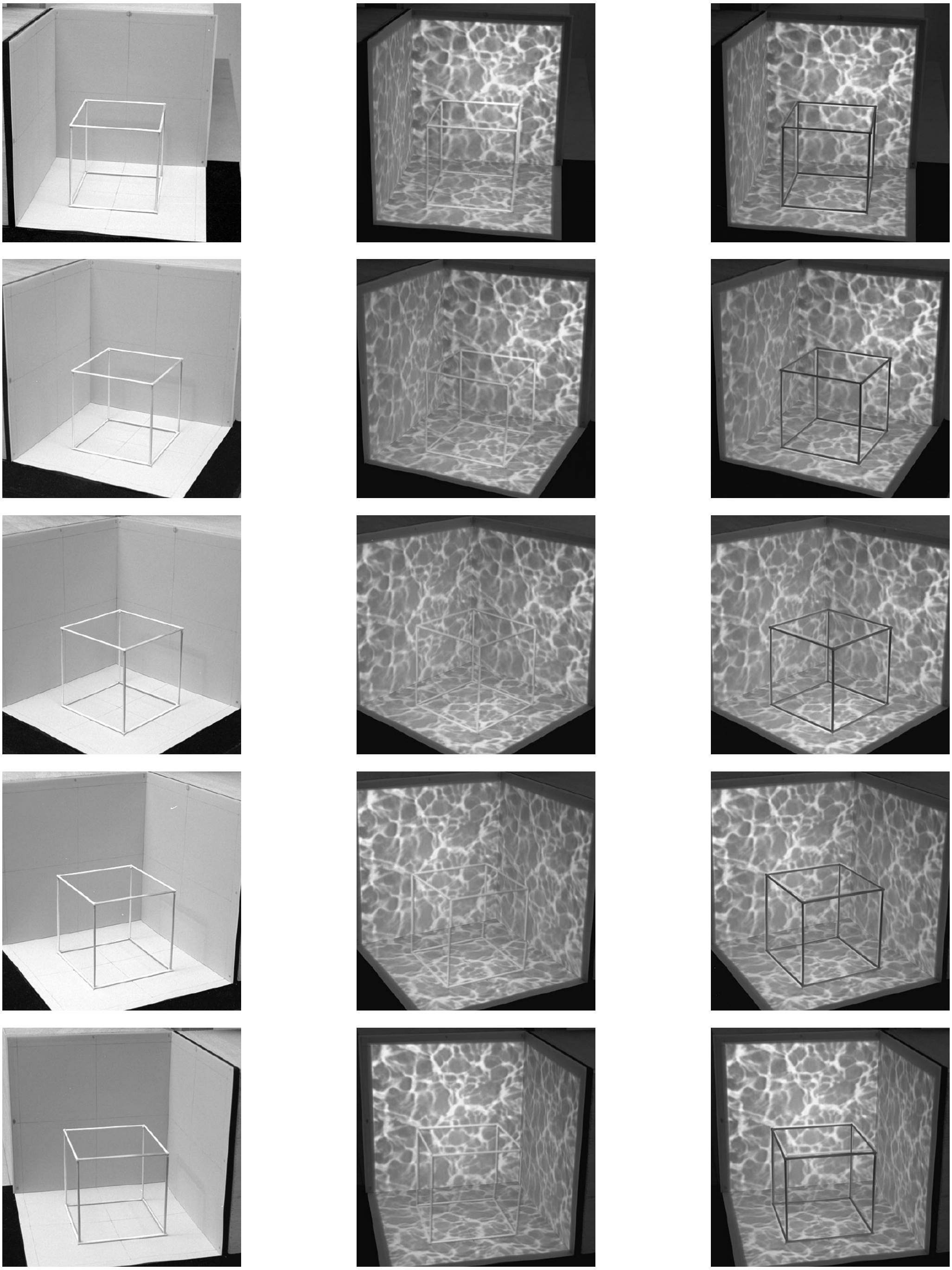


Figure 7.5
For five viewpoints the real, the virtual and the two frames combined are shown.

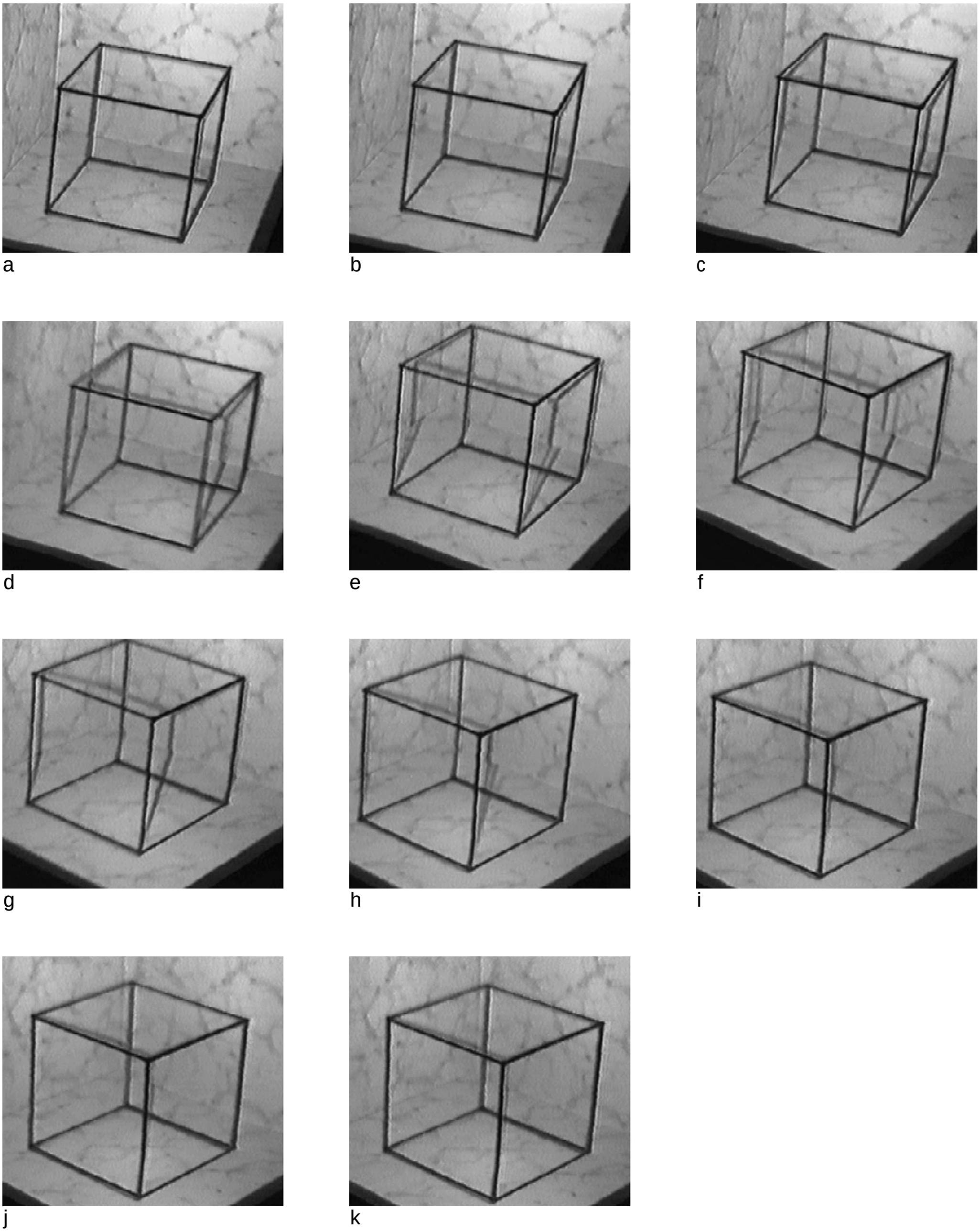


Figure 7.6
Sequence shows lag of virtual image. At the start and end points the observer is stationary.

This shear effect becomes even more disturbing as part of the vertical rib moves to one of the vertical projection screens. Suddenly, points on this part of the rib will have equal distance to the vertical projection screen and consequently will lag by an equal distance. As a result, under observer movement a kink will be visible on the seam of the two projection screens, no matter how well lined up they are. I refer again to Figure 7.6 in which the kink in the vertical ribs can be clearly seen, especially in Figure 7.6d and g. It must be pointed out that this occurs only under observer movement. To a stationary observer the scene appears undistorted.

Another dynamic distortion cue which becomes apparent in Figure 7.6 is that the three projection screens are updated sequentially rather than simultaneously. This is most pronounced in Figures e and f of Figure 7.6 which show the vertical ribs not matching up at the seams between the projection screens.

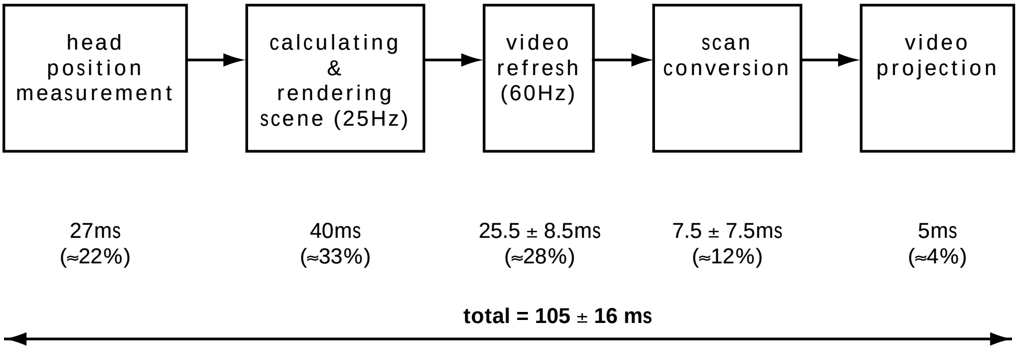
Investigating delay

To find out which part of the system contributed most heavily to the total delay, the component parts of the system were analyzed. Figure 7.7 shows a system diagram. In the analysis the following components were distinguished: Dynasight head tracking, computer image calculation and rendering, screen refresh, scan conversion (the process of converting a computer monitor signal into a video signal) and video projection. The aim of this analysis was not to find the exact delays of each component but to find an estimate of how the total delay was distributed over the components. The reasoning being that only those component(s) which caused the largest delay(s) would need replacing as replacing all components would be very costly.

As the Dynasight optical tracker samples with 37Hz the delay in this component is approximately 27ms.

For simple scenes such as the desk chair and the carpal bones shown in chapter 6 a frame rate of 25Hz could be achieved. Assuming this frame rate the delay caused by computer image calculation and rendering could taken to be approximately 40ms.

The refresh rate of the display is 60Hz and adds 8.5±8.5ms synchronisation delay and 17ms processing delay. The delay added by the display is thus 25.5±8.5ms.



changes:

Dynasight → mechanical head position detection (non-headfree but with a very small delay)	• faster processor • faster graphics board	none	faster and higher resolution scan converter	none

Figure 7.7
The parts of the system which contribute to the total delay

Finally, the measurement of the delay in the scan conversion and projection is detailed in Appendix I. The scan conversion takes up approximately 15ms and the projection 5ms.

Remedies

To minimise the distortion of virtual objects inside Cubby a number of changes were made. After Cubby was calibrated by means of the previously described procedure no further changes were made affecting static distortion cues. Therefore, in this section on remedies attention is focused on those changes which affected flatness cues and those which affected dynamic distortion.

Figure 7.8
The second, more robust Cubby set-up.

Remedies against flatness cues

To alleviate the aforementioned flatness cues a number of changes were made to Cubby. To lessen projection misalignment two measures were taken. First, a new and more robust set-up was built (Figure 7.8). This figure also shows close-up of the display area. A metal frame was added to lend rigidity to the draughting film used for the projection screens. To facilitate lining up the projectors with the screens the set-up was equipped with adjustable projector mounts (Figure 7.9). With this new set-up projection misalignment was indeed less visually disturbing than with its predecessor. Second, it was found that the scan converters change the aspect ratio of the rendered images from 1:1 to 0.97:1 and that consequently the projected images were not quite square. This was compensated for by changing the aspect ratio of the rendered images in software.

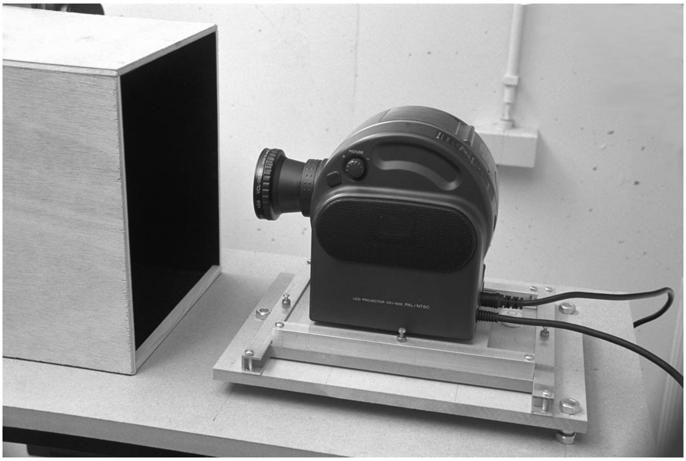


Figure 7.9
Projector mounted on adjustable platform to facilitate adjustment. The box enveloping the projection beam is ligned with black felt to prevent internal reflections.

Finally, to minimise reflections, those parts of the construction exposed to projected light - most notably the boxes which shield the projection beams - were covered with black felt (Figure 7.9).

Remedies against dynamic distortion

When looking at the delay it should be noted that it is not possible to make major changes to some components without major investments. For example, raising the refresh rate of the video board and scan converter combination significantly above 60Hz would be highly expensive. It was therefore decided to concentrate on two parts. The first is trying mechanical head-tracking. The second is faster calculation and rendering to try to ensure that the total delay does not get even worse when using more complex models.

Head tracking

The DynaSight optical head tracker was replaced by a mechanical head tracker (Figure 7.10 and Figure 7.11). The position of the mechanical tracker is read out by means of an analog to digital converter board (PCI1200 by National Instruments). While trackers such as the DynaSight rely on a relatively slow serial link to the computer, an A/D converter board is connected directly to the computer's bus and as such is not hampered by slow data communication.

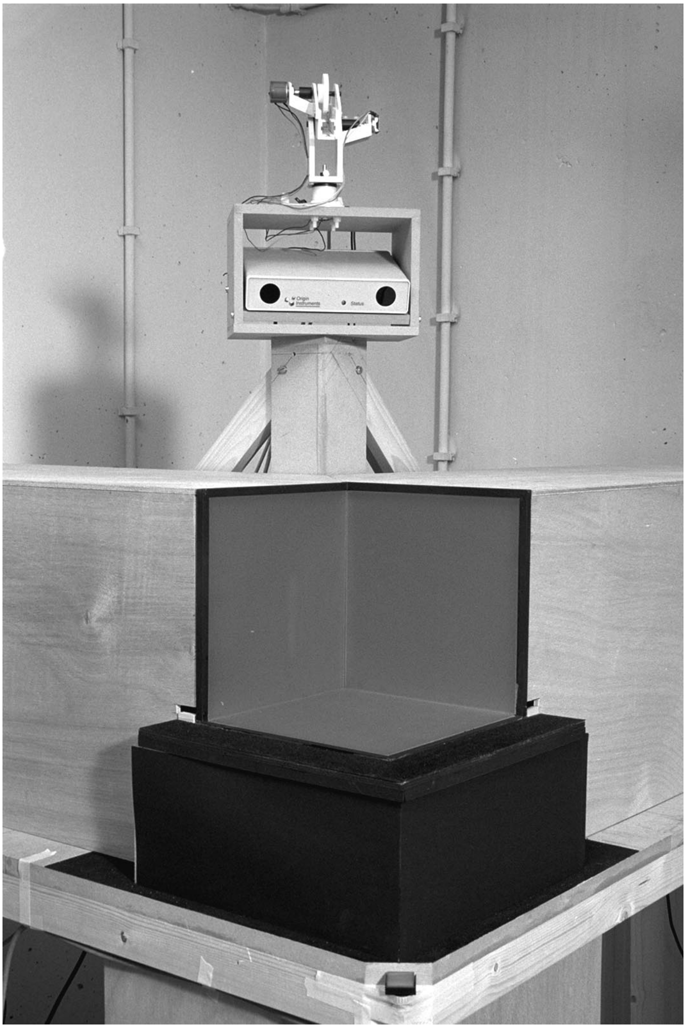


Figure 7.10
The display space with both the Dynasight optical tracker and the mechanical tracker

A mechanical head tracker does not allow for head free tracking and therefore undoes Cubby's advantage of a minimum of head gear (Figure 7.12). However, it does allow for minimisation of delays and a better understanding of the effects of delay in virtual window systems.

Computer set-up

A number of changes were made to the computer set-up. The essence of these changes was that instead of using a display board and a scan converter per projection screen, the set-up used a single display board with a single scan converter to feed all three projectors. The single display

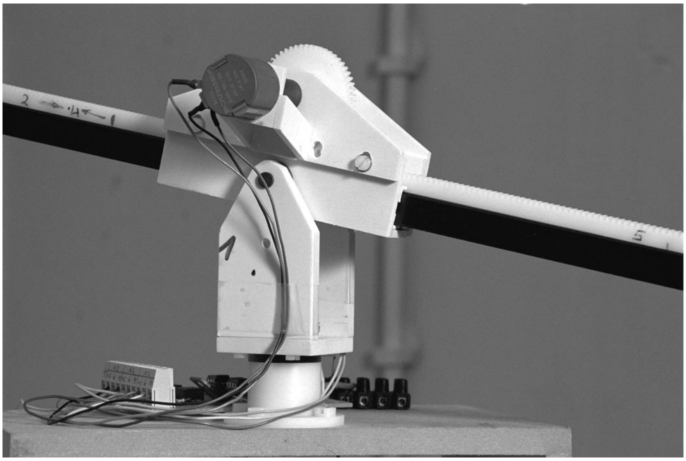


Figure 7.11
The mechanical tracker.

Figure 7.12
Observer with headware for optical tracker (left) and for mechanical and optical tracker (right).

board contained all three images and the projectors were positioned in such a way that only the relevant part of the total display would reach its projection screen (see Appendix II for technical drawings). The advantage of this set-up was that the costs of display boards and scan converters were less by a factor of three. The disadvantage was that

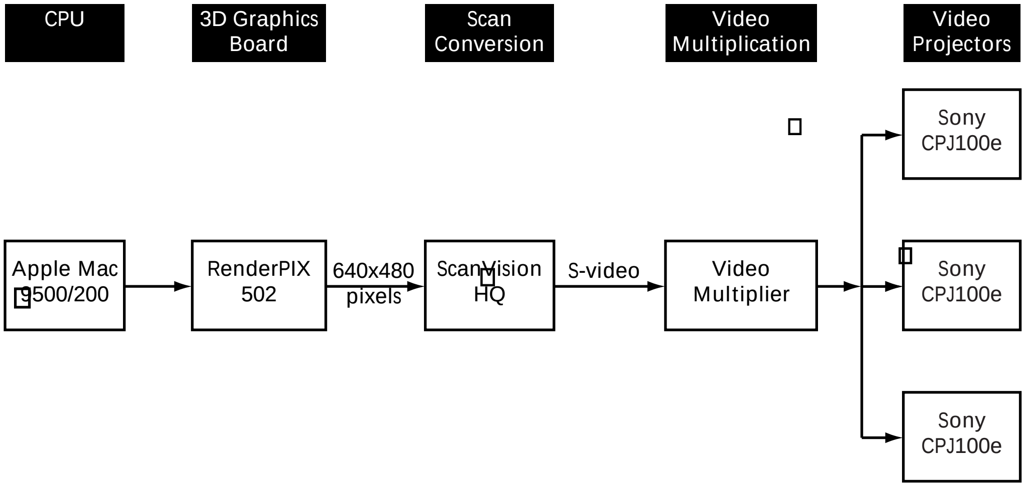


Figure 7.13
The new computer set-up.

the resolution of the projectors was not fully used as an image for one projection screen would take up less than a quarter of a projector’s CCD.

The processor card (PowerPC 604/132MHz) of the PowerMacintosh 9500 was replaced by a faster version (PowerPC 604e/200MHz by Newer Technologies). The three Televisor Zoom scan converters by Displays Technologies were replaced by a ScanVision HQ by Analog Way. Apart from a reduction in delay, this converter was chosen for its ability to convert higher resolutions (1024x768 instead of 640x480) to allow upgrading to higher resolution projectors in the future. The three ATI XClaim graphics boards and the three Apple QD3D accelerators were replaced by a single RenderPIX 502 graphics board by Newer Technologies.A diagram of the new set-up is shown in Figure 7.13.

Summarised, the combined changes to head-tracking and the computer set-up could be expected to reduce the total delay by approximately one quarter. As a result the total delay would end up around 80±16ms.

Conclusions

Delays in the various components of Cubby appeared to be the main cause of the rubbery behaviour seen in Cubby. By improving a number of these components this dynamic distortion was reduced. However, though with the new set-up virtual objects deformed considerably less it was not possible to eliminate the distortion completely. Perceptually this distortion is most notable on verticals. Distortion of virtual objects appears to be more noticeable in Cubby than in immersive VR systems because the observer can see the apparatus surrounding the display space which provides him with a frame of reference.

In the next chapter...

In chapter 8 an experiment is conducted to investigate depth perception within Cubby. Depth perception of a virtual scene using the optical tracker or the mechanical tracker is compared to depth perception of a real scene.

8

Testing Cubby in Depth

Summary

For Cubby to be useful as a simulator it is important that depth perception is accurate and reliable. In an experiment depth perception in Cubby was compared to depth perception in the real world, using both a headfree infrared tracker and a non-headfree mechanical tracker. To establish a performance baseline subjects completed the same experiment using two eyes in the real condition.

Introduction

While the fishtank projection method (McKenna, 1992; Ware, 1993) used in Cubby is supposed to be mathematically and optically accurate, little is known about how people perceive virtual objects which are displayed using this projection method. Consistent and accurate depth information is important in medical visualisation. Unlike for example a game player in virtual reality who can adapt to distorted depth perception, the user of a medical VR system needs to be confident that the skills he practises in the simulator will hold when he is confronted with the physical reality of a patient’s body. If a model of a virtual body is to be used for pre-operative visualisation we need to be certain that the user is not faced with surprises when he views the real body. Therefore, accurate and reliable depth perception is necessary if Cubby is to be used as a medical simulator. For example, when the user is setting out a path for radiation treatment or for a robotic probe, he needs to have an accurate idea of the distance of this path to the critical structures which it is to avoid. Another example is when the user is moving an instrument, say a scalpel, towards an organ, and he needs to be able to judge the distance between the tip of the scalpel and the organ. It can reasonably be assumed that reliable visualisation is a prerequisite for confident manipulation. In the experiment described in this chapter, subjects had to judge the distance between a highly abstracted version of a surgical instrument's tip and that of an organ.

Technical specifications

The Cubby set-up used in the experiment was built around an Apple PowerMacintosh 9500/200MHz and Apple's QuickDraw3D graphics library. As the accelerated 3D graphics board (Newer Technologies RenderPIX502) mentioned in the previous chapter was not yet available the experiment was carried out with one of the existing video boards (ATI XClaim) and two accelerator boards (Apple QuickDraw3D accelerator boards). The monitor signal of the video board was converted to S-Video by means of a scan converter (ScanVision HQ by Analog Way). The S-Video signal was multiplied into three identical signals each of which was sent to a video projector (Sony CPJ-100e). These video projectors had approximately 180,000 pixels, which amounts to 490 x 367 pixels in an aspect ratio of 4:3. The resulting resolution on the back projection screens was approximately 23dpi (approximately one pixel per millimetre). To reduce the distance between the projector lens and the back projection screen the projectors were equipped with wide-angle lenses (Standard lens f=54.9mm, wide angle factor = 0.6).

Two methods of head position detection were used. The first one was headfree by means of a Dynasight optical radar by Origin Instruments. This device tracks the position (three DOFs) of a small (7mm ø) reflective disc which can be attached to a pair of glasses or directly to the forehead. The second one was a non-headfree mechanical tracker. When using this method of head position detection the user wore a helmet which was connected to the mechanical tracker by means of a rod. The rod could rotate around a suspension point with two degrees of freedom as well as slide to and from the suspension point, resulting in three DOFs in total. The movements of the rod and thus of the observer were recorded by means of potentiometers connected to a PCI1200 data acquisition board by National Instruments.

Using this set-up the virtual scenes in the experiment could be displayed with approximately 35 fps.

Group 1 (16 subjects)	
scenes virtual **viewing** monocular **tracker** headfree	**scenes** real **viewing** monocular **tracker** n/a
sample scene 1	sample scene 1
sample scene 2	sample scene 2
scene 1	scene 1
scene 2	scene 2
scene 3	scene 3
scene 4	scene 4
scene 5	scene 5
scene 6	scene 6

Group 2 (16 subjects)	
scenes virtual **viewing** monocular **tracker** non-headfree	**scenes** real **viewing** monocular **tracker** n/a
sample scene 1	sample scene 1
sample scene 2	sample scene 2
scene 1	scene 1
scene 2	scene 2
scene 3	scene 3
scene 4	scene 4
scene 5	scene 5
scene 6	scene 6

Figure 8.1
The three conditions were divided over two groups of subjects. The order of both the sample scenes and the trials was randomised.

Experiment

Design

In a visualisation experiment distance judgements made in Cubby were compared to those in the real world. In addition performance with a head-free tracker was compared to performance with a non-headfree mechanical tracker. There were three conditions in total:

1. virtual scene with headfree Dynasight tracker
2. virtual scene with non-headfree mechanical tracker
3. real scene

Subjects were divided into two groups as shown in Figure 8.1. Group 1 completed the virtual, headfree condition and the real condition, while Group 2 completed the virtual, non-headfree condition and the real condition. The conditions were divided over the two groups for two reasons. First, it was our objective to keep the duration of the experiment within reason. Second, we wished to minimise the risk of subjects getting familiar with the scene.

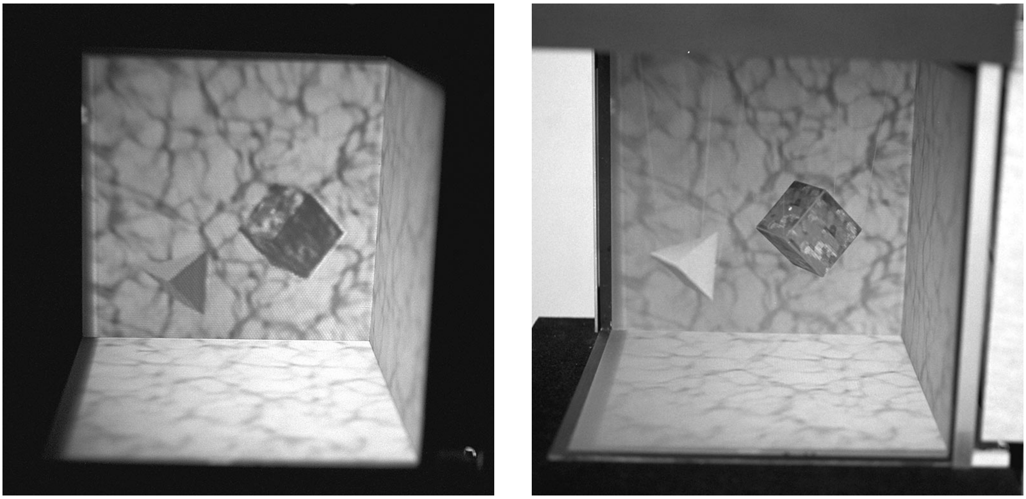


Figure 8.2
The pyramid and cube of the virtual scene (left) and of the real scene (right).

The type of scene which was used consisted of highly abstracted versions of the tip of a surgical instrument and an organ. The instrument tip was represented by a green pyramid while the organ was represented by a textured red cube. On the middle of one of the faces of this red cube a white dot of 3mm diameter could be seen. The objects were displayed against a white marble textured background (Figure 8.2). Subjects were asked to judge the distance between the tip of the pyramid and the white dot on one of the faces of the cube, and to adjust a blank measuring instrument (Figure 8.3) to match that distance. The tip of the pyramid always pointed towards Cubby, just as a user of Cubby would point an instrument towards the virtual body displayed in Cubby. The orientation of the axis between the tip of the pyramid and the dot on the cube was such that a subject could not look at the axis perpendicularly. Therefore, the distance between the objects could not simply be measured in a full orthogonal side view. The orientations of the pyramid and the cube with respect to the connecting axis were also varied. As a result it was impossible for the subjects to derive the orientation of the connecting axis from the orientation of the objects themselves. Because the virtual objects did not cast shadows and because of Cubby being monoscopic subjects had

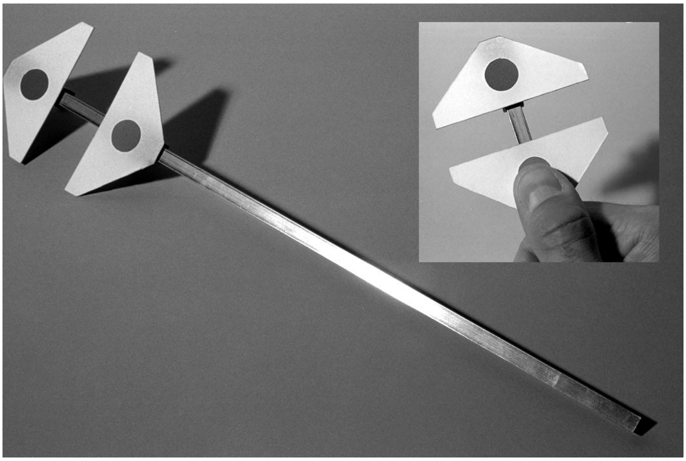


Figure 8.3
The blank measuring instrument consisting of two beaks on a handle. The beak at the top end of the handle was fixed while the other one could move. The inset shows how a subject could adjust the bottom beak with his thumb while holding the handle in the palm of his hand.

to make their distance estimates through position estimates of the virtual objects based on other depth cues such as movement parallax, perspective, texture gradient and occlusion. In each scene the size of at least one of the objects would be changed so that subjects could not make consistent distance estimates on the basis of an assumed size of the cube or pyramid. The edge length of the cubes varied between 30mm and 48mm. The pyramids had a square base and were as high as they were wide. The base edge varied between 30mm and 42mm. The distances between the objects varied from 30mm to 100mm. For each of the virtual scenes displayed in Cubby an isomorphic real scene containing a real pyramid and a real cube was built. These pyramids and cubes were built out of cardboard and weighed down with sand, so that - once stabilised by the experimenter - they would not be caused to sway by the movements of the subject. The texture which was used on the virtual cube was printed at the same resolution as the backprojections (23 dpi) and pasted on the real cube. The real scene was put up in Cubby's display space and the marble background textures displayed in the virtual condition were used for the real condition too. The projection of the textured backgrounds eliminated shadows cast by the real objects, thus mimicking the virtual condition in which no shadows were present. The real objects were suspended from a horizontal plate above Cubby's display space by means of 0.2mm fishing wire (Figure 8.4). Because the suspension points in the

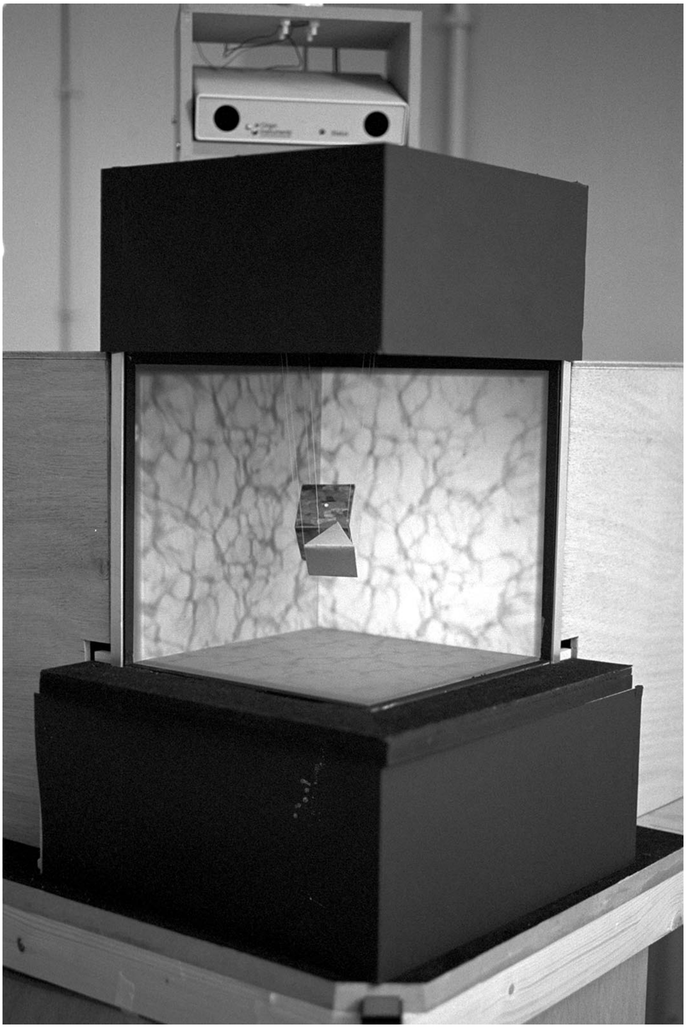


Figure 8.4
The cardboard pyramid and cube suspended on fishing wire.

plate were concealed, it was difficult for subjects to establish the position of the objects other than through the depth cues perspective, occlusion, texture gradient and movement parallax. Subjects did not get any stereoscopic cues as they wore an eye patch just as in the virtual condition

Subjects

There were sixteen subjects in Group 1 and sixteen in Group 2. All subjects were students from Delft University of Technology, except for two, who were members of staff. There were 25 male and seven female subjects. Subjects were paid five guilders for participating (approximately two loafs of bread) and could win 25 guilders if they were the best in their condition.

Procedure

Subjects were seated on a stool. The plate from which the real objects were suspended in the real condition was not present in the virtual condition so that the head detection systems would not be hindered. To prevent that subjects would have an advantage in the virtual condition by looking at the scene from above, they were told to remain seated but encouraged to move as much as possible to benefit from the parallax depth-cue offered by Cubby. To prevent that in the real condition subjects would be able to move further than in the virtual conditions with their limited tracker range, two room dividers were set up which physically limited a subject's movements to the range of the tracker (Figure 8.5). Subjects were given the blank measuring instrument (Figure 8.3) which had two beaks, one marked with a green dot, the other marked with a red dot. The beak with the red dot was fixed at one end of the measuring instrument. Subjects could thus move the beak with the green dot only and were told to adjust the distance between the beaks to the distance between the tip of the pyramid and the face of the cube as accurately and as quickly as possible. It should be noted that the blank measuring instrument was given to the subjects with its beaks closed.

Because installing and removing the suspension system for the real scenes could potentially cause projector misalignment, trials with virtual scenes and trials with real scenes were not mixed but offered as groups. Half of the subjects commenced with the virtual scenes while the other half started with the real scenes. In both conditions subjects were first given two sample scenes to allow them to practise after which six more scenes followed. The order of presentation of the two sample scenes and that of the six trials were randomised.

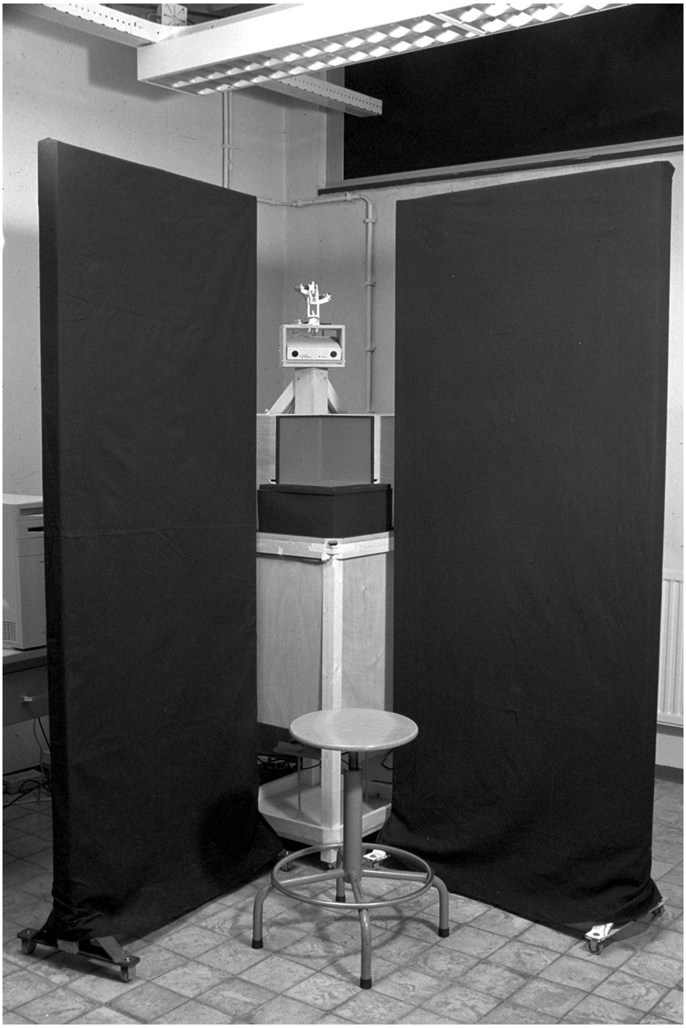


Figure 8.5
The set-up showing the two room dividers which prevented the subjects from being able to move further in the real condition than in the virtual conditions.

During the installation of a real scene subjects had to face away so that they would not get information in advance of starting the trial.

Subjects had 60 seconds to complete a trial. After 50 seconds the computer would give an auditory alert, after which subjects had 10 seconds to make there final adjustment.To indicate that they had finished a trial subjects had to push a button which would cause the elapsed time to be recorded. Subjects then handed over the blank measuring instrument to the experimenter who subsequently recorded the adjusted distance in millimetres.

Lighting and colour of the real and the virtual scenes were matched by eye as well as possible. However, it should be noted that a perfect match was difficult to achieve as the real objects have pigment colour while the virtual objects were generated by projected light. In both conditions it was tried to achieve good contrast between the faces of the objects. To ensure that the real objects would not only be lit from the back by the projection screens an indirect incandescent spotlight was used to illuminate the faces of the objects which were turned towards the subject. This light was on in both the real and the virtual condition.

Hypotheses

The following comparisons were to be made:

a. Within Group 1: virtual scenes with headfree tracker vs. real scenes
b. Within Group 2: virtual scenes with non-headfree tracker vs. real scenes
c. Between Group 1 and Group 2: virtual scenes with headfree tracker vs. virtual scenes with non-headfree tracker

The results of each trial were expressed as follows:

$$\text{error} = \frac{\|\text{adjusted distance} - \text{true distance}\|}{\text{true distance}}$$

The adjusted distance is the distance which was adjusted by the subjects on the blank measuring instrument. In case of the real scenes the true distance is the physical distance between the objects, in case of the virtual scene it is the distance the objects were put apart in the modelling program. The value for the error which results from the above formula is dimensionless. This was done because the distance between cube and pyramid differed per scene. If the error would not have been made dimensionless, a small distance would have led to a small error, while a large distance would have led to a large one.

This leads to the following hypotheses (Table 8.1).

Table 8.1 Hypotheses

H_{a0}	μ_{Ehf}	=	μ_{Er}	H_{a1}	μ_{Ehf}	≠	μ_{Er}
H_{b0}	μ_{Enhf}	=	μ_{Er}	H_{b1}	μ_{Enhf}	≠	μ_{Er}
H_{c0}	μ_{Ehf}	=	μ_{Enhf}	H_{c1}	μ_{Ehf}	≠	μ_{Enhf}
H_{d0}	σ^2_{Ehf}	=	σ^2_{Er}	H_{d1}	σ^2_{Ehf}	≠	σ^2_{Er}
H_{e0}	σ^2_{Enhf}	=	σ^2_{Er}	H_{e1}	σ^2_{Enhf}	≠	σ^2_{Er}
H_{f0}	σ^2_{Ehf}	=	σ^2_{Enhf}	H_{f1}	σ^2_{Enhf}	≠	σ^2_{Er}
Note: E = error, r = real, hf = headfree, nhf = non-headfree							

Results

Distance

From the 16 subjects in Group 1 one was excluded from the analysis as due to a technical malfunction. From the 16 subjects in Group 2 one was excluded from the analysis as he hardly moved at all and thus did not make use of the essential movement parallax depth cue. The results are grouped together in Table 8.2. They are shown graphically in Figure 8.6.

a. Within Group 1: virtual scenes with headfree tracker vs real scenes — Table 8.2 includes the results of Group 1 for the virtual scenes with the headfree tracker and for the real scenes. With a t-test no significant difference in mean could be found (Table 8.3). However, an F-test shows that the variance in the real condition is significant lower than in the virtual headfree condition (p<0.01) (Table 8.3). These results can also be seen in Figure 8.6.

b. Within Group 2: virtual scenes with non-headfree tracker vs real scenes. — Table 8.2 includes the results of Group 2 for the virtual scenes with the non-headfree tracker vs. the real scenes. With a t-test no significant difference could be found (Table 8.4). An F-test shows that differences in variance are not significant either (Table 8.4). These results can also be seen in Figure 8.6.

Figure 8.6
Means and 95% confidence intervals for error.

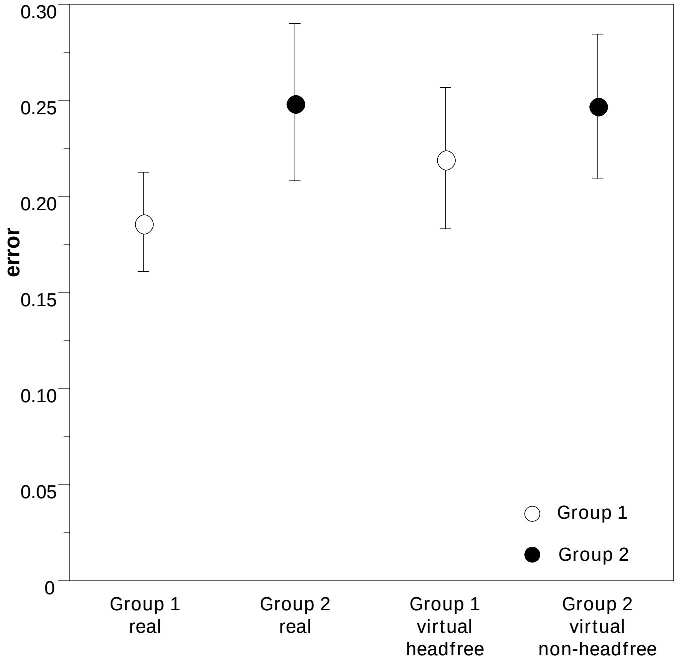


Table 8.2 Results Group 1 (real and virtual headfree) and Group 2 (real and virtual non-headfree)

	Mean	Variance
Group 1 real	0.186	0.015
Group 1 virtual headfree	0.220	0.030
Group 2 real	0.249	0.038
Group 2 virtual non-headfree	0.247	0.032

Group 1: SS 1-15
Group 2: SS 15-30

c. Between Group 1 and Group 2: virtual scenes with headfree tracker vs virtual scenes with non-headfree tracker — Table 8.2 includes the results of Group 1 for the virtual scenes with the headfree tracker and of Group 2 for the virtual scenes with the non-headfree tracker. With a t-test no significant difference in means could be found

Table 8.3 t-tests and F-tests for Group 1 real vs Group 1 virtual headfree

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 1 real vs Group 1 virtual headfree	-0.033	0.1408	1.996	0.0015

Table 8.4 t-tests and F-tests for Group 2 real vs Group 2 virtual non-headfree

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 2 real vs Group 2 virtual non-headfree	0.002	0.9437	1.210	0.3697

(Table 8.5). An F-test shows that differences in variance are not significant either (Table 8.5). These results can also be seen in Figure 8.6.

Table 8.5 t-tests and F-tests for Group 1 virtual headfree vs Group 2 virtual non-headfree

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 1 virtual headfree vs Group 2 virtual non-headfree	-0.027	0.2998	1.057	0.7965

Comparing the real conditions of Group 1 and Group 2 — In the real condition Group 1 should perform approximately the same as Group 2. However, a t-test shows that in Group 1 the error mean is significantly lower than in Group 2 ($p<0.05$) (Table 8.6). In addition, an F-test shows that variance in Group 1 is also significantly lower than in Group 2 ($p<0.0001$) (Table 8.6).

An order effect? — To investigate the reason for performance differences in both real conditions of Group 1 and Group 2 possible order effects will now be considered. Table 8.7 shows the four groups that emerge when the order of conditions is taken into account.

Table 8.6 t-tests and F-tests for Group 1 real vs Group 2 real

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 1 real vs Group 2 real	-0.063	0.0114	2.538	<0.0001

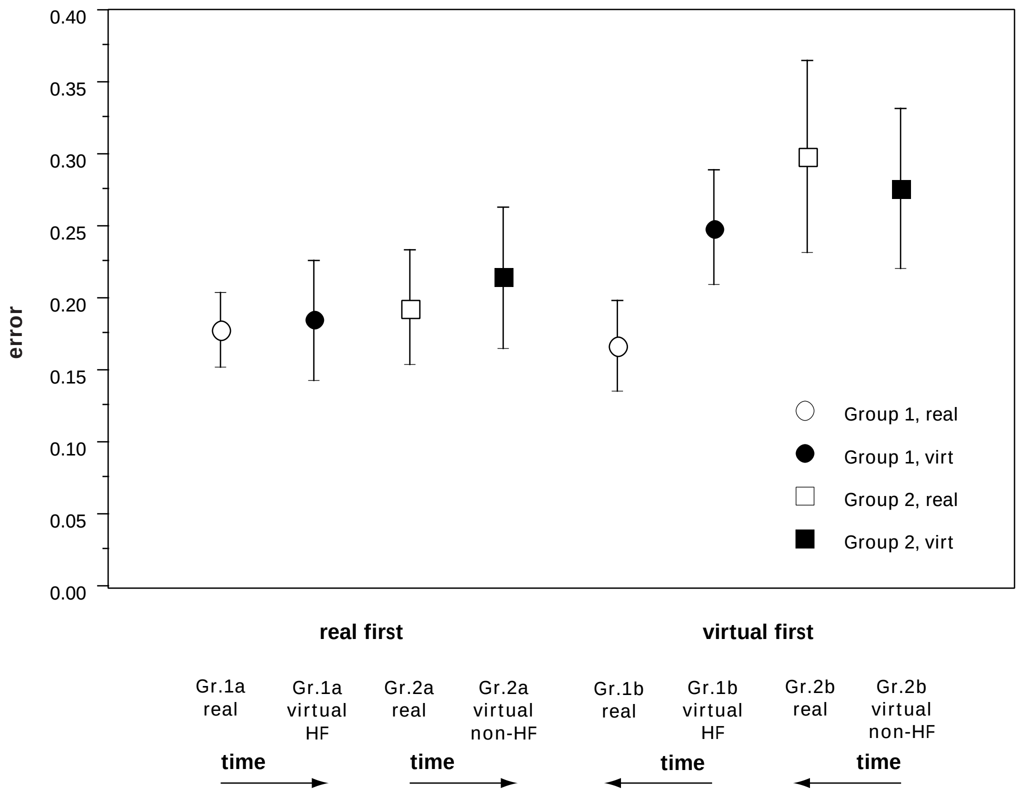
Table 8.7 Possible orders within Group 1 and Group 2

	first condition	**second condition**
Group 1a	real	virtual headfree
Group 1b	virtual headfree	real
Group 2a	real	virtual non-headfree
Group 2b	virtual non-headfree	real

Four groups make eight possible conditions when order is taken into account. There are four groups, each of which completes two conditions. The results for these eight possibilities are shown in Table 8.8. These results are shown graphically in Figure 8.7.

It can be seen that the performance in Group 2 is affected by the order of the conditions while performance in Group 1 is not. Subjects in Group 2 made distance estimates with significant higher accuracy (t-test, p<0.05) in the real condition when this condition was offered as the first condition than when it was offered as the second condition (Table 8.9). In other words, in spite of the fact that subjects had the chance to practise with virtual scenes and the non-headfree tracker, in the real condition which followed accuracy was negatively affected. This could be the result of subjects feeling limited in their movements in the real condition because they were limited in the preceding virtual condition with the mechanical tracker.

Subjects in Group 2 who were offered the virtual condition as the second condition, i.e., after they had completed the real condition, did not perform significantly different from those who were offered the virtual condition as the first condition.



Group 1a: SS 1-8
Group 1b: SS 9-15
Group 2a: SS 16-23
Group 2b: SS 24-30

Figure 8.7
Means and 95% confidence intervals for error.

Variance for the real condition of Group 2 was also significantly affected (F-test, p<0.001) by the order in which the conditions were presented. When the real scenes were presented as the second condition, i.e., after subjects had completed the virtual condition with the non-headfree tracker, variance on distance estimates was significantly higher than when the real scenes were presented as the first condition. Again, this could be the result of subjects

Table 8.8 Results taking condition order into account

	first condition	Mean	Variance	second condition	Mean	Variance
Group 1a	real	0.189	0.013	virtual headfree	0.203	0.034
Group 1b	virtual headfree	0.239	0.026	real	0.899	0.042
Group 2a	real	0.193	0.017	virtual non-headfree	0.214	0.025
Group 2b	virtual non-headfree	0.276	0.036	real	0.298	0.053
Group 1a: SS 1-8 Group 1b: SS 9-15 Group 2a: SS 16-23 Group 2b: SS 24-30						

being limited in their movements in the real condition because they were limited in the preceding virtual condition with the non-headfree tracker.

No differences in variance could be found between the non-headfree tracker condition presented as the first and presented as the second condition.

Table 8.9 t-tests and F-tests for condition order per condition

	t-test		F-test	
	Mean diff.	P-Value	F-Value	P-Value
Group 1a vs Group 1b (real)	0.005	0.8468	1.333	0.3480
Group 1a vs Group 1b (virtual headfree)	-0.036	0.3242	1.286	0.4106
Group 2a vs Group 2b (real)	-0.105	0.0103	3.195	0.0002
Group 2a vs Group 2b (virtual non-headfree)	-0.062	0.0997	1.443	0.2280

Because the order in which the conditions were presented had a significant influence in case of the non-headfree tracker, new comparisons were made between the virtual and real conditions split by first and second condition. The second condition for Group 2 was omitted from these comparisons because the real condition suffered from a significant order effect as was shown previously. The results of the t-tests are shown in Table 8.10. For the real condition no significant difference could be found between Group 1a and Group 2a if it was offered as the first condition. Differences between the real and the virtual headfree condition within Group 1 are neither significant if these conditions are offered first nor if they are offered second. However, in Group 2 the mean error was significantly higher in the virtual non-headfree condition than in the real condition if these conditions were offered first ($p<0.05$). When comparing the virtual non-headfree condition and the virtual headfree condition no significant difference was found for the order in which these conditions are offered first.

Table 8.10 t-tests and F-tests between subgroups

	Mean Diff.	P-Value	F-Value	P-Value
Group 1a real vs. Group 2a real (1st condition)	-0.004	0.8720	1.252	0.4631
Group 1a real vs. Group 1b virtual headfree (1st condition)	-0.050	0.0923	1.969	0.0279
Group 1a virt. headfree vs. Group 1b real (2nd condition)	0.019	0.5846	1.902	0.0370
Group 2a real vs. Group 2b virt. non-headfree (1st condition)	-0.083	0.0192	2.193	0.0108
Group 1b virt. headfree vs Group 2b virt. non-headfree (1st condition)	-0.037	0.3270	1.393	0.2777

The results of the F-tests are also shown in Table 8.10. Again, the second condition for Group 2 was omitted from these comparisons because the real condition suffered from a significant order effect as was shown previously. For the real condition no significant difference could be found between Group 1a and Group 2a if it was offered as the first condition. Differences between the real and the virtual headfree condition within Group 1 are significant if these conditions are offered first and also if they are offered second (p<0.05). Differences between the real and the virtual headfree condition within Group 2 are significant if these conditions are offered first (p<0.05). When comparing the virtual non-headfree condition and the virtual headfree condition variances no significant difference could be found for when these conditions are offered first.

Time

From the duration of the trials no clear pattern emerged. Group 1b took significantly more time to complete the trials in both the virtual headfree and the real condition compared to any of the other subgroups. For this I do not have a satisfactory explanation.

Discussion

From the results it can be seen that it is not safe to make the comparisons I set out to make, since the error in the real condition of Group 1 and Group 2 differed significantly in both means and variances. This was surprising as these real conditions were identical for both groups. However, comparisons within all conditions between those subjects who completed a condition as their first condition and those who completed it as their second condition, showed that there were significant differences in error in the real condition of Group 2 both in terms of means and variances. The subjects who completed the real condition as their second condition performed significantly worse than those who completed it as their first condition. This could be the result of subjects being limited in their movements in the preceding non-headfree condition and carrying on this behaviour into their second, real condition. As none of the other conditions exhibited this, the second condition offered to Group 2 was omitted from further analysis. After this the real conditions of Group 1 and Group 2 no long differed significantly.

For the headfree condition no difference in mean could be found when compared with the real condition, though these two conditions did differ significantly in variance.

However, the non-headfree condition did differ significantly from the real condition, both in means and variance. Again, this could be the result of subjects being more limited in their movements by the non-headfree tracker than by the headfree tracker.

No significant differences could be found when directly comparing the headfree and non-headfree trackers. The reason for this can be seen when looking at Figure 8.7 and only considering the data of the conditions which were offered first. Error increases when moving from the real, through the headfree, to the non-headfree condition. As a result, while the differences in means between the real and headfree conditions and between the headfree and the non-headfree condition are not significant, the differences between the real and the non-headfree condition are.

The poor performance with the non-headfree tracker could well be the result of its movement hampering aspect, since the dominant depth cue in Cubby, being movement parallax, is dependent on free observer movement.

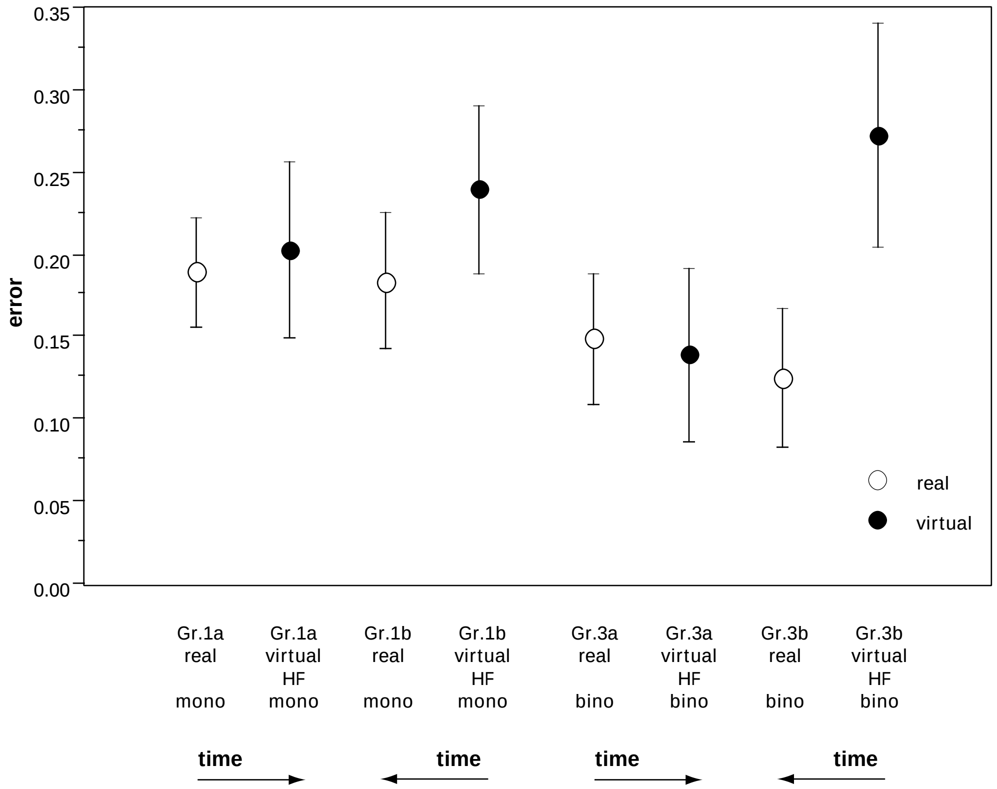
As for the differences in variance it must be remembered that, while the real and virtual conditions were made to resemble each other as closely as possible, there are still many differences between them. We name delay, limited spatial and temporal resolution and flatness cues.

Having compared the real and virtual conditions with both headfree and non-headfree trackers viewed monocularly, the question remained how well subjects would perform when they view the scene binocularly. For the real scene binocular viewing would give subjects stereoscopic depth cues, while for the virtual scene it could result in the conflict between apparatus and surroundings which are viewed stereoscopically and the scene itself which is monoscopic.

Binocular conditions

A third group of six naive subjects had to complete these binocularly viewed real condition and binocularly viewed virtual condition. For the virtual condition the headfree tracker was used. The conditions for Group 3 were therefore identical to Group 1 except for the monocular/binocular difference. The results for Group 3 together with those of Group 1 are shown in Table 8.11, again split by condition order. These results are shown graphically in Figure 8.8 together with those of Group 1.

I would like to make three comparisons. First, how do the *real monocular* and the *real binocular* condition compare? Second, how do the *virtual headfree monocular* and



Group 1a: SS 1-8
Group 1b: SS 9-15
Group 3a: SS 31-33
Group 2b: SS 34-36

Figure 8.8
Means and 95% confidence intervals for error.

the *real binocular* condition compare? This is interesting for comparison of a simulated operation in Cubby to a real one which is performed binocularly. And finally, how do the *virtual headfree monocular* and the *virtual headfree binocular* conditions compare? This last question is interesting from a practical point of view: can subjects make accurate depth estimates with Cubby's monoscopic images while looking with two eyes, or is performance higher when they wear an eye patch?

Table 8.11 Results Group 1 and 3 taking condition order into account

	first condition			second condition		
		Mean	**Variance**		**Mean**	**Variance**
Group 1a	real mono	0.189	0.013	virtual headfree mono	0.203	0.034
Group 1b	virtual headfree mono	0.239	0.026	real mono	0.184	0.018
Group 2a	real bino	0.149	0.006	virtual headfree bino	0.139	0.011
Group 2b	virtual headfree bino	0.272	0.019	real bino	0.125	0.007

Group 1a: SS 1-8
Group 2b: SS 9-15
Group 3a: SS 31-33
Group 3b: SS 34-36

The results of these three comparisons are discussed next.

Real monocular vs real binocular — No significant differences in means could be found, neither when performed as the first (Group 1a), nor when performed as the second condition (Group 1b). When comparing the variances, there are no significant differences when the real condition is completed first, though an F-test shows that with binocular viewing variance is significantly lower than with monocular viewing ($p<0.05$) if the real condition is completed as the second condition. The results are shown in Table 8.12.

Table 8.12 t-tests and F-tests for real monocular vs real binocular

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 1a vs 3a (real, first condition)	0.040	0.1796	2.100	0.0721
Group 1b vs 3b (real, second condition)	0.059	0.0895	2.433	0.0354

Virtual headfree monocular vs real binocular — Significant differences in means were found, for both the first and the second condition ($p<0.05$). Moreover, when comparing the variances, there are significant differences for both the first ($p<0.001$) and the second condition ($p<0.005$). The results are shown in Table 8.13.

Table 8.13 t-tests and F-tests for virtual headfree monocular vs real binocular

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 1b (virtual mono) vs 3a (real, bino) first condition	0.110	0.0115	5.003	0.0003
Group 1a (virtual mono) vs 3b (real, bino) second condition	0.083	0.0478	4.039	0.0011

Virtual headfree monocular and virtual binocular — No significant differences in means between these conditions could be found, neither when performed as the first (Group 1a), nor when performed as the second condition (Group 1b). When comparing the variances, there are no significant differences when the conditions are completed as the first conditions. However, an F-test shows that with binocular viewing variance is significantly lower than with monocular viewing ($p<0.01$) if the virtual condition is completed as the second condition. This is surprising as performance in the virtual binocular condition was expected to be worse than in the virtual monocular condition, due to conflicts between the stereoscopically perceived surroundings and apparatus on the one hand, and the monosopic virtual scene on the other hand. I have no satisfactory explanation for this result. It may be that the stereoscopically viewed concave space actually pulls the virtual scenes into 3D, and that this effect is stronger than the monoscopic-stereoscopic depth cue conflict acting as a flatness cue. However, it could also be that adjustment of the blank measuring instrument is more reliable when viewed binocularly. The results are shown in Table 8.14.

Table 8.14 t-tests and F-tests for virtual monocular vs virtual binocular

	t-test		F-test	
	Mean diff.	**P-Value**	**F-Value**	**P-Value**
Group 1a vs 3a (virtual, second condition)	0.064	0.1697	2.799	0.0087
Group 1b vs 3b (virtual, first condition)	-0.033	0.4497	1.372	0.4442

Conclusion

In this experiment we compared distance estimates between virtual objects in Cubby, using both a headfree and a non-headfree tracker, and estimates between real objects. A prerequisite for good performance in a system based on movement parallax is an observer who is and who feels unhampered in his movements. Performance of subjects in the real condition appeared to be significantly less when preceded by the virtual condition with the mechanical non-headfree tracker, than when the real condition was offered as the first condition. Unlike the virtual condition with the headfree tracker, which only a showed a significant difference in variance compared to the real condition, the virtual condition with the non-headfree tracker, also showed a difference in means.

The experiment also showed that if the real condition is performed monocularly only the variance is lower than in the virtual headfree condition, while if the real condition is performed binocularly both mean and variance are significantly lower. Thus performance in Cubby cannot match performance with a binocularly viewed real scene. However, performance in Cubby compares well to a monocularly viewed real scene. The higher variance may be the result of a number of factors such as delay, limited spatial and temporal resolution and flatness cues.

Finally, there was less variance in the virtual headfree binocular condition than in the virtual headfree monocular condition for which I have no satisfactory explanation.

It should be noted that the sign of the error was not considered in the analysis. In other words, the results — errors of approximately 20% for the monocularly viewed scenes — do not tell whether subjects systematically underestimated or underestimated distances in Cubby.

In the next chapter...

instrument based manipulation is added to Cubby in such a way that the display and manipulation spaces are unified. In an experiment a unified Cubby, a non-unified Cubby, a unified fishtank display and a non-unified fishtank display are compared.

9

Manipulation in Cubby

> *"In the primate and the human, the five-pronged shapes that specify the hands are especially meaningful. Their deforming contours and the underlying invariants make possible what psychologists have called, very inadequately, eye-hand coordination. More exactly, they are the basis of the visual control of manipulation. And when an object grasped by the hand is used as a tool, it becomes a sort of extension of the hand, almost a part of the body." (Gibson, 1986)*

Summary

To allow manipulation in Cubby a six DOF, hybrid instrument is added with a physical barrel and a virtual tip. By means of this instrument the objects in the virtual scene can be manipulated directly because they appear in front of the projection screens. As the virtual objects and the instrument can share the same space the display and manipulation spaces can be unified. Non-immersive, unified systems based on head-tracking are commonly associated with problems such as occlusion anomalies, clipping, and tracker jittering caused by a CRT display. With Cubby, these problems are minimised or eliminated through the use of the aforementioned hybrid instrument, projection technology, and multiple, orthogonal screens. Manipulation in Cubby is tested by means of an experiment in which subjects had to complete a three dimensional, virtual puzzle of an icosahedron. Completing the puzzle requires picking up three pieces and translating and orienting them into the correct places by means of the hybrid instrument. The experiment includes four conditions. In a condition either Cubby or a single screen fishtank display is used, and the display and manipulation spaces are either separated or unified (Figure 9.7). In a system in which the display and manipulation spaces are separated there is an offset between the manipulating instrument and the virtual object which is being manipulated. The results show that subjects were able to manipulate virtual objects with higher accuracy in the Cubby unified condition than in the single screen unified condition. Subjects also performed more accurately in the Cubby non-unified condition than in the single screen non-unified condition. No differences in performance could be found between the unified and non-unified conditions for either Cubby or the single screen fishtank display. However, when subjects are asked to rank the conditions according to preference, a pattern emerges which shows not only that Cubby is preferred

over the single screen fishtank set-up, but also that within a set-up the unified condition is preferred over the non-unified condition.

Introduction

In many non-immersive 3D systems the display space and the manipulation space are separated. The virtual objects appear behind the monitor screen which prevents the user from getting at the objects directly. As a consequence the user has to manipulate the virtual objects by means of an input device which occupies a different space. Moreover, if the input device has less than six degrees of freedom there is a problem with mapping the degrees of freedom of the input device onto the six degrees of freedom of the virtual object to be manipulated. When such systems are used as surgical simulators the actions needed to manipulate the virtual body on the screen do not match those needed during the actual operation. Therefore the user is practising how to instruct the computer rather than practising the skills which he needs in the operating theatre.

Even if the input device does have six degrees of freedom the proprioceptive information may not match the visual information. An example of such a six DOF input device is the SpaceBall by the SpaceTec IMC Corporation. As the SpaceBall is an isometric input device the user manipulates objects with only miniscule finger movements. Only if the six DOFs input device is of the isotonic, position variety can the skills to handle the input device resemble those needed to do the actual operation.

Yet even with isotonic position input devices with six DOFs, the hand-eye coordination in the simulator will be different from that in the operating room. This is because the space in which the hand moves is separated from the one in which the virtual objects reside. There is evidence in the literature that *rotation* of the manipulation space with regard to the display space has a negative effect on performance. Tendick et al. (1993) show for an endoscopy task that when the angle between the display and manipulation space exceeds 45°, performance decreases considerably.

A side effect of the separation of display and manipulation space is that if the user focuses his attention on the screen he will miss out on the visual feedback offered by his manipulating hand. To offer feedback most 3D systems show some kind of cursor on screen. This may range from a simple arrow similar to the cursor used in graphical user interfaces, to a 3D cursor which shows both position and orientation of the input device. However, even when compared to what would be considered an advanced 3D cursor, the visual feedback which the user obtains from his manipulating hand when handling objects in everyday life, is much richer.

Unification of the display and manipulation spaces

Unified or non-unified?

I will call systems which allow the unification of the display and manipulation spaces ‘unified systems’. First, the requirements for a non-immersive 3D system to qualify as a unified system are specified. Second, a number of disadvantages commonly associated with unified systems are discussed. Third, some experimental non-immersive VR systems which unify the display and manipulation spaces are described. Fourth, those features of Cubby are highlighted which allow it to minimise or overcome the disadvantages commonly associated with unified systems.

Ingredients for a non-immersive unified system

Not every non-immersive 3D system can be made into a unified system. It must satisfy two requirements. Firstly, it must allow some kind of look-around facility whereby the virtual objects appear to be rigidly connected to their physical surroundings. Note that this excludes systems which are based on stereo only as these cannot maintain the illusion that a hand or instrument touches a particular spot on a virtual object when the user moves his head. Systems based on movement parallax can maintain this illusion. Secondly, it must be possible for physical entities, such as a hand or instrument, to share the same space as the virtual scene.

Cubby satisfies these requirements. It creates the illusion that virtual objects stand within its display space. These virtual objects are directly accessible as they are displayed in front of the projection screens.

Disadvantages of unification

There are a number of disadvantages associated with integrating the display and manipulation spaces in non-immersive 3D systems. The first one is the occlusion conflict between the physically present hand and the virtual objects. While the hand occludes the virtual objects, the virtual object cannot occlude the hand.

The second disadvantage of unification concerns user fatigue. User fatigue in unified systems is the result of an absolute rather than a relative coupling between the input device and the cursor. A relative input device reports changes in position, whereas an absolute input device reports its position relative to a fixed coordinate system[1]. When the display and manipulation space are to be unified the hand and the cursor cannot be relatively coupled. From the point of view of user fatigue the advantage of a relative coupling is that he can make use of a clutch. Through declutching the user can position his hand anywhere which feels comfortable. Afterwards he can re-engage the clutch and the coupling between the hand and the virtual object will be re-established. In contrast, this cannot be done with absolute coupling and therefore it is possible that the user needs to keep his hand outstretched in an uncomfortable position. Note that this may be more the result of neglect of ergonomic aspects of desktop VR systems than an inherent shortcoming of the absolute coupling in unified systems. While the standard computer configuration of putting a monitor on top of

---

[1] To understand the terms relative and absolute coupling in relation to a 3D system, an analogy can be drawn with input devices in graphical user interfaces. A mouse is a relative input device, while a drawing tablet is an absolute input device. A mouse implements a relative coupling between hand movement and cursor movement: it reports changes in position. The user can interrupt the coupling by lifting the mouse from the desk. While the mouse is in the air, the tracking is lost and it cannot report changes in position to the system until it is put down again. A drawing tablet on the other hand, establishes an absolute coupling between hand movement and cursor movement: it reports its position relative to a fixed coordinate system. Lifting the pen from the drawing tablet does not interrupt the coupling. When the pen is picked up and moved to another location, the cursor will move to another location too, because the tablet reports the new location to the system.

the computer may be acceptable when used together with a two DOF input device which rests on the work surface, such a configuration is ergonomically poor when used in conjunction with a six DOF isotonic position input device which needs to be held in the air.

The third disadvantage applies to single screen, movement parallax based 3D systems. It concerns clipping of virtual objects by the monitor's edges under head movements. This kind of clipping should be avoided as it makes the 3D impression collapse. To make virtual objects accessible to a physical entity such as a hand or a stylus the virtual scene needs to be positioned in front of the monitor screen. However, if the virtual scene is positioned completely in front of the screen, the viewing volume within which the user can move his head without virtual objects being clipped by the screen's edges, is much smaller than when the virtual scene is positioned half in front, half behind the screen. Assuming a fixed screen size, this means that either the user is more limited in his movements or that the virtual scene needs to be smaller.

The fourth disadvantage is that the electro-magnetic trackers, which are used for tracking hand and head movements in most non-immersive 3D systems, do not work well in the proximity of a CRT display. The magnetic field of the CRT display causes distortion of the field of the tracking system resulting in inaccuracy and jittering. In a system with relative coupling inaccuracy is far less of a problem than in a unified system in which the coupling is absolute. Jittering is of course highly undesirable for a system which aspires to be a surgical simulator for which precision is essential.

Existing unified, desktop-sized, non-immersive systems

In a system by Schmandt (1983), one of the pioneers of unified systems, the user sees the stereoscopic image of a CRT under 45° reflected in a half-silvered mirror parallel to the floor (Figure 9.1). The user can reach underneath the mirror into the 3D scene and paint in 3D or input vertices with a six DOF wand. Schmandt's system does not make use of movement parallax. The assumption is that the user will not move his head and that the user's point of view, which is measured at the beginning of a session, will stay valid during a session.

Kameyama et al. (1993a) use two concave mirrors to translate the image of an autostereoscopic volume scanning LCD panel into another free space, so that virtual

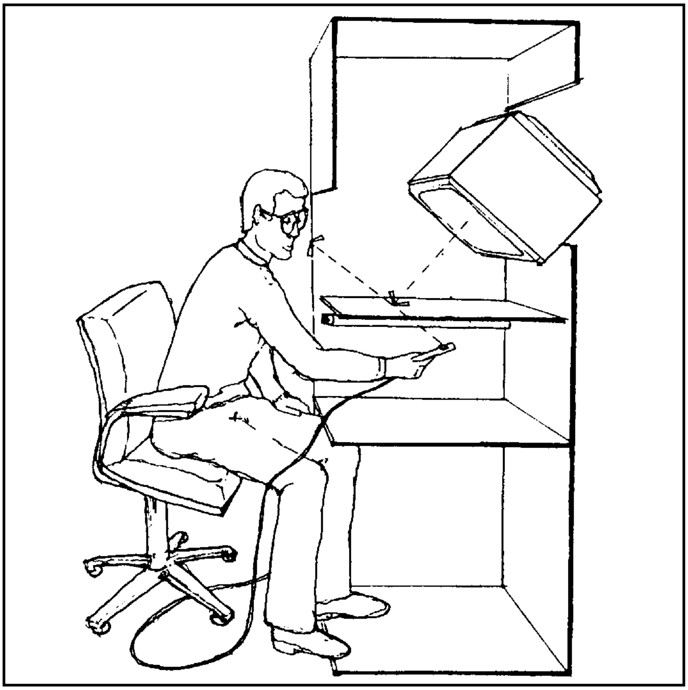


Figure 9.1
Schmandt's system (© 1983, Association for Computing Machinery).

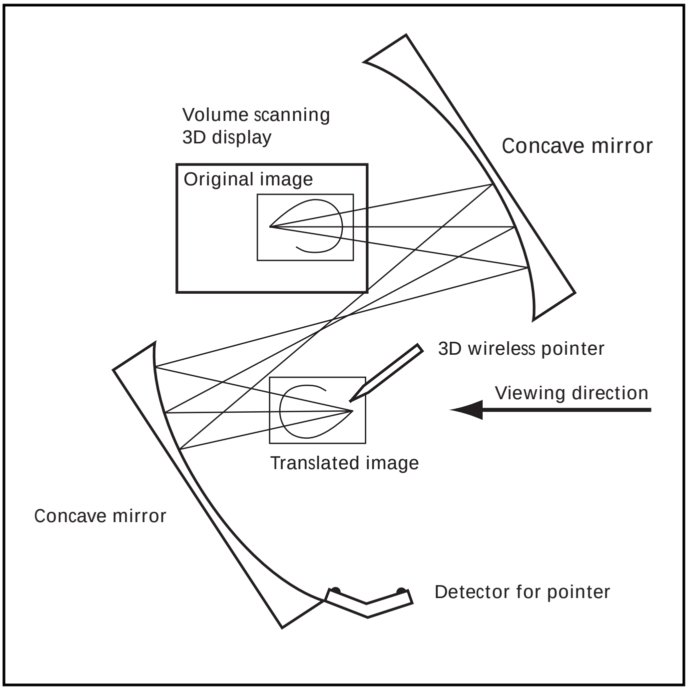


Figure 9.2
Diagram of Kameyama's volume scanning display with optical relay system.

objects can be manipulated with a wireless 3D mouse (Figure 9.2). Because of the volume scanning this system offers both stereoscopy and movement parallax. Kameyama et al. (1993b) also describe a similar set-up but with the optical relay system replaced by a half-silvered mirror.

Ishii et al. (1994) have developed a unified system with a stereoscopic, movement parallax, single screen display and a mechanical six DOF pointing device with force feedback. They use a virtual pointer which is rendered as an extension of the physical pointing device. In a positioning task they compare a stylus with a virtual pointer, a stylus without a virtual pointer and a joystick controlled virtual pointer. In the first two cases there is force feedback while in the latter there is auditive feedback only. The results show that the stylus with the virtual pointer yields the highest performance, followed by the joystick and finally by the stylus without the virtual pointer. Ishii et al. (1994) say that the poor performance of the stylus without the virtual tip is the result of there always being some deviation between the stylus tip in the real and in the virtual space. Consequently the user must take that offset into account when finding the target. This is difficult when there is no virtual tip which indicates the position of the tip of the stylus according to the system.

Finally, there are two unified systems which both use head-tracked stereoscopy on a screen the size of a drafting table. These are the Responsive Workbench (Krüger and Fröhlich, 1994) and the ImmersaDesk (Czernuszenko et al., 1997). The main difference between the two systems is

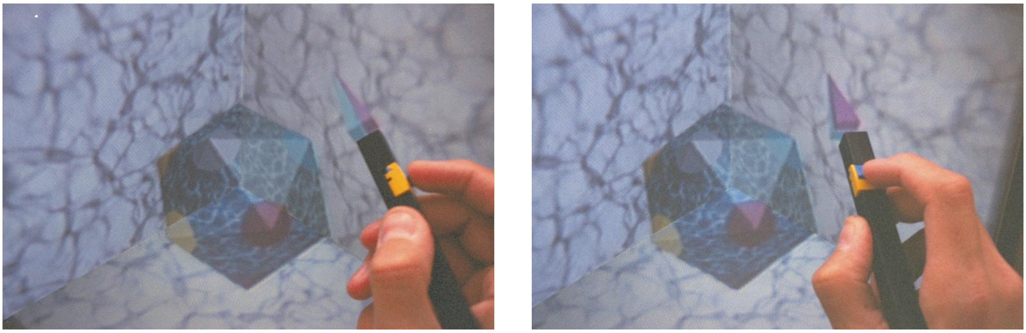


Figure 9.3
The virtual tip is rendered as an extension of the physical barrel.

that while the Workbench has a horizontal screen, the screen of the ImmersaDesk is placed under 45°. Accordingly, the Workbench is better suited to tasks which in the real world are performed on a table, while with the ImmersaDesk it is easier to view both the front and the top of a virtual model without clipping occurring.

Cubby as a unified system

Having listed the disadvantages commonly associated with unified systems and a number of experimental unified set-ups, I now turn to how in Cubby these disadvantages have been alleviated.

To reduce the occlusion problem, Cubby uses a hybrid instrument with a physical barrel and a virtual tip. The tip is rendered as an extension of the physical barrel(Figure 9.3). The virtual pointer is rendered with the scene and thus occlusion as in every day life can be implemented. It is possible to move the virtual part of the hybrid instrument behind a virtual object without occlusion anomalies occurring. Without a virtual tip, occlusion anomalies would occur as soon as the physical instrument moved behind a virtual object. Because Cubby allows viewing from many angles the user can choose a viewpoint from where objects in the virtual scene are not in conflict with the physical part of the pointer. As the virtual tip and the scene are rendered simultaneously there is no delay between them. The delay is cushioned between the physical barrel and the virtual tip. Since the user tends to concentrate on the tip the hybrid instrument lends itself to accurate manipulation. This approach is similar to that of Ishii et al. (1994) with the difference that with Cubby the instrument is not mounted on a mechanical, force-feedback arm.

To minimise user fatigue Cubby’s horizontal projection screen is mounted flush in a frame which forms a hand rest. While currently covered with a single layer of felt only, this frame could be upholstered for improved comfort. Also, when the user releases a virtual object, the object does not fall down but stays in place. Although this is not a faithful reproduction of object behaviour in everyday life, it is beneficial from a fatigue point of view, as it allows the user to rest and pick up from where he left.

As was discussed in Chapter 6, Cubby suffers less from clipping than single screen, movement parallax based systems. Even though the virtual scene appears in front of the screens, the user can move within a large viewing volume without clipping occurring. When a virtual object is clipped by the inner edge of one screen, the clipped part re-appears on the adjacent screen.

Because Cubby's display is based on projection technology rather than CRT technology, the electro-magnetic tracking system used for tracking the instrument is not affected by it. Cubby's display space and the table supporting it are made completely out of wood and plastics, with the projectors being the nearest metal objects. Therefore the tracker's magnetic field is not warped and high accuracy can be maintained within Cubby's display and manipulation space.

Experiment

Aim

The aim of the experiment described in the remainder of this chapter is to investigate whether Cubby's three orthogonal screens and unified nature, improve performance during manipulation. Can the extra expense and complexity of Cubby's three screens be justified? Does the unification of display and manipulation space have a beneficial effect on performance, or might they just as well remain separated? In the experiment subjects had to complete a three dimensional puzzle of an icosahedron, a regular twenty-sided object (Figure 9.4). In order to solve the puzzle both the position and the orientation of the pieces was important. An analogy can be drawn with tasks in for example craniofacial surgery, in which a stray piece of skull or a prosthesis needs to be fitted into the existing structure of the skull (Groen, 1998; Van Hattem, 1995). In case of a prosthesis, craniofacial surgery also requires modelling the form of the prosthesis. This is an aspect which is not reflected in the puzzle task in this experiment.

Experiment terminology

To prevent confusion I will first define the puzzle related terminology as it is used in this chapter.

1. A *polygon* is one single triangle of the total of twenty

Figure 9.4
The icosahedron shown from two different viewpoints.

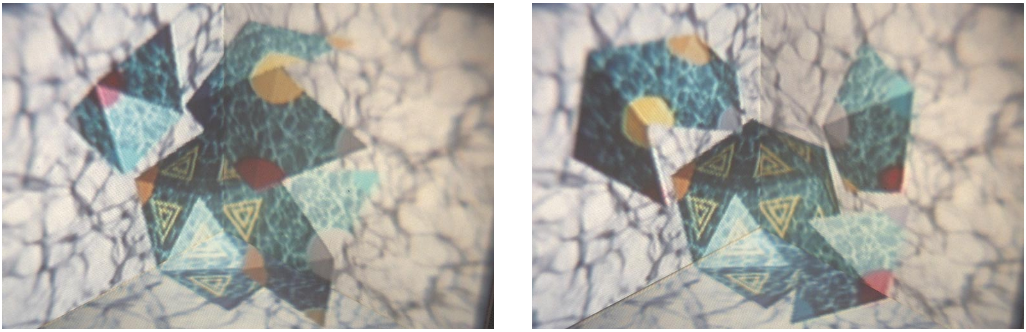


Figure 9.5
Two puzzle arrangements, a 10, 4, 3, 3 (left), and a 10, 5, 3, 2 (right). The puzzle base is the piece consisting of 10 polygons with the concentric triangles on each polygon.

triangles in the icosahedron.

2. One part of the icosahedron cannot be moved. This part consists of ten polygons and is called the *puzzle base*. To communicate to the subject that polygons belong to the puzzle base and cannot be moved, they are marked by three concentric, yellow triangles (Figure 9.5).
3. A *puzzle piece* consists of two to four polygons forming a single piece. Subjects are always offered three puzzle pieces which together contain ten polygons (Figure 9.5).
4. A *puzzle arrangement* is a puzzle base with a particular set of three puzzle pieces which together form a complete icosahedron. For example, a puzzle arrangement could be formed by the puzzle base which consists of ten polygons, two puzzle pieces consisting of four pol-

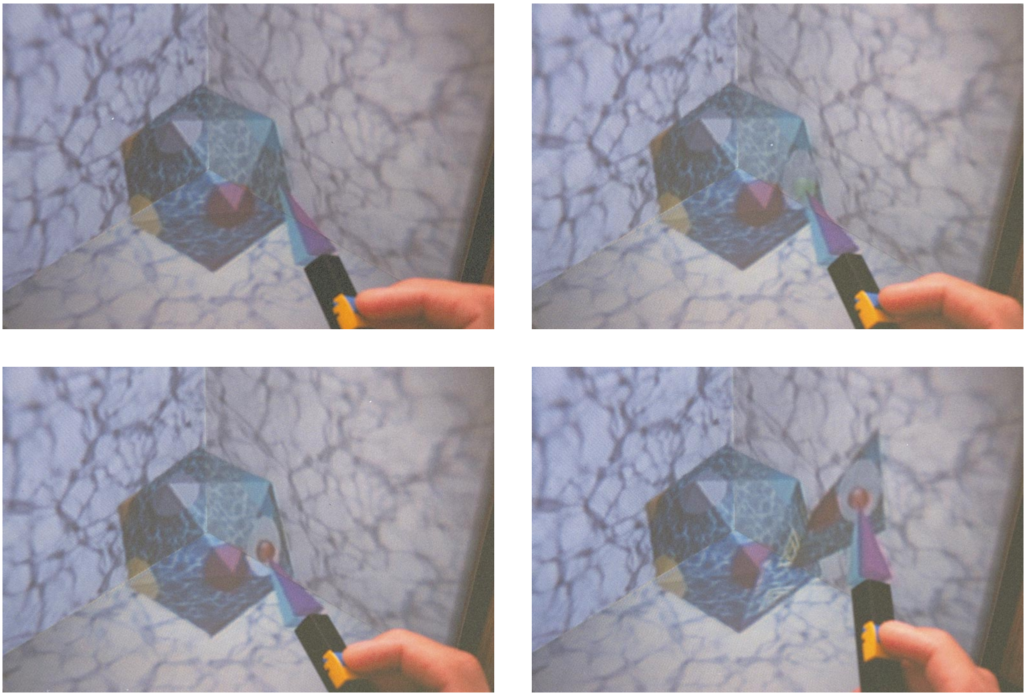


Figure 9.6
Sequence showing visual feedback. The instrument approaches the polygon (top left). As the virtual tip enters the sensitive zone of a polygon the inscribed circle lights up, a small green sphere appears at the point of contact and a collision sound is heard (top right).The closer the tip gets to the polygon, the brighter the inscribed circle and the more saturated the colour of the sphere. After the user presses the button on the instrument, the sphere turns from green to red (bottom left), and the puzzle piece follows the instrument with six degrees of freedom.

ygons and a third puzzle piece consisting of two polygons (10, 4, 4, 2). Or a puzzle arrangement could be formed by a puzzle base, a piece consisting of four polygons, and two puzzle pieces consisting of three polygons (10, 4, 3, 3) (Figure 9.5).

Puzzle behaviour

The interaction between the hybrid instrument and the puzzle is shown in Figure 9.6. When the virtual tip approaches a puzzle piece the nearest polygon lights up, a small green sphere appears which indicates the point of contact, and a collision sound is heard. As the distance between the virtual tip and the polygon decreases, the light intensity of the polygon increases and the colour of the sphere becomes more saturated. This mechanism of visual and auditive feedback is meant to compensate for the lack of haptic feedback. The feedback is a sign to the user that he can pick up a puzzle piece. By pressing the button on the instrument the user closes the tip. The metaphor to think of is a pair of tweezers. As soon as the user closes the tip, the puzzle piece follows the instrument with six DOF.

Design

Task

The subject was asked to put the icosahedron together in such a way that the textures on adjacent faces of the puzzle would correspond. The textures were applied to the puzzle pieces such that the puzzle pieces could only be put together in one particular manner.

Subjects were asked to put together the icosahedron as accurately and as quickly as possible within a time span of three minutes.

Conditions

In total there were four conditions which fit into the orthogonal design shown in Figure 9.7. On the horizontal axis of Figure 9.7, there is either a single screen fishtank display (Figure 9.8, Figure 9.10 and Figure 9.9) or Cubby's three screen Fish Tank VR display (Figure 9.11). For convenience the implementation of the single screen fish tank display as used for this experiment will be referred to as Solo. On the vertical axis the display and manipulation space are either separated or unified. These two independent variables reflect the two questions which are central in this experiment. First, is there any point in using three head tracked displays instead of the conventional single head tracked display? Do the three head tracked screens lead to increased performance during manipulation, do they not make any difference, or do they perhaps lead to decreased performance as a result of, for example, the seams between the screens? Second, does the unification of display and manipulation space lead to improved performance? Or might they just as well remain separated, as is the case in most non-immersive 3D systems?

The four conditions can be summarised as follows:

1. Solo with the display and manipulation space separated. The coupling between instrument and virtual tip is absolute and cannot be decoupled. The display and manipulation space are equal in size but the manipula-

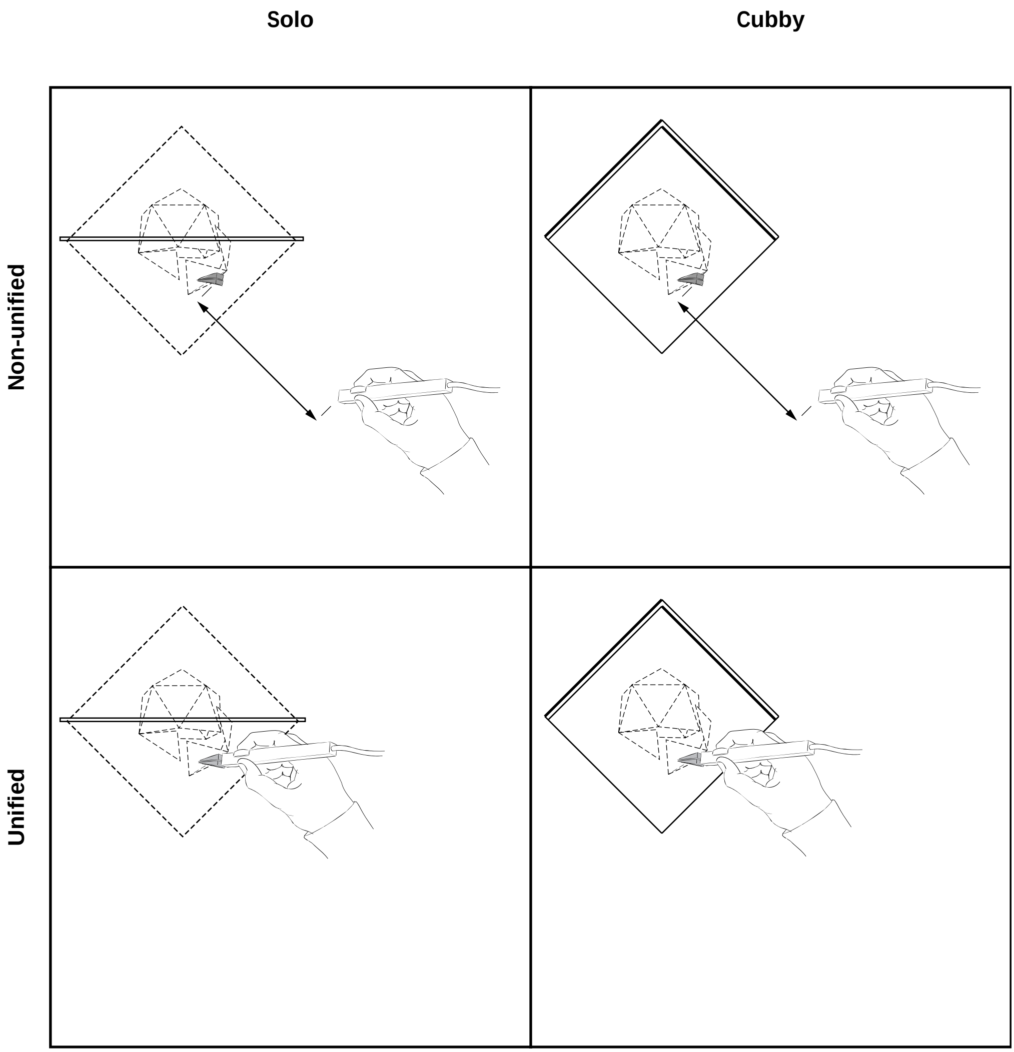


Figure 9.7
The four conditions of the manipulation experiment

tion space has been translated by 200mm in the direction shown in Figure 9.7.

2. Cubby with the display and manipulation space separated. The coupling between instrument and virtual tip is absolute and cannot be decoupled. The display and manipulation space are equal in size but the manipulation space has been translated by 200mm in the direction shown in Figure 9.7.

Figure 9.8
The icosahedron seen from two different points of view in Solo.

3. Solo with the display and manipulation space unified. Since this condition is unified, the coupling between instrument and virtual tip is absolute by necessity and cannot be decoupled. The icosahedron was oriented in such a way that the puzzle pieces could be manoeuvred into place without the screen forming a barrier for the physical barrel.
4. Cubby with the display and manipulation space unified. Since this condition is unified, the coupling between instrument and virtual tip is absolute by necessity and cannot be decoupled.

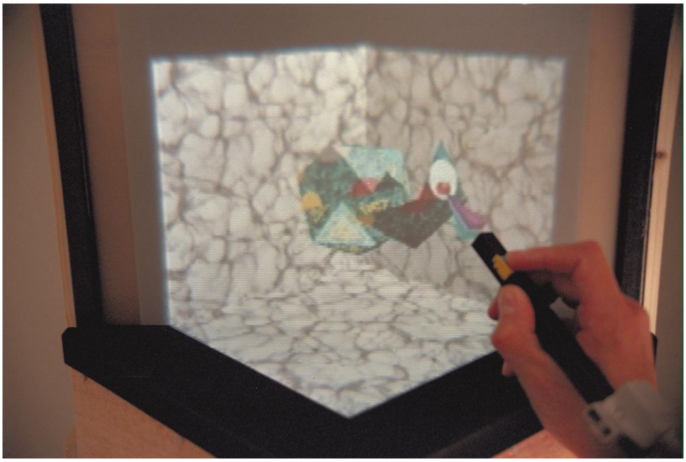


Figure 9.9
Manipulating the icosahedron in Solo.

Figure 9.10
An observer manipulating virtual objects in Solo.

There were twelve subjects who each completed all four conditions. In each condition the subjects were first offered two trials to practise, followed by five trials which were used as data. Two puzzle arrangements were reserved for the practice trials and five for the genuine ones. The same two practice puzzle arrangements and five genuine ones were used in all four conditions. For each subject both the order of the practice puzzles and that of the genuine puzzles were randomised. The puzzle

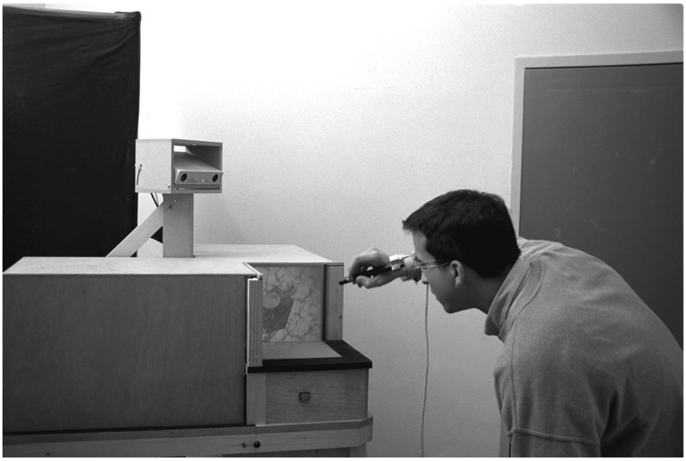
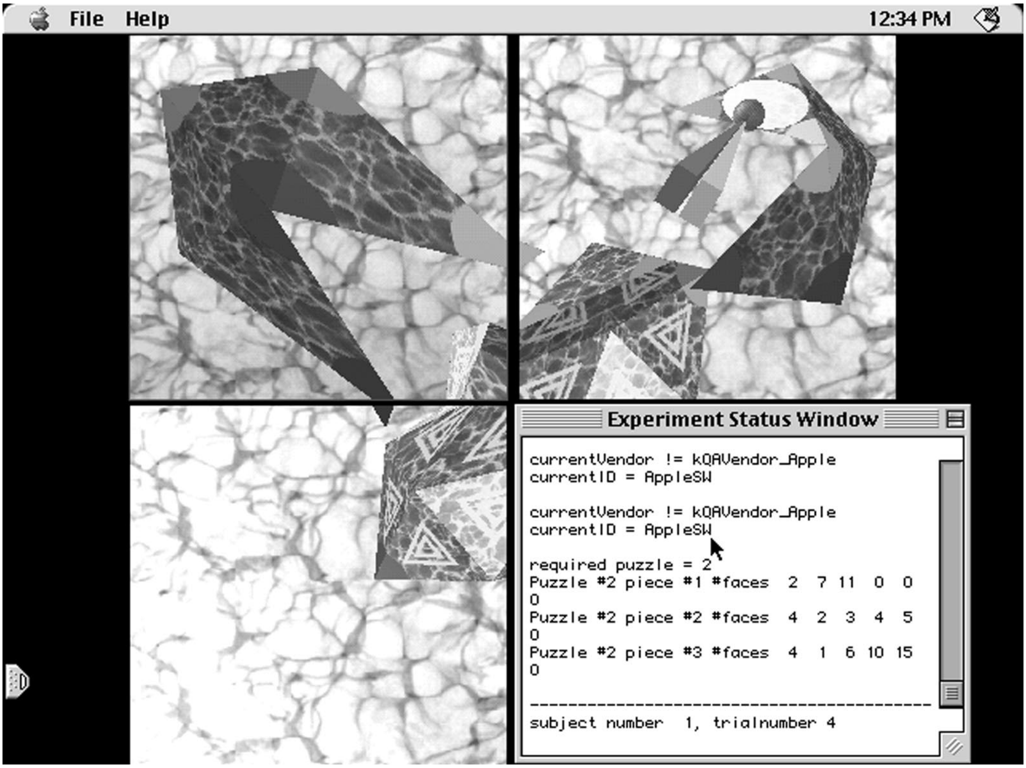


Figure 9.11
Manipulating the icosahedron in Cubby.



```
currentVendor != kQAVendor_Apple
currentID = AppleSW

currentVendor != kQAVendor_Apple
currentID = AppleSW

required puzzle = 2
Puzzle #2 piece #1 #faces  2  7 11  0  0
0
Puzzle #2 piece #2 #faces  4  2  3  4  5
0
Puzzle #2 piece #3 #faces  4  1  6 10 15
0

------------------------------------------
subject number  1, trialnumber 4
```

Figure 9.12
The three images which are projected on the three screens of Cubby, as seen on a conventional monitor. Note how the images look distorted and are difficult to make sense of when not displayed in Cubby.

arrangements were designed by eye, such that the starting positions and orientations of the puzzle pieces were approximately equally difficult.

The three puzzle pieces to be positioned by the subject were different in each trial. However, when put together the orientation of the resulting icosahedron was the same in all trials, with the coloured apices ending up in the same positions. A learning effect could be expected with performance increasing with the order of the conditions. Therefore the conditions were offered in counterbalanced order.

After a week the twelve subjects came back for a second session in which they repeated all four conditions. A second session was held to make positive that subjects had reached their maximum on the learning curve in all four conditions. An interval of one week was chosen because the first session took up four days (three subjects a day), because of the wish to complete the experiment as soon as possible, and to keep the interval between the two sessions the same for all subjects. In the second session the conditions were offered in the reverse order of the first session.

The main dimensions of the icosahedron and the display space are shown in Figure 9.13. Both in Solo and Cubby the icosahedron is displayed in the middle of the display space. The icosahedron has a diameter of approximately 10cm, or half the length of the edge of the cubic display space. In Solo the icosahedron thus projects 5cm out of the screen and 5cm into it.

Subjects

Of the twelve subjects eight were male and four were female. All twelve subjects were students from the faculty of industrial design engineering, except for one female who was an architecture student. Subjects were paid fifty guilders for two sessions. The best performer could win an additional fifty guilders.

Apparatus

The images were rendered at approximately 30fps by a computer (Apple PowerMacintosh 9500) fitted with a processor upgrade (Newer Technologies MaxPower200Mhz) and an accelerated 3D graphics board (Newer Technologies RenderPIX502). Measurement of the hybrid instrument's position and orientation, as well as detection of its button state, were done through an electro-magnetic, six DOFs tracker (Ascension Flock of Birds). A

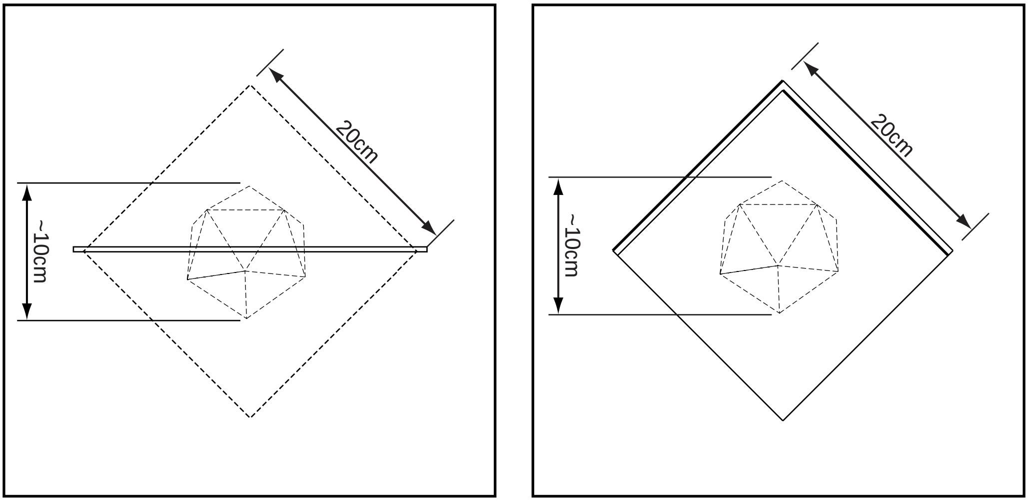


Figure 9.13
The positioning of the icosahedron within Solo (left) and within Cubby (right).

receiver was mounted on the back end of the physical barrel of the hybrid instrument. The emitter, relative to which the receiver reports its position and orientation, was mounted underneath the screen in Solo's case, and within one of the light shields in case of Cubby. All other components were the same as in the set-up used in chapter 8.

Procedure

In all four conditions subjects wore a pair of glasses with one eye blocked. They thus viewed the puzzle monoscopically. Subjects were first shown an assembled icosahedron and were asked to pull it apart to get familiar with the hybrid instrument, the behaviour of the puzzle and the particular condition. Then they were asked to put it together again. Subsequently, they were given the two practice trials after which the genuine trials followed. After they had completed a trial to their satisfaction, they had to press a button. When the three minute time limit was reached the trial was automatically terminated. Twenty seconds before, an auditive warning would sound. This procedure was repeated for each condition. After having completed two conditions, subjects were allowed to pause for ten minutes. In total it would take a subject approximately two hours to complete one session of the experiment. At the end of each session subjects were asked to rank the four conditions in order of preference.

Dependent variables

The following variables were recorded for each trial:

1. the position of each puzzle piece
2. the orientation of each puzzle piece

The positions and orientations of each puzzle piece as adjusted by a subject, were calculated relative to the centre of gravity of its constituent polygons. From the geometry of the icosahedron model, for each puzzle piece the positions and orientations relative to their centre of gravity for which all the pieces fit together exactly were calculated. Through comparison of the measured positions and orientations with the 'ideal' positions and orientations, the positional and rotational error were found.

Hypotheses

It was expected that subjects would perform more accurately when using Cubby's display method than when using Solo's display method. This was expected on the ground that when viewed under large angles the puzzle appears less distorted in Cubby than in Solo. Let's call the front-rear diagonal of the display space the neutral axis. With Cubby, if the user makes a large angle with the neutral axis he will make a small angle with the normal of the screen he is looking at. In other words, he looks at one of the screens almost head on. With Solo, however, if the subject makes a large angle with the neutral axis, he is close to the screen and makes a large angle with the screen normal. The result is that small errors in position measurement lead to large displacements on screen.

It was also expected that within one type of display method, subjects would perform more accurately in the unified than in the non-unified condition. This was expected on the ground that hand-eye coordination in the unified conditions more closely resembles everyday life than the hand-eye coordination in the non-unified conditions.

The hypotheses are be drawn up for both the rotational error (Table 9.1) and the positional error (Table 9.2). When one of the alternative hypotheses is accepted, the conditions will be compared in pairs.

Table 9.1 Hypothesis for mean rotational error

H_0	$M_{rotCU}=M_{rotCNU}=M_{rotSU}=M_{rotSNU}$
H_1	not H_0
Note:	rot = rotational error, CU=Cubby unified, CNU=Cubby non-unified, SU=Solo unified, SNU=Solo non-unified.

Table 9.2 Hypothesis for mean positional error

H_0	$M_{posCU}=M_{posCNU}=M_{posSU}=M_{posSNU}$
H_1	not H_0
Note:	pos = positional error

Results

Analysis of the data showed that in many trials subjects had not got round to adjusting all of the puzzle pieces within the three minutes. Puzzle pieces which had not been adjusted showed high positional and rotational errors. These were deemed to not be representative of performance in a certain condition, yet they had a considerable impact on the mean error. This problem was reduced in a two step process. The first step was to exclude from the analysis those puzzle pieces which had a positional error of 10000 dynasight steps (50cm) or more, and those which had a rotational error of 30 degrees or more. Such large differences were considered to be unintentional. In this way, 6.7% of the data in session 1, and 2.1% of the data in session 2 was eliminated from analysis. The second step was to use the median instead of the mean as a measure of central tendency, because the mean is sensitive to extreme values of positional or rotational error

Results for positional and rotational error

Session 1

First, it is determined whether the four conditions have equal medians through the extension of the median test. This is done both for rotational (Table 9.3) and positional error (Table 9.4).

Table 9.3 The extension of the median test for rotational errors, session 1

	CU	CNU	SU	SNU
#trials (rotational error>common median)	74	64	94	104
#trials (rotational error<=common median)	104	88	75	69
total	178	152	169	173
df=3; χ^2=18.06; p<0.001**				

Table 9.4 The extension of the median test for positional errors, session 1

	CU	CNU	SU	SNU
#trials (positional error>common median)	83	64	86	103
#trials (positional error<=common median)	95	88	83	70
total	178	152	169	173
df=3; χ^2=10.95; p<0.05*				

The null-hypothesis can be rejected both for the rotational and positional errors: the four conditions do not share the same median. I will therefore proceed by comparing the following pairs: CU-CNU (Table 9.5 and Table 9.6), SU-SNU (Table 9.7 and Table 9.8), CU-SU (Table 9.9 and Table 9.10) and CNU-SNU (Table 9.11 and Table 9.12).

Table 9.5 The median test for rotational error in CU vs. CNU, session 1

	CU	CNU
#trials (positional error>common median)	88	77
#trials (positional error<=common median)	90	75
total	178	152
df=1; χ^2=0.012; 0.4< p<0.45		

Table 9.6 The median test for positional error in CU vs. CNU, session 1

	CU	**CNU**
#trials (positional error>common median)	91	74
#trials (positional error<=common median)	87	77
total	178	152
df=1; χ^2=0.074; 0.35< p<0.4		

Table 9.7 The median test for rotational error in SU vs. SNU, session 1

	SU	**SNU**
#trials (rotational error>common median)	81	90
#trials (rotational error<=common median)	88	83
total	169	173
df=1; χ^2=0.42; 0.25< p<0.35		

Table 9.8 The median test for positional error in SU vs. SNU, session 1

	SU	**SNU**
#trials (positional error>common median)	78	93
#trials (positional error<=common median)	91	79
total	169	173
df=1; χ^2=1.84; 0.1< p<0.5		

CU and SU thus differ significantly in terms of rotational error. CNU and SNU differ significantly both in rotational and positional error.

All other pairwise comparisons show no significant differences.

Table 9.9 The median test for rotational error in CU vs. SU, session 1

	CU	SU
#trials (rotational error>common median)	76	97
#trials (rotational error<=common median)	102	72
total	178	169
df=1; χ^2=6.92; 0.0005< p<0.005*		

Table 9.10 The median test for positional error in CU vs. SU, session 1

	SU	SNU
#trials (positional error>common median)	85	88
#trials (positional error<=common median)	93	81
total	178	169
df=1; χ^2=0.49; 0.25< p<0.35		

Table 9.11 The median test for rotational error in CNU vs. SNU, session 1

	CNU	SNU
#trials (rotational error>common median)	61	102
#trials (rotational error<=common median)	91	71
total	152	173
df=1; χ^2=10.73; 0.0005< p<0.005*		

Session 2

Again I start with the extension of the median test to investigate whether the four conditions share the same median. The test is carried out for rotational error (Table 9.13) and for positional error (Table 9.14).

Since the null-hypothesis can be rejected for the rotational error only, the pairwise testing will only be carried out on the rotational error, not on the positional error.

Table 9.12 The median test for positional error in CNU vs. SNU, session 1

	CNU	SNU
#trials (positional error>common median)	63	100
#trials (positional error<=common median)	89	73
total	152	173
df=1; χ^2=8.36; 0.0005< p<0.005*		

Table 9.13 The extension of the median test for rotational error, session 2

	CU	CNU	SU	SNU
#trials (rotational error>common median)	79	67	111	96
#trials (rotational error<=common median)	98	104	69	81
total	177	171	180	177
df=3; χ^2=21.17; p<0.001**				

Table 9.14 The extension of the median test for positional error, session 2

	CU	CNU	SU	SNU
#trials (positional error>common median)	80	79	102	92
#trials (positional error<=common median)	97	92	78	85
total	177	171	180	177
df=3; χ^2=6.45; 0.1<p<0.2				

Again the pairs used in testing are CU-CNU (Table 9.15), SU-SNU (Table 9.16), CU-SU (Table 9.17) and CNU-SNU (Table 9.18).

CU-SU and CNU-SNU thus differ significantly in terms of rotational error. Cubby allows subjects to manipulate virtual objects with higher accuracy than is possible with Solo. All other pairwise comparisons show no significant differences.

Table 9.15 The median test for rotational error in CU vs. CNU, session 2

	CU	CNU
#trials (rotational error>common median)	91	83
#trials (rotational error<=common median)	86	88
total	177	171
df=1; χ^2=0.18; 0.25< p<0.35		

Table 9.16 The median test for rotational error in SU vs. SNU, session 2

	SU	SNU
#trials (rotational error>common median)	98	81
#trials (rotational error<=common median)	82	96
total	180	177
df=1; χ^2=2.35; 0.05< p<0.1		

Table 9.17 The median test for rotational error in CU vs. SU, session 2

	CU	SU
#trials (rotational error>common median)	73	106
#trials (rotational error<=common median)	104	74
total	180	177
df=1; χ^2=10.42; 0.0005< p<0.005*		

Results for the ranking of conditions in order of decreasing preference

At the end of a session subjects were asked to rank the four conditions according to preference. The condition preferred the most by subject was put in first place, and the condition preferred the least was put in fourth and last place. When subjects rated two conditions the same, each condition would be given the mean of the two rankings. For example, if a subject judged that two conditions

Table 9.18 The median test for rotational error in CNU vs. SNU, session 2

	CNU	SNU
#trials (rotational error>common median)	74	100
#trials (rotational error<=common median)	97	77
total	171	177
df=1; χ^2=5.56; 0.005< p<0.01*		

shared the most preferred rank, these two conditions would both be rated 1.5. From these rankings a mean rank was found. These are shown in Figure 9.14 for session 1 and in Figure 9.15 for session 2. Not only do subjects prefer Cubby over the single screen Fish Tank display, within these set-ups they also prefer the unified conditions over the non-unified conditions.

Discussion

The results show that subjects can manipulate virtual objects with significantly higher accuracy in the Cubby unified condition than they can in the single screen unified condition. Likewise, they perform with higher accuracy in the Cubby non-unified condition than in the single

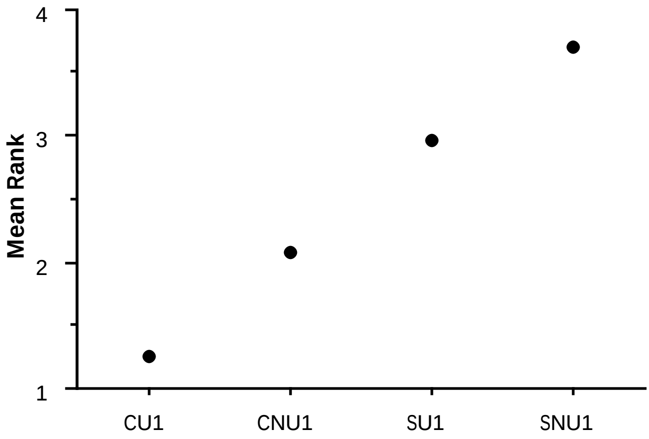
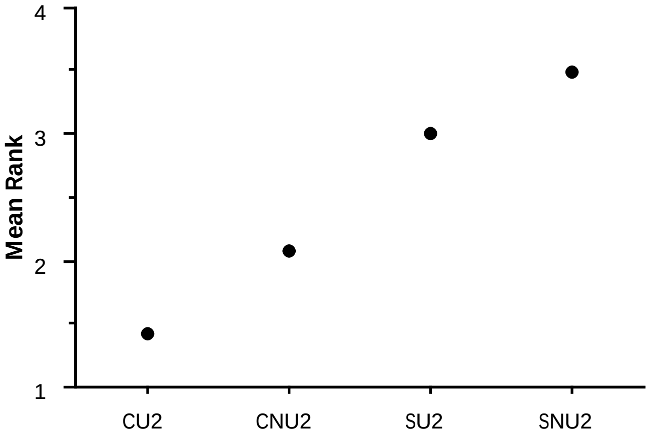


Figure 9.14
The mean rank for session 1.

Figure 9.15
The mean rank for session 2.



screen non-unified condition. No differences in performance could be found between the unified and non-unified conditions for either Cubby or the single screen fishtank display. However, when subjects are asked to rank the conditions according to preference, a pattern emerges which shows not only that Cubby is preferred over the single screen fishtank set-up, but also that within a set-up the unified condition is preferred over the non-unified condition.

The higher performance in Cubby and the fact that Cubby was preferred over Solo cannot be solely attributed to the puzzle appearing less distorted under large viewing angles in Cubby than in Solo. There is a more pragmatic second factor which is related to the infra-red head tracking and the fact that subjects keep their body close to the scene. This behaviour is not restricted to the unified condition in which the subject needs to put his hand in the scene. It occurs in both the unified and the non-unified conditions because subjects look closely at the scene to try to see things more accurately. With Cubby, the Dynasight tracker was mounted behind the display space at approximately head-height. Because Cubby's display space was open, the Dynasight's infra-red beam was not obscured, even when the subject held his eye close to the scene at puzzle-height. However, with Solo, it was not possible to mount the Dynasight tracker in the same location, since the projection screen would obscure the infra-red beam if

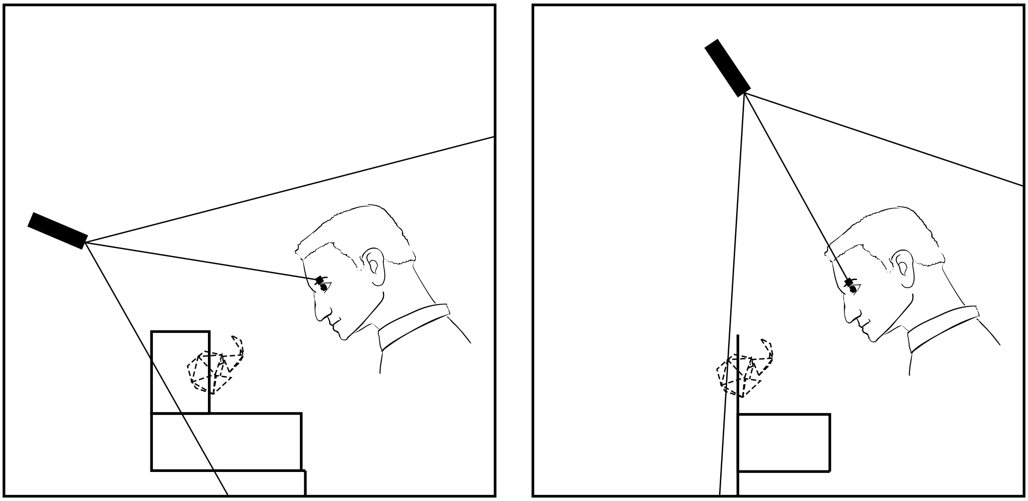
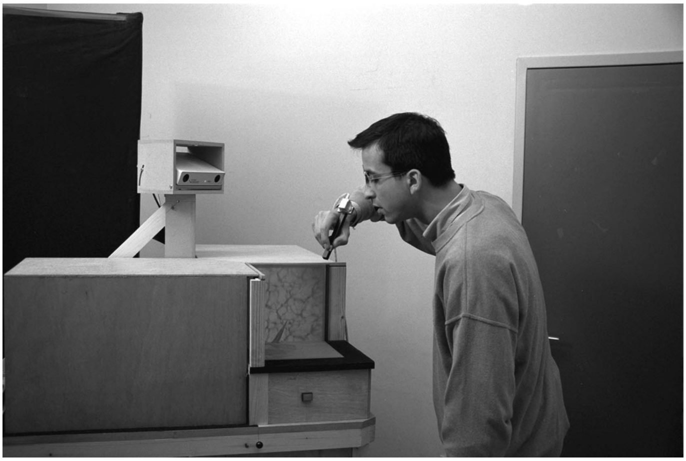


Figure 9.16
The positioning of the Dynasight base unit, the Dynasight reflective disc and the puzzle, for Cubby (left) and for Solo (right).

the user would hold his eye close to the scene at puzzle-height. Moving the tracker towards the user was no solution to this problem, since it would limit the freedom of lateral movement. To avoid the infra-red beam being obscured, and to achieve freedom of lateral movement in Solo comparable to that in Cubby, the Dynasight tracker was mounted high up under an angle, and the reflective disc attached to the spectacle frame was turned upwards towards the tracker. Figure 9.16 shows the position of the Dynasight tracker for both Solo and Cubby. While this did indeed provide the desired freedom of lateral movement, it had the side effect that if the subject tried to look down on the scene and tilted his head far forward, this could lead to occlusion of the reflective disc. The likelihood of occlusion depended on the subject’s height, on the form of the head, on how the frame fitted his face. Subjects with long hair were asked to wear a hair band to prevent occlusion as much as possible. Still, Solo is much more sensitive to this problem than Cubby in which the Dynasight tracker is mounted lower and in which the reflective disc is not obscured by the subject’s head when he looks down on the scene.

Figure 9.17
A user looking down on the virtual scene from above.



It should be emphasised that some of Cubby's qualities were not exploited in this experiment. In order to make a comparison between Cubby and a single screen Fish Tank VR display possible, the experiment was set up in such a way that two of Cubby's main features were left unused. The first feature concerns the size of the space which is effective both as display and manipulation space. With Cubby's display method the whole of the space formed by the projection screens is both display and manipulation space. With Solo the space formed by the virtual background planes is not completely accessible to the user's instrument in the unified condition. If a virtual object is more than the length of the virtual tip of the hybrid instrument behind Solo's projection screen, it is impossible to manipulate it. The user simply cannot reach it. In the experiment this was prevented by keeping the starting positions of the puzzle pieces within reach of the instrument. Cubby's second feature concerns possible viewing angles. With Cubby it is possible to look down on the display space (Figure 9.17). This is something which Solo does not allow. In the experiment the puzzle was placed in such a way that subjects did not need to look down onto the scene in order to solve the puzzle. However, in surgical simulation it would be an advantage to be able to look down onto the scene.

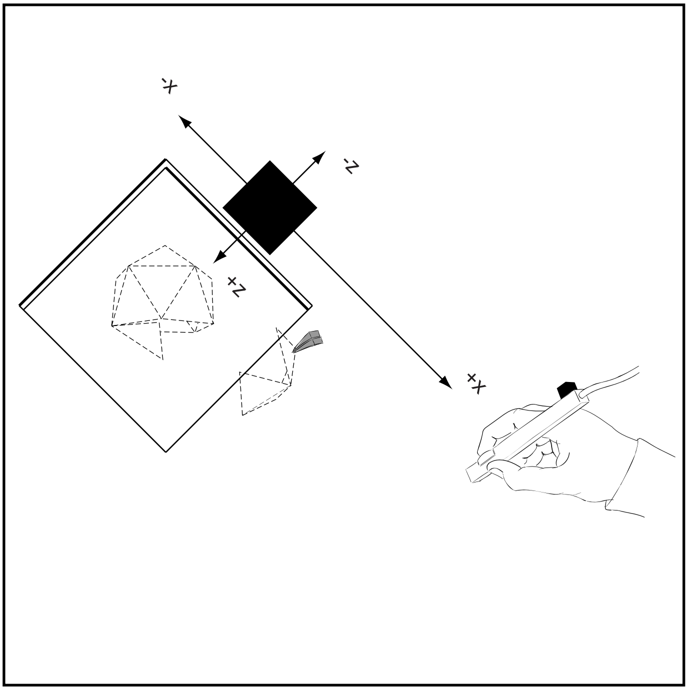


Figure 9.18
In the Cubby non-unified condition the back end of the instrument with the receiver could end up in negative Z-space. The big black box represents the emitter of the tracking system, the small black box on the back of the instrument represents the receiver.

In a handful of trials in the Cubby non-unified condition, subjects complained about mysteriously losing a puzzle piece. It was found afterwards that they had moved the receiver behind the emitter of the tracking system used for the hybrid instrument (Figure 9.18). The emitter has two so-called hemispheres, one covering the positive z-axis the other the negative z-axis. The receiver only gives consistent positional readings within one hemisphere. When the boundary between the two hemispheres is crossed, the x coordinate changes sign. This only happened in the Cubby non-unified condition where the subjects kept the instrument outside Cubby's display space while still seeing the virtual tip within it. It was possible that they oriented the instrument in such a way that the back of the instrument, carrying the receiver, ended up in the other hemisphere. To the user this gave a highly disturbing effect: both the virtual tip and the currently held puzzle piece simply vanished. If the subject kept the button pressed he could move the instrument in such a way that the receiver re-entered the correct hemisphere, after which both tip and puzzle piece would re-appear. However, when confronted with this strange visual behaviour the subject was likely to release the button in confusion. The result was that the puzzle piece was ‘lost in cyberspace’: it had become difficult to retrieve because, being positioned on the opposite side of the display space, it was invisible to the user. Nevertheless, the puzzle piece did not cease to exist, it ‘lived on’ somewhere else. If the user could move Cubby’s display space or had a PDA like the one by Fitzmaurice et al. (1993), and were to hold it near the puzzle piece, he could ascertain himself of its reclusive presence. Another way to make the puzzle piece visible for the user, would be to make the textures on Cubby’s projection screens transparent. The user would then be able to see the puzzle piece through the screens, on the other side of Cubby’s display space.

Informal observation showed that users quickly become accustomed to a virtual world existing in parallel with the real world. When playing about with Cubby some users would bring a puzzle piece within inches from their faces, and discover that the puzzle piece would indeed take up a large visual angle and cover almost all of the projection screens. A few subjects tried positioning a puzzle piece behind their heads, then stepped back to get it back into view again. Clearly they understood that ‘out of sight’ did not mean ‘out of existence’. One subject even ‘parked’ a puzzle piece behind his ear to momentarily stretch his arm. Then – without moving his head – he reached back behind his ear, groping about within the few cubic inches where he was sure to have left the puzzle piece. Indeed he found back the puzzle piece by auditive feedback. He then picked up the puzzle piece from behind his ear, moved it back into the display space and continued putting the puzzle back together again.

Conclusions

Subjects can manipulate virtual objects in Cubby with significantly higher accuracy than in a single screen Fish Tank set-up. No significant difference in performance between the unified and non-unified conditions could be found. However, when asked to rank the four conditions in order of preference, subjects not only indicate a preference for Cubby over the single screen Fish Tank set-up, but also a preference for unified over non-unified within the set-ups. The higher performance in Cubby and the subjects' preference for Cubby may be the result of reduced distortion and more freedom of movement, as compared to the single screen set-up. Finally, it needs to be pointed out that in order to make a comparison between Cubby and a single screen Fish Tank VR display possible, two of Cubby's features were left unexploited. Cubby's first advantage is that the full space set-up by the three orthogonal screens can be accessed by the virtual instrument, as opposed to the single screen system in which this space is much more limited. Cubby's second asset is that the user can look down onto the scene. With the single screen system this is not possible. Clearly, from an application point of view these are important benefits. Finally, informal observation showed that users quickly get used to Cubby and the idea of a virtual world existing in parallel with the real world.

10

Conclusions

On Cubby and medical 3D

Cubby is only a proof of concept. It lacks certain characteristics essential for a medical simulator, such as high definition, tissue behaviour modelling, and integration with medical scanners. All these characteristics form research areas in themselves. What I believe is important is that none of the essence of Cubby clashes with the demands made by these characteristics. Integrating some or all of these characteristics in Cubby should not present any significant problems.

On the unification of the display and the manipulation spaces

In designing Cubby, the emphasis has been on the display method using head-coupled movement parallax on three orthogonal, desktop-sized screens. As became evident, this display method allows more intuitive manipulation than existing movement parallax systems can offer. Cubby should be seen as proof that, by radically changing the visualisation method, manipulation can suddenly become much more intuitive. Although it was shown in an experiment that it is possible to manipulate with greater accuracy in Cubby than in a single-screen Fish Tank set-up, the performance advantage of unified display and manipulation spaces over non-unified spaces could not be confirmed experimentally. However, subjects did have a preference for a unified version over a non-unified version of the same set-up.

On the technical implementation of Cubby

Apple's QuickDraw3D technology proved highly valuable as a prototyping tool for Cubby. However, if Cubby is to be used for medical voxel-based 3D reconstructions, another graphics library will need to be used, since QuickDraw3D currently does not support voxel rendering, only surface rendering.

Cubby's hardware set-up grew organically over the last few years, and many changes were made along the way. As computer technology has improved rapidly, some of Cubby's components from its early days are already out-of-date, and seem a strange choice today. Cubby's components are tightly interlinked and changes to one component of the set-up tend to require changes to others. For example, when the first Cubby prototype was put together, purchasing three projectors with a computer interface did not fit within the budget. At the time, it was considerably less expensive to purchase three projectors with a video interface, and three scan converters to convert the computer monitor signal to video. Today it would actually be less expensive to purchase projectors which can be connected directly to the computer. This would eliminate the scan converter(s), thus reducing the number of Cubby's components, as well as improving the quality of the projected images. In turn, this would have an effect on the graphics board. Currently, only a single graphics board is used for all three images. Using three graphics boards and three scan converters would be prohibitively expensive. However, if the scan converters were eliminated, using three graphics boards instead of a single graphics board in combination with a scan converter, the total cost would be approximately the same. As each projector would then be fed directly by its own graphics board, resolution would be doubled.

On Cubby as an electronic product, rather than a computer

With Cubby, the computer has been pushed into the background. When working with Cubby, the user does not see a computer, nor does he have to adapt his behaviour to instruct the computer. For the experiment described in chapter 9, the computer generating the perspectives and the physical Cubby set-up did not even share the same room. By moving the computer next door, the illusion of a stand-alone electronic product was strengthened. With Cubby, the elements which most disturb the illusion of natural behaviour and the absence of a computer are perhaps not delay and resolution, but the glasses which cover one eye and the tethered instrument.

On taking Cubby into production

Before Cubby can become a viable alternative to existing medical 3D systems, a number of practical problems need to be solved. The most important of these is that the 3D impression suffers considerably from projector misalignment. Unfortunately, this misalignment is a recurring problem because of dimensional instability of the table under temperature fluctuations. While this problem would decrease if the projector-screen distance were smaller, it would be preferable to dispense with the projectors altogether. In the future, a solution may lie with active matrix LCDs Depp and Howard, 1993).What would be needed is an LCD with two borderless edges. Such an LCD would also need a wide viewing angle. Another solution may be to bring in technology from the field of video walls. In a video wall, lenses are put in front of conventional CRT displays to create the illusion that the displays match seamlessly.

On testing human-computer interfaces without computers

I think that human-computer interfacing would greatly benefit if interfacing principles would be tested first without computers. Often interfaces can be simulated through the use of physical artifacts. Avoiding programming altogether can be a great time-saver. Not having to worry about how to implement the software also allows a more creative and less restricted conceptual phase. The experiment of chapter 5 showed how the number of degrees of freedom influenced performance without using computers. Implementing the essence of chapter 3 with computers would have taken far more time and would have involved a team of experts. They would have had to solve problems such as finger-tip position detection and how finger movements determine orientation.

On the theory of affordances and product design

It is clear that the way in which electronic consumer products such as digital watches, microwave ovens and video recorders are currently designed results in a usability problem. In the literature, there is an abundance of examples of products which are confusing and difficult to use. Sadly, there are very few examples of products which are intuitive in use. In chapter 3, I tried to illustrate a completely different approach to formgiving by means of a concept for a video deck. The effects of this formgiving on intuitiveness in use have not yet been experimentally tested. Such experimental testing is essential to lend some credibility to the application of affordances in product design. I would like to think that affordances in product design are style-transcendent. Accordingly, the video recorder concept in chapter 3 should not be judged on the grounds of looking beautiful or ugly, trendy or outdated. A video recorder may afford the required actions regardless of whether it is in a baroque, modernist or deconstructivist style. I think that the theory of affordances can help industrial designers to create products which are easier to use, yet have an aesthetic which reflects the product's creator, its company's identity and its era.

On affordances and Cubby

From informal observation, it appears that users are able to pick up very quickly how to manipulate objects in Cubby with the stylus. As soon as the user puts on his glasses and is encouraged to move around, the virtual scene jumps ‘into perspective’. As soon as he moves the stylus into the display space, the virtual tip appears in line with the stylus, indicating that moving the stylus does indeed have an effect. As soon as the virtual tip touches one of the virtual objects, the user hears the effect and sees a polygon light up and the hinge point appear.

In terms of affordances, the current formgiving of the stylus is still crude. While manipulation of virtual objects with the stylus appears to be highly intuitive, many users do need to be told how to hold the stylus. Often they hold the stylus in the palm of their hand with the thumb on top of the button in a kind of hammer grip, rather than as a pen. This is interesting, as it is an example of how it is not self-evident that ergonomically superior interaction is afforded by a product. Once users have tried the pen grip, they acknowledge its ergonomic superiority over the hammer grip, and do not return to the latter. Apparently, the stylus currently affords a power grip rather than a precision grip. It clearly needs to be redesigned to communicate its pen-like nature more successfully.

Appendix I — Delay Measurements

Summary

This appendix documents the procedure by which the delays in the scan conversion and video projection stages of Cubby MkII were measured. The scan converter was a Televisor Zoom by Displays Technologies and the video projector was a Sony CPJ-100E. The delays were measured for the two components individually.

Scan converter

The delay was measured by means of an oscilloscope with one channel on the VGA-in pin and the other channel on the video-out pin. As input signal a computer generated black and white flicker signal with a mark-space ratio of 1:10 was used. White lasted 17ms and black lasted 170ms. The input and output signals are shown in Figure A1.1. The worst case delay was approximately 15ms or 7.5ms on average.

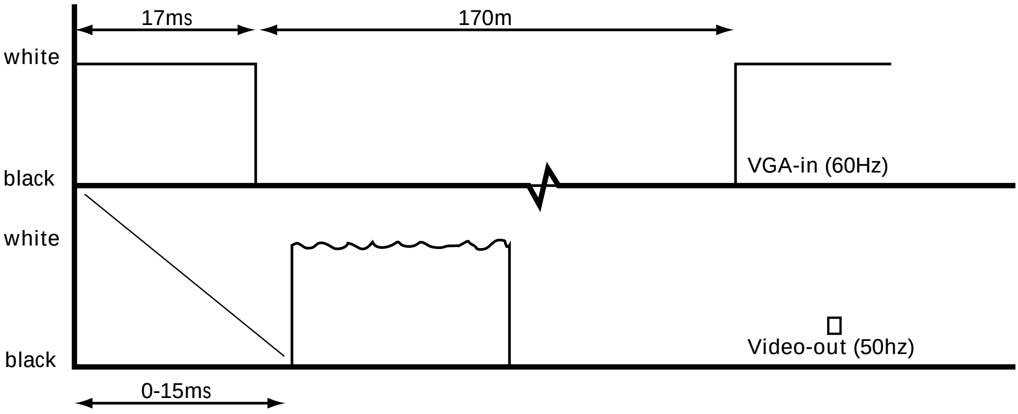


Figure A1.1
The input and output signals of the scan converter.

Video projector

The delay was measured by means of an oscilloscope with one channel on the video-out of the scan converter, which is the same as the video-in of the projector, and the other channel connected to a light dependent resistor which was held near the edge of the projector lens. The input and output signals are shown in. The delay caused by the video projector was approximately 5ms.

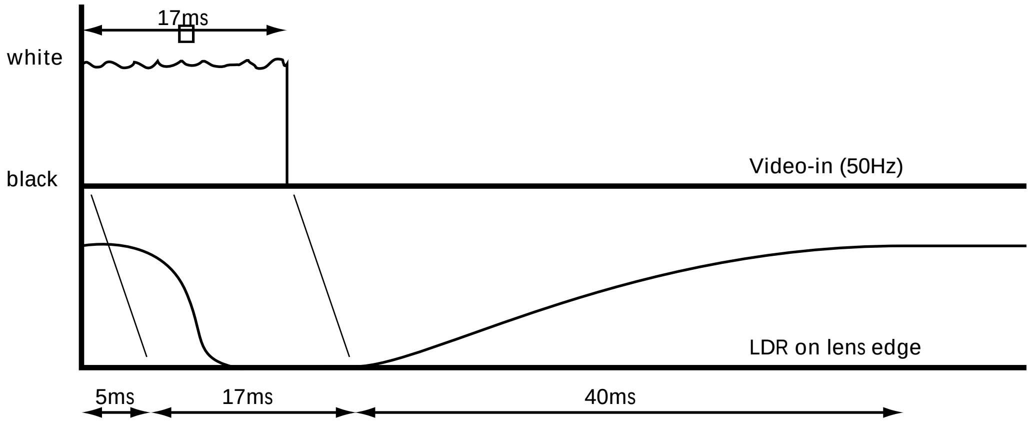


Figure A1.2
The input and output signals of the video projector.

Appendix 2 — Technical Drawings Cubby

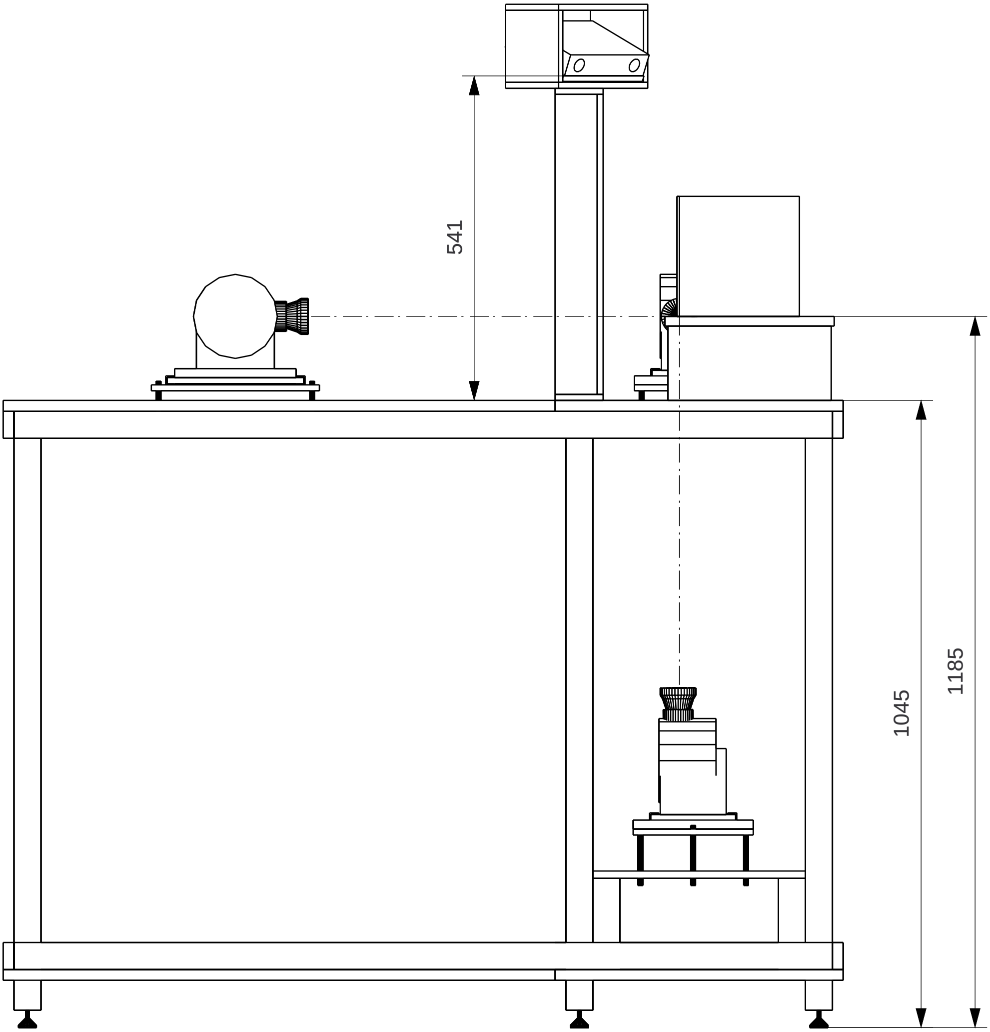


Figure A2.1
Cubby front view.

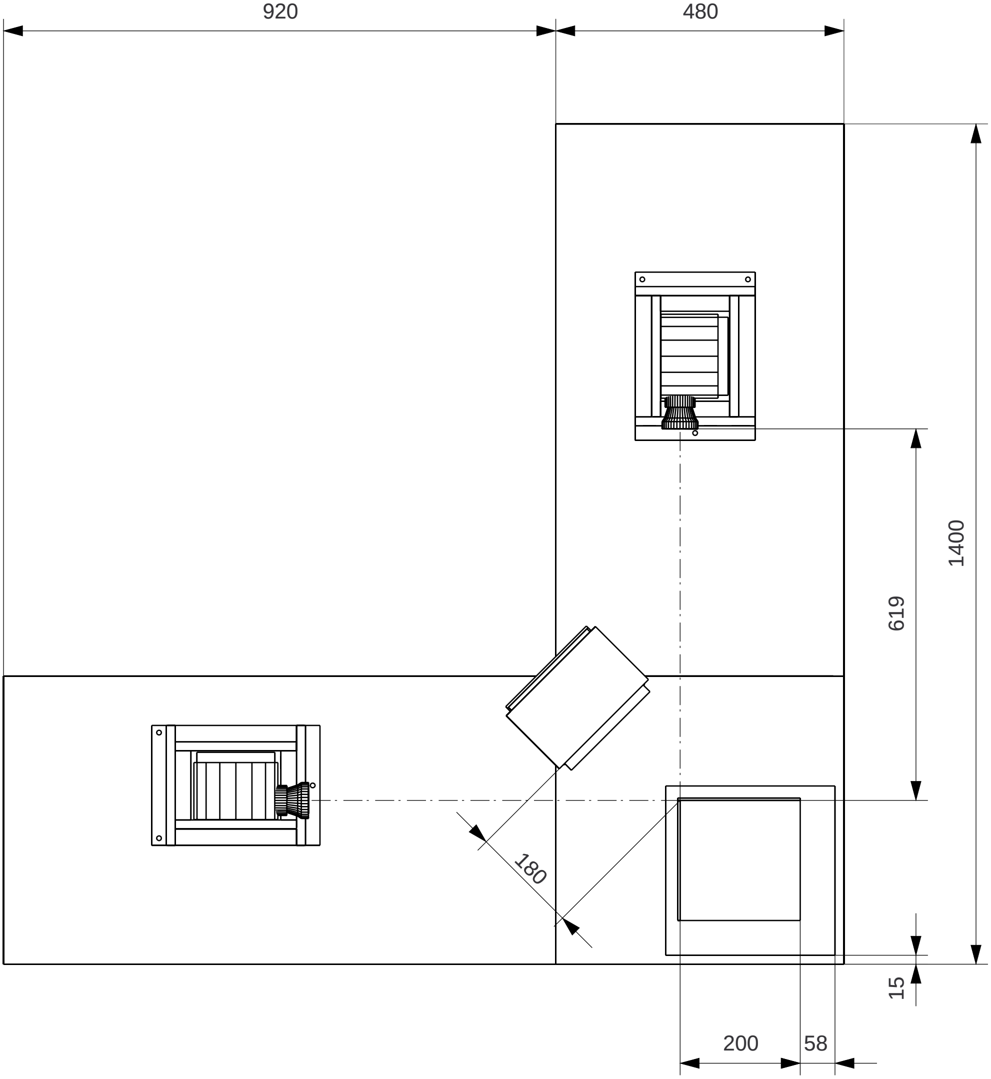


Figure A2.2
Cubby top view.

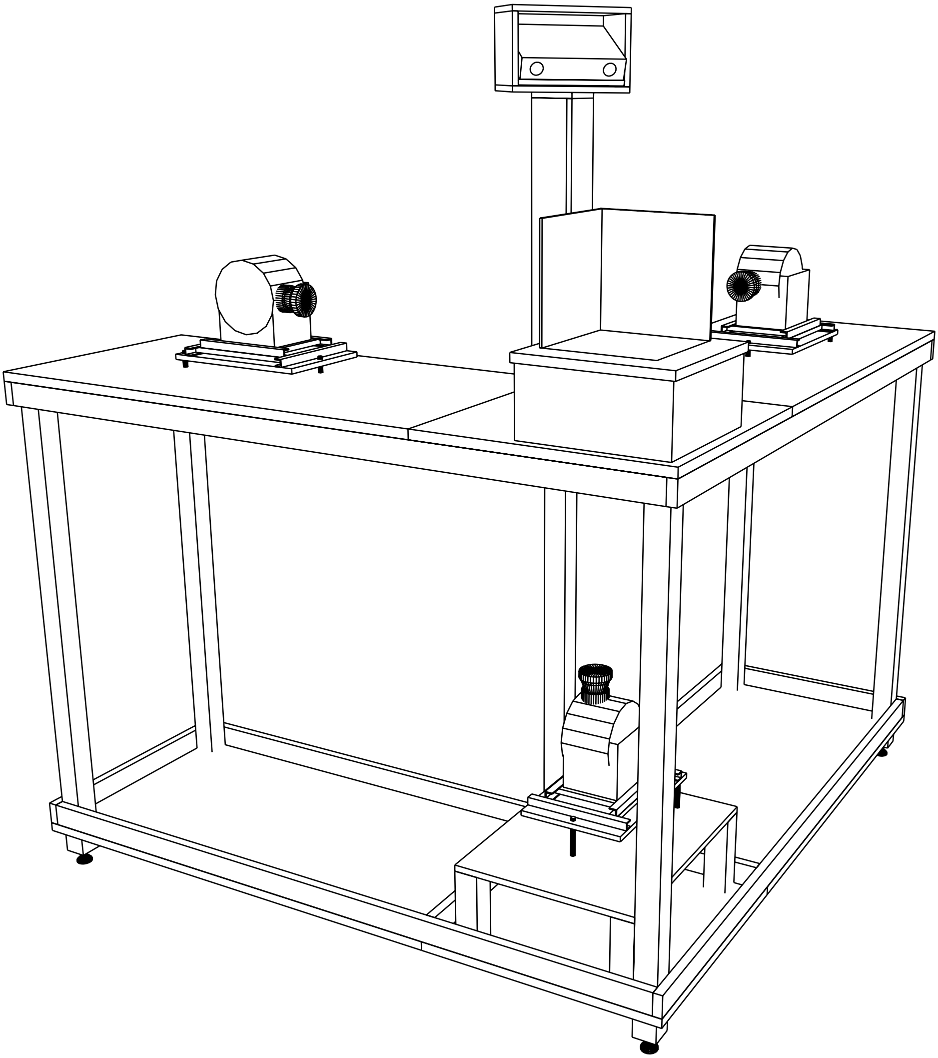


Figure A2.3
Cubby perspective view.

Bibliography

Ahlberg, C., & Schneiderman, B. (1994). The Alphaslider: A compact and rapid selector. *Proceedings CHI'94*, 365-371.

Aldersey-Williams, H., Wild, L., Boles, D., McCoy, K., McCoy, M. Slade, R., & Diffrient, N. (1990). *The New Cranbrook Design Discourse*. New York: Rizzoli International Publications.

Amselem, D. (1995, Spring). A window on shared virtual environments. *Presence, 4* (2), 130-145.

Barthes, R. (1973). *Mythologies*. London: Paladin.

Beaton, R.J., DeHoff, R.J., Weiman, N., & Hildebrandt, P.W. (1987). An evaluation of input devices for 3-D computer display workstations. *SPIE True 3D Imaging Techniques and Display Technologies*, 761, 94-101.

Bentum, M. (1995). *Interactive visualization of volume data*. Unpublished doctoral dissertation, Universiteit Twente, Enschede, The Netherlands.

Brebner, J., & Sandow, B. (1976). The effect of scale side on population stereotypes. *Ergonomics*, 9, 231-238.

Bröckl-Fox, U., Kettner, L., Klingert, A., & Kobbelt, L. (1994). Using three-dimensional hand-gesture recognition as a new 3D input technique. In N.M. Thalmann and D. Thalmann (Eds.), *Artificial Life and Virtual Reality* (pp. 173-187). New York: John Wiley.

Buxton, W. (1986). There's more to interaction than meets the eye: Some issues in manual input. In D.A. Norman and S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction* (pp. 319-337). Hillsdale, NJ: Lawrence Erlbaum.

Buxton, W., & Myers, B.A. (1986). A study of two-handed input. *Proceedings of CHI'86*, 321-326.

Byrne, K. (1990, July). A 'semantics' of visual design: the care and feeding of studio projects within a communication-theory context. *Design Studies, 11* (3), 141-163.

Chen, M. (1993). 3-D rotation using a 2-D input device. *Develop*, 14, 40-53.

Chen, M. (August 1988). A study in interactive 3-D rotation using 2-D control devices. *Computer Graphics, 22* (4), 121-129.

Clymer, F. (1950). *Treasury of early American automobiles*. New York: McGraw-Hill.

Cohen, J. (1979). *Three-dimensional display of parallel sliced images*. Unpublished doctoral dissertation, University of North Carolina.

Cook, L.T., Lee, K.R., Cytacki, E.P., & Dwyer, S.J. (1987). Three-dimensional display of radiation therapy planning. *SPIE True 3D Imaging Techniques and Display Technologies*, 761, 153-156.

Cruz-Neira, C., Sandin, D.J., & DeFanti, T.A. (1993). Surround-screen projection based virtual reality: The design and implementation of the CAVE. *Proceedings of SIGGRAPH'93*, 135-142.

Czernuszenko, M., Pape, D., Sandin, D.J., DeFanti, T.A., Dawe, G.L., & Brown, M. (1997, May). The ImmersaDesk and Infinity Wall projection-based virtual reality displays. *Computer Graphics, 31* (2), 46-49.

Davey, B. Comeau, R., Gabe, C., Olivier, A., & Peters, T. (1994). Interactive stereoscopic image-guided neurosurgery. *SPIE*, 2164, 167-176.

DeFanti, T.A., Sandin, D.J., & Cruz-Neira, C. (1993, October). A 'room' with a 'view'. *IEEE Spectrum*, 30-33.

Depp, S.W., & Howard, W.E. (1993, March). Flat-panel displays. *Scientific American*, 27-38.

Djajadiningrat, J.P., & Gribnau, M.W. (1998, July). Desktop VR using QuickDraw 3D, Part I. Using the View Plane Camera to implement a head-tracked display. *MacTech*.

Djajadiningrat, J.P., Overbeeke, C.J., & Smets, G.J.F. (1997a). Cubby: a medical virtual environment based on multiscreen movement parallax. In K.S. Morgan et al. (Eds.), *Proceedings of Medicine Meets Virtual Reality* (pp. 387-394). Amsterdam: Elsevier.

Djajadiningrat, J.P., Overbeeke, C.J., & Smets, G.J.F. (1997b). The importance of the number of degrees of freedom for rotation of objects. *Behavior and Information Technology, 16* (6), 337-347.

Djajadiningrat, J.P., Smets, G.J.F., & Overbeeke, C.J. (1997c). Cubby: a multiscreen movement parallax display for direct manual manipulation. *Displays*, 17, 191-197.

Dondis, D.A. (1973). *A primer of visual literacy*. Cambridge, M.A.: The MIT Press.

Ehricke, H., Daiber, G., Sonntag, R., Straßer, W., Lochner, M., Schad, L.R., & Lorenz, W.J. (1992). Interactive 3D-graphics workstations in stereotaxy: Clinical requirements, algorithms and solutions. *SPIE Visualization in Biomedical Computing*, 1808, 548-558.

Eichenlaub, J.B. (1990). An autostereoscopic display for use with a personal computer. *SPIE Stereoscopic Displays and Applications*, 1256, 156-163.

Evenblij, M. (1995, December 9). Operatiecomputer moet nog anatomie leren. *De Volkskrant*.

Fisher, S. (1992). Viewpoint dependent imaging: An interactive stereoscopic display. *SPIE Processing and Display of Three-Dimensional Data*, 367, 41-45.

Fitzmaurice, G.W., Zhai, S., & Chignell, M.H. (1993, July). Virtual Reality for Palmtop Computers. *ACM Transactions on Information Systems, 11* (3), 197-218.

Friedländer, U. (1989). Produkt-Semantik und Steinzeit. *Form, Zeitschrift für Gestaltung*, 128, 13-15.

Friedländer, U. (Spring 1984). An historical perspective on the new wave in design. *Innovation. The Journal of the Industrial Designers Society of America*, 12-15.

Gentner, D., & Nielsen, J. (1996). The anti-Mac interface. *Communications of the ACM, 39* (8), 70-82.

Gibson, J.J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J.J. (1986). *The ecological approach to visual perception*. Hillsdale, New Jersey: Lawrence Erlbaum.

Gribnau, M.W., & Djajadiningrat, J.P. (1998, August). Desktop VR using QuickDraw 3D, Part II. Using the Pointing Device Manager to implement a head-tracked display. *MacTech*.

Groen, P. (1988). *Ontwikkeling van het D3D-systeem [Development of the D3D-system]. Manipulatie van ruimtelijke medische beelden [Manipulation of three-dimensional medical images]*. Unpublished master’s thesis, Delft University of Technology, Delft, The Netherlands.

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

Herman, G.T. (1993). 3D Display: A survey from theory to applications. *Computerized Medical Imaging and Graphics, 17* (4/5), 231-242.

Hinckley, K., Pausch, R., Goble, J.C., & Kassell, N. (1994). Passive real-world interface props for neurosurgical visualization. *Proceedings of CHI’94*, 452-458.

Hounsfield, G.N. (1980, October 3). Computed medical imaging. *Science*, 210, 22-28.

Ichinose, S., Tetsutani, N., & Ishibashi, M. (1989). Full-color stereoscopic video pickup and display technique without special glasses. *Proceedings of the SID, 30* (4), 319-323.

Ishii, I., Karasawa, T., & Makino, H. (1994). A method of handling virtual objects that incorporates the sense of touch. *Systems and Computers in Japan, 25* (10), 72-81.

Jones, M., & Wyatt, A. (1994). *3D madness*. Indianapolis: Sams Publishing.

Kameyama, K., & Ohtomi, K. (1993b). A shape modeling system with a volume scanning display and multisensory input device. *Presence, 2* (2), 104-111.

Kameyama, K., Ohtomi, K., & Fukui, Y. (1993a). Interactive volume scanning 3-D display with an optical relay system and multidimensional scaling. *SPIE Stereoscopic Displays and Applications IV*, 1915, 12-20.

Kim, W.S., Tendick, F. Ellis, S.R., & Stark, L.W. (1987, October). A comparison of position and rate control for telemanipulations with consideration of manipulator system dynamics. *IEEE Journal of Robotics and Automation, RA-3* (5), 426-436.

Krippendorff, K., & Butter, R. (1984, Spring). Product semantics: Exploring the symbolic qualities of form. *Innovation. The Journal of the Industrial Designers Society of America*, 4-9.

Krüger, W., & Fröhlich, B. (1994, May). The Responsive Workbench. *Computer Graphics and Applications, 14* (3), 12-15.

Lannoch, H., & Lannoch, H. (1983). Die Form folgt nicht mehr der elektronischen Funktion. *Form, Zeitschrift für Gestaltung*, 104, 28.

Lannoch, H., & Lannoch, H. (1987). Vom geometrischen zum semantischen Raum. *Form, Zeitschrift für Gestaltung*, 118, 12-15.

Lipscomb, J.S. (1989). Experience with stereoscopic display devices and output algorithms. *SPIE Three-Dimensional Visualization and Display Technologies*, 1083, 28-34.

Lee, D.N., & Lishman, J.R. (1977). Visual control of locomotion. *Scandinavian Journal of Psychology*, 18, 224-230.

Mackinlay, J., Card, S.K., & Robertson, G.G. (1990). A semantic analysis of the design space of input devices. *Human-Computer Interaction*, 5, 145-190.

Mann, R.W. (1985). Computer-aided surgery. *RESNA 8th Annual Conference*, 160-162.

Matias, E., MacKenzie, I.S., & Buxton, W. (1994). Half-QWERTY: A one-handed keyboard facilitating skill transfer from QWERTY. *Proceedings of CHI'94*, 88-94.

Mayall, W.H. (1968). *Machines and perception in industrial design*. London: Studio Vista.

McAllister, D.F. (1992, May). 3-D displays. *Byte*, 183-188.

McKenna, M. (1992). Interactive viewpoint control and three-dimensional operations. *Proceedings 1992 Symposium on Interactive 3D graphics*, 53-56.

McKinnon, G.M., & King, M.L. (1988). Manual control of telemanipulators. *Proceedings of the International Symposium on Teleoperation and Control*, 263-276.

Merril, J.R. (1993, November/December). Surgery on the cutting-edge. *Virtual Reality World*, 34-38.

Michaels, C., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, New Jersey: Prentice-Hall.

Molenaar, R. (1997). *Beeldsysteem voor de neurochirurgie [Imaging system for neurosurgery]*. Unpublished master's thesis, Delft University of Technology, Delft, The Netherlands.

Mountford, S.J., Spires, S., & Korner, K. (1986). *Visage: A three-dimensional graphics editor - evaluation and review* (MCC Technical Report #HI-105-86-P). Austin, TX: Microelectronics and Computer Technology Corporation.

Nilsson, J., & Khakhar, D. (1993). Application of living book in medical education. *Proceedings of the INTERACT'93 and CHI'93*, 97-98.

Norman, D.A. (1988). *The psychology of everyday things*. New York: Basic Books.

Norman, D.A. (1992). *Turn signals are the facial expressions of automobiles*. Reading, MA: Addison-Wesley.

Overbeeke, C.J., Smets, G.J.F., & Stratmann, M.H. (1987). Depth on a flat screen II. *Perceptual and Motor Skills*, 65, 120.

Overbeeke, C.J., & Stratmann, M.H. (1988). *Space through movement*. Unpublished doctoral dissertation, Delft University of Technology, Delft, The Netherlands.

Pasman, W. (1997). *Enhancing x-ray baggage inspection by interactive viewpoint selection*. Unpublished doctoral dissertation, Delft University of Technology, Delft, The Netherlands.

Pasman, W., Smets, G.J.F., & Stappers, P.J. (1997a). Effects of image quality, number of selectable viewpoints, and way to select the viewpoint in X-ray luggage inspection. *Presence, 6* (3), 268-281.

Pasman, W., Stappers, P.J., & Smets, G.J.F. (1997b). Distortion in on- and off-axis movement-coupled perspective displays [Abstract]. *Proceedings of the Ninth International Conference on Perception and Action*, 77.

Petropoulos, H., & Brebner, J. (1981). Stereotypes for direction-of-movement of rotary controls associated with linear displays: the effects of scale presence and position, of pointer direction, and distances between the control and the display. *Ergonomics, 24* (2), 143-151.

Raichle, M.E. (1994, April). Visualizing the mind. *Scientific American*, 36-42.

Scheuer, W. (1989). Semantics. *Form, Zeitschrift für Gestaltung*, 128, 12.

Schmandt, C.M. (1983, July). Spatial input/display correspondence in a stereoscopic computer graphic work station. *SIGGRAPH'83 Computer Graphics, 17* (3), 253-261.

Schubert, R., Bomans, M., Höhne, K.H., Pommert, A., Riemer, M., Schiemann, Th., Tiede, U., & Lierse, W. (1993). A new method for representing the human anatomy. *Computerized Medical Imaging and Graphics, 17* (4/5), 243-249.

Shaw, R.E., & McIntyre, M. (1974). Algoristic foundations to cognitive psychology. In W. Weimer & D. Palermo (Eds.), *Cognition and the Symbolic Processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Smets, G.J.F. (1995). Industrial design engineering and the theory of direct perception and action. *Ecological psychology, 7* (4), 329-374.

Smets, G.J.F., Overbeeke, C.J., & Gaver, W. (1994). Form-giving: Expressing the nonobvious. *Proceedings of CHI'94*, 79-84.

Smets, G.J.F., Overbeeke, C.J., & Stratmann, M.H. (1987). Depth on a flat screen. *Perceptual and Motor Skills*, 64, 1023-1034.

Smets, G.J.F., Stratmann, M.H., & Overbeeke, C.J. (1988). Method of causing an observer to get a three-dimensional impression from a two-dimensional representation. *U.S. Patent 4,757,380*.

Spalteholz, W. (1954). *Handatlas und Lehrbuch der Anatomie des Menschen*. Bohn Stafleu van Loghem.

Stoakley, R., Conway, M.J., & Pausch, R. (1995). Virtual Reality on a WIM: Interactive worlds in miniature. *Proceedings of CHI'95*, 265-272.

Sturman, D.J., & Zeltzer, D. (1994). A survey of glove-based input. *IEEE Computer Graphics & Applications, 14* (1), 30-39.

Suetens, P., Vandermeulen, D., Oosterlinck, A., Gybels, J., & Marchal, G. (1988). A 3-D display system with stereoscopic, movement parallax and real-time rotation capabilities. *SPIE Medical Imaging II*, 914, 855-861.

Tendick, F., Jennings, R.W., Tharp, G., & Stark, L. (Winter 1993). Sensing and manipulation problems in endoscopic surgery: experiment, analysis, and observation. *Presence, 2* (1), 66-81.

Thalmann, N.M., & Thalmann, D. (1994). Towards virtual humans in medicine: A prospective view. *Computerized Medical Imaging and Graphics, 18* (2), 97-106.

Van Hattem, E. (1995, July 14). De schedel als industrieel ontwerp [The skull as industrial design]. *Intermediair, 31* (28), 35.

Van Nes, F.L., & Van Itegem, J.P.M. (1990). Hidden functionality: how an advanced car radio is really used. *IPO Annual Progress Report*, 25, 101-112.

Veldhoen, L. (1998, January 3). Moderne mummificatie. Opschudding over nieuwe conververingsmethode [Modern mummification. Commotion around new preservation method]. *NRC Handelsblad*, p.5.

Von Neumann, J. (1966). *Theory of self-reproducing automata*. Urbana, Ill.: University of Illinois Press.

Voorhorst, F.A. (1998). *Affording action, implementing perception-action coupling for endoscopy*. Unpublished doctoral dissertation, Delft University of Technology, Delft, The Netherlands.

Voorhorst, F.A., Overbeeke, C.J., & Smets, G.J.F. (1997). Spatial perception during laparoscopy. In K.S. Morgan et al. (Eds.), *Proceedings of Medicine Meets VR* (pp. 379-386). Amsterdam: Elsevier.

Wade, N.J. (1987). On the late invention of the stereoscope. *Perception*, 16, 785-818.

Ware, C. (1990). Using hand position for virtual object placement. *The Visual Computer*, 6, 245-253.

Ware, C., Arthur, K., & Booth, K.S. (1993). Fish tank virtual reality. *Proceedings of the INTERCHI'93*, 37-42.

Ware, C., & Jessome, D.R. (1988, November). Using the bat: A six-dimensional mouse for object placement. *IEEE Computer Graphics and Applications*, 65-70.

Wheatstone, C. (1838). Contributions to the physiology of vision. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, 128, 371-394.

Wickens, C.D. (1990). Three-dimensional stereoscopic display implementation: Guidelines derived from human visual capabilities. *SPIE Stereoscopic Displays and Applications*, 1256, 2-11.

Worthington, C., Peters, T.M., & Ethier, R. (1985). Stereoscopic digital subtraction angiography in neurological assessment. *American Journal for Neuroradiology*, 6, 802-808.

Zhai, S., & Milgram, P. (1993a). Human performance evaluation of isometric and elastic rate controllers in a 6 DOF tracking task. *Proceedings SPIE Telemanipulator'93*, 2057, 130-141.

Zhai, S., & Milgram, P. (1993b). Human performance evaluation of manipulation schemes in virtual environments. *Proceedings IEEE Virtual Reality Annual International Symposium '93*, 155-161.

Zhai, S., & Milgram, P. (1994). Input techniques for HCI in 3D environments. *Proceedings of CHI'94*, 85-86.

Zubal, G., Harrell, C.R., Smith, E.O., Rattner, Z., Gindi, G., & Hoffer, P.B. (1993). *Computerized three-dimensional segmented human anatomy. Medical Physics, 21* (2), 299-302.

Zuiderveld, K.J. (1995). *Visualization of multimodality medical volume data using object-oriented methods*. Unpublished doctoral dissertation, University of Utrecht, Utrecht, The Netherlands.

Cubby: What you see is where you act

Interlacing the display and manipulation spaces

Summary

In this thesis, a search for an intuitive interface for medical 3D systems is documented. Central in achieving this interface is the use of head-coupled movement parallax. Head-coupled movement parallax gives the observer a 3D impression of a virtual scene on a conventional monitor screen by coupling the parallax shifts on the screen to his head movements. By moving his head, the observer can look around a virtual scene as if it were real. The design philosophy behind the work in this thesis is that in order to make a medical 3D system intuitive it should be less a computer and more a product. The computer needs to be hidden from the user, and a product tailored to the user's natural behaviour needs to surface.

Chapter 1 provides background information on the use of three-dimensional computer models in the medical sciences. Through literature and personal communication, an account is given of how radiologists, generally considered the primary user group of medical 3D systems, view the use of 3D computer graphics. On the basis of this information, it is argued that other physicians, most notably surgeons, are likely to benefit more than radiologists, as the latter are highly trained in making a mental 3D reconstruction out of 2D material. Several areas of application within the medical sciences are discussed, namely visualisation, pre-operative simulation, operative support and education. Two examples of medical procedures are discussed which already benefit from 3D computing, namely stereotactic and craniofacial surgery.

In the second chapter 3D displays are reviewed. The suitability of 3D displays for use in a medical environment is assessed according to two criteria. The first criterion is that a medical 3D system should not hamper the user in his mobility and his communication with others. A distinction is made between desktop and immersive virtual reality (VR) systems, and I argue the advantages of desktop VR for medical applications. Two movement parallax systems, the

Delft Virtual Window System and Fish Tank VR, are compared. The second criterion is that the display method should allow the display and manipulation spaces to be unified, so that virtual objects can be directly manipulated, either by hand or through an instrument. This unification of display and manipulation space could allow the user to manipulate virtual objects with more confidence and higher accuracy.

In accordance with the idea to make a 3D system more product-like, the third chapter is concerned with human-product interaction (HPI). The value of Gibson's theory of affordances for industrial design engineering is argued, with an emphasis on formgiving and interaction. An overview of established 'good practice' is provided to make clear what affordances have to offer in addition to existing practice in HPI. Particular attention is given to the field of product semantics, which I consider to be part of established practice in HPI. Product semantics and affordances are contrasted as they both claim to improve usability through formgiving. While the two approaches stem from completely different theoretical backgrounds - semiotics and Gibson's theory of direct perception respectively - on a practical level they appear very similar. I endeavour to show the difference between the two with an emphasis on the limitations of the use of metaphor and the potential of affordances for inviting action. An example of a videodeck is given to illustrate how an affordance-conscious design approach can differ from existing good practice, and how it can improve HPI. The current trend in HPI to hide components in a 'black box' and to use displays with abstract representations of the product's internal state is criticised.

In chapter 4, a number of design concepts are presented. One of these concepts concerns a hand-held computer, designed by F.A. Voorhorst and myself, based on the Delft Virtual Window System. The pros and cons of a 3D display for a hand-held computer, and of our particular implementation relative both to other hand-coupled and to head-coupled movement parallax systems are discussed. The remaining concepts are for desktop computers based on head-coupled movement parallax. These concepts explore interfaces which make visualizing and cross-sectioning of a virtual body more intuitive than is the case with existing medical work stations.

In one of the interfaces of chapter 4, the user rotates a virtual object by means of an encasing sphere. Chapter 5 documents an experiment in which subjects rotate a transparent, physical sphere encasing an object. There are five conditions which differ with respect to the number of fingers which the subjects are allowed to use for rotation. There is a free condition without restrictions on the number of fingers, conditions with three, two and one finger, and an orthogonally restricted condition. This latter condition corresponds to the decomposed rotation offered in many current interfaces. The conditions correspond to differing numbers of simultaneously available degrees of freedom (DOFs). It is shown that, for quick and intuitive rotation, the number of simultaneous degrees of freedom should be three, which can be realised with two and with three fingers. Types of control which offer less than three rotational DOFs simultaneously fall behind in performance. Orthogonally restricted rotation offered the least performance.

In chapter 6, Cubby is introduced. Cubby is a desktop-sized virtual reality system with three orthogonal screens, forming a cubic space of 200x200x200mm. Through the use of movement parallax on all three screens, the illusion is created that virtual objects stand within the cubic space. Because the virtual objects appear in front of the screens, Cubby makes it possible to unify the display and manipulation space. The use of three orthogonal screens reduces the clipping of virtual objects under observer movements.

When looking at virtual objects in Cubby's early prototypes, observers complained that the virtual objects appeared to deform. In chapter 7, several potential causes of this deformation are investigated. These include flatness cues, static distortion causes and dynamic distortion causes. The most serious cause of distortion is found to be delay. Several changes are made to improve Cubby. These include a more robust set-up to remedy projector misalignment; thick projection sheets to eliminate cockling; the elimination of reflections; faster components; and a mechanical head tracker to remedy delay.

For Cubby to be useful as a surgical simulator, it is important that depth perception is accurate and reliable. Chapter 8 describes an experiment in which subjects have to judge the distance between two virtual objects displayed in Cubby, and between isomorphic cardboard objects suspended in Cubby. There are three conditions: virtual with a mechanical tracker; virtual with a headfree tracker; and a real condition. In these three conditions, subjects view the scene with one eye. In a control experiment, a virtual headfree and a real condition are tested in which subjects view the scene with both eyes. The results show that subjects are limited in their movements by the mechanical tracker. The results also show that, when compared to the monocular real condition, in the virtual headfree condition only the variance of the error is higher, whereas when compared to the binocular real condition, both mean and variance of the error are significantly higher. Thus, performance in Cubby cannot match performance with a binocularly viewed real scene. However, performance in Cubby compares well to a monocularly viewed real scene. The higher variance may be the result of a number of factors such as delay, limited spatial and temporal resolution, and flatness cues. There was less variance in the virtual headfree binocular condition than in the virtual headfree monocular condition, for which I have no satisfactory explanation.

In chapter 9, unification of display and manipulation space is implemented in Cubby. The literature on unified systems is covered. The three main hurdles in unified systems are clipping, occlusion anomalies through the mixture of physical and virtual objects, and distortion of tracking system’s magnetic field caused by the monitor’s magnetic field. It is described how Cubby overcomes or reduces these problems. An experiment is detailed in which subjects have to manipulate virtual objects by means of a hybrid instrument with six degrees of freedom. The hybrid instrument consists of a physical barrel and a virtual tip. It allows for accurate manipulation, as the virtual tip is rendered with the virtual scene, and the delay is cushioned between the tip and the barrel. The experimental task consists of assembling a 3D puzzle. In the four conditions, subjects have to work with either Cubby or a single-screen Fish Tank display, which is either unified or non-unified. The results show that, in Cubby, virtual objects can be manipulated with higher accuracy than in a single screen Fish Tank system. No significant difference in performance between the unified and the non-unified versions of either Cubby or the single screen display could be established experimentally. However, calculation of the mean rank showed that subjects preferred the unified over the non-unified versions in the case of both Cubby and the single-screen display, even though there was no significant difference in performance.

J.P. Djajadiningrat, 1998

Cubby: Wat je ziet is waar je handelt

Het verweven van de weergave- en manipulatieruimte

Samenvatting

In dit proefschrift wordt een zoektocht naar een intuïtieve interface voor medische 3D systemen beschreven. Centraal in het bereiken van deze interface staat het gebruik van hoofdgekoppelde bewegingsparallax. Hoofdgekoppelde bewegingsparallax geeft de waarnemer een 3D indruk van een virtuele scène op een conventioneel monitorscherm door de parallaxveranderingen op het scherm te koppelen aan zijn hoofdbewegingen. Door zijn hoofd te bewegen kan de waarnemer rond een virtuele scène kijken als ware het een reële scène. De ontwerpfilosofie achter het werk in dit proefschrift is, dat een 3D systeem, om intuïtief te zijn, minder een computer en meer een produkt zou moeten gelijken. De computer dient te worden verborgen voor de gebruiker, en een produkt dat is afgestemd op het natuurlijke gedrag van de gebruiker dient naar voren te treden.

Hoofdstuk 1 verschaft achtergrondinformatie over het gebruik van driedimensionale computermodellen in de medische wetenschappen. Middels literatuur en persoonlijke communicatie wordt een overzicht gegeven van hoe radiologen, die algemeen worden beschouwd als de primaire gebruikersgroep van 3D systemen, het gebruik daarvan zien. Op basis van deze informatie wordt beargumenteerd dat andere artsen, met name chirurgen, waarschijnlijk meer baat hebben bij medische 3D systemen dan radiologen, omdat de laatstgenoemden een sterk geoefend voorstellingsvermogen hebben voor het maken van 3D reconstructies uit 2D materiaal. Diverse toepassingsgebieden binnen de medische wetenschappen worden besproken, namelijk visualisatie, preoperatieve simulatie, peroperatieve ondersteuning en onderwijs. Er worden twee voorbeelden besproken van medische handelingen die nu al baat hebben bij 3D computing, namelijk stereotactische en craniofaciale chirurgie.

In het tweede hoofdstuk worden 3D displays besproken. De geschiktheid van 3D displays voor het gebruik in een medische omgeving wordt bepaald aan de hand van twee criteria. Het eerste criterium is dat een medisch 3D systeem de gebruiker niet mag beperken in zijn bewegingsvrijheid of in zijn communicatie met anderen. Er wordt een onderscheid gemaakt tussen desktop en immersive virtual reality (VR) systemen en ik bepleit de voordelen van desktop VR voor medische toepassingen. Twee bewegingsparallaxsystemen, het Delft Virtual Window System en Fish Tank VR, worden vergeleken. Het tweede criterium is dat de methode van beeldweergave het mogelijk zou moeten maken dat de weergave- en de manipulatieruimte worden samengevoegd, zodat virtuele objecten direct kunnen worden gemanipuleerd, hetzij met de hand, hetzij met behulp van een instrument. Deze samenvoeging van de weergave- en manipulatieruimtes stelt de gebruiker in staat om virtuele objecten met meer zekerheid en een grotere nauwkeurigheid te manipuleren.

In overeenstemming met de idee om een 3D systeem meer produktachtig te maken, is het derde hoofdstuk gewijd aan mens-produkt interactie (human-product interaction, HPI). Bediscussieerd wordt welke waarde Gibsons theorie over affordances heeft voor industrieel ontwerpen, met een nadruk op vormgeven en interactie. Er wordt een overzicht gegeven van huidige ‘good practice’, om duidelijk te maken wat affordances aan meerwaarde kunnen toevoegen aan de gangbare werkwijze in HPI. Er wordt extra aandacht geschonken aan produktsemantiek, hetgeen ik beschouw als een onderdeel van de gangbare werkwijze in HPI. Produktsemantiek en affordances worden tegen elkaar afgewogen, aangezien ze beide aanspraak maken de bruikbaarheid van produkten te verbeteren middels vormgeving. Ofschoon de twee benaderingen voortkomen uit geheel verschillende theoretische achtergronden - respectievelijk semiotiek en Gibsons directe perceptie-theorie - lijken ze in de praktijk sterk overeen te komen. Ik poog het verschil tussen de twee te laten zien, met een nadruk op de beperkingen van het gebruik van de metafoor en op de mogelijkheden van affordances om tot handelingen uit te nodigen. Er wordt een voorbeeld gegeven van een videodeck om te illustreren hoe een affordance-bewuste ontwerpbenadering kan verschillen van bestaande ‘good practice’ en hoe deze HPI kan verbeteren. De huidige trend in HPI om componenten te verbergen in een ‘zwarte doos’ en om displays te gebruiken met abstracte weergaves van de inwendige toestand van het produkt worden bekritiseerd.

In hoofdstuk 4 wordt een aantal ontwerpconcepten gepresenteerd. Eén van deze concepten betreft een handheld computer, ontworpen door F.A. Voorhorst en mijzelf, gebaseerd op het Delft Virtual Window System. De voor- en nadelen van een 3D display voor een handheld computer, en van onze specifieke implementatie, ten opzichte van zowel andere handgekoppelde als hoofdgekoppelde bewegingsparallaxsystemen, worden besproken. De resterende concepten zijn voor desktop computers op basis van hoofdgekoppelde bewegingsparallax. Deze concepten verkennen interfaces die de visualisatie en het doorsnijden van een virtueel lichaam intuïtiever maken dan het geval is met bestaande medische werkstations.

In één van de interfaces van hoofdstuk 4 roteert de gebruiker een virtueel object door middel van een omhullende bol. Hoofdstuk 5 beschrijft een experiment, waarin proefpersonen een transparante, fysieke bol roteren, die een object omhult. Er zijn vijf condities die verschillen ten aanzien van het aantal vingers dat de proefpersonen mogen gebruiken voor rotatie. Er is een vrije conditie, zonder beperkingen ten aanzien van het aantal vingers, er zijn condities met drie, twee en één vinger(s) en een orthogonaal beperkte conditie. Deze laatste conditie stemt overeen met de ontlede rotatiemogelijkheid die in veel huidige interfaces wordt aangeboden. De condities komen overeen met verschillende aantallen van gelijktijdig beschikbare vrijheidsgraden (degrees of freedom, DOFs). Getoond wordt dat, voor snelle en intuïtieve rotatie, het aantal gelijktijdige vrijheidsgraden drie zou moeten zijn, hetgeen kan worden bewerkstelligd met twee en met drie vingers. Bedieningsvormen die minder dan drie rotatoire DOFs tegelijkertijd bieden, resulteren in een lagere prestatie. Orthogonaal beperkte rotatie resulteerde in de slechtste prestatie.

In hoofdstuk 6 wordt Cubby geïntroduceerd. Cubby is een desktop-sized virtual reality systeem met drie orthogonale schermen, die een kubusvormige ruimte vormen van 200x200x200 mm. Door middel van het gebruik van bewegingsparallax op alle drie de schermen wordt de indruk gewekt dat virtuele objecten in de kubusvormige ruimte staan. Omdat de virtuele objecten vóór de schermen verschijnen, maakt Cubby het mogelijk om de weergave- en manipulatieruimtes samen te voegen. Het gebruik van drie orthogonale schermen vermindert het afkappen van virtuele objecten tijdens bewegingen van de waarnemer.

Bij het bekijken van virtuele objecten in vroege prototypes van Cubby klaagden waarnemers dat de virtuele objecten vervormden. In hoofdstuk 7 worden verscheidene mogelijke oorzaken van deze vervorming onderzocht. Deze omvatten vlakheidscues, statische vervormingsoorzaken en dynamische vervormingsoorzaken. De belangrijkste oorzaak van vervorming blijkt vertraging te zijn. Verscheidene veranderingen ter verbetering van Cubby worden doorgevoerd. Deze omvatten een meer solide opzet om slechte oplijning van de projectoren te verhelpen; dikke projectieschermen om bobbelen te elimineren; het uitsluiten van reflecties; snellere componenten; en een mechanische head-tracker om vertraging te verminderen.

Om Cubby te maken tot een bruikbare chirurgiesimulator is het van belang dat diepteperceptie nauwkeurig en betrouwbaar is. Hoofdstuk 8 beschrijft een experiment waarin proefpersonen de afstand moeten schatten tussen twee virtuele voorwerpen die in Cubby zijn weergegeven, en tussen isomorfe kartonnen voorwerpen die in Cubby zijn opgehangen. Er zijn drie condities: virtueel met een mechanische tracker; virtueel met een headfree tracker; en een reële conditie. In deze drie condities bekijken proefpersonen de scène met één oog. In een controle-experiment worden een virtuele headfree conditie en een reële conditie getest, waarin proefpersonen de scène bekijken met beide ogen. De resultaten laten zien dat proefpersonen in hun bewegingen worden beperkt door de mechanische tracker. De resultaten laten ook zien dat, in vergelijking tot de monoculaire reële conditie, in de virtuele headfree conditie alleen de variantie van de fout hoger is, terwijl vergeleken met de binoculaire reële conditie zowel het gemiddelde als de variantie van de fout significant hoger zijn. Derhalve kan met Cubby het prestatieniveau met een binoculair bekeken reële scène niet worden gehaald. Het prestatieniveau met Cubby komt echter redelijk overeen met dat van een monoculair bekeken reële scène. De hogere variantie zou het resultaat kunnen zijn van verschillende factoren zoals vertraging, beperkt ruimtelijk en temporeel oplossend vermogen, en vlakheidscues. Er was minder variantie in de virtuele headfree binoculaire conditie dan in de virtuele headfree monoculaire conditie, waarvoor ik geen bevredigende verklaring heb.

In hoofdstuk 9 wordt de samenvoeging van de weergave- en manipulatieruimtes toegepast in Cubby. De literatuur over samengevoegde systemen wordt behandeld. De drie belangrijkste hindernissen bij het realiseren van samengevoegde systemen zijn afkapping, occlusietegenstrijdigheden ten gevolge van het door elkaar lopen van fysieke en virtuele objecten, en vervorming van het magnetische veld van tracking systemen door dat van de monitor. Beschreven wordt hoe Cubby deze problemen elimineert of beperkt. Een experiment wordt beschreven waarin proefpersonen virtuele objecten manipuleren met behulp van een hybride instrument met zes graden van vrijheid. Het hybride instrument bestaat uit een fysieke handgreep en een virtuele punt. Het maakt precieze manipulatie mogelijk, doordat de virtuele punt wordt getekend in de virtuele scène, en de vertraging wordt opgevangen tussen de punt en de handgreep. De experimentele taak bestaat uit het in elkaar zetten van een 3D puzzel. In de vier condities moeten proefpersonen werken met Cubby of een enkelscherms Fish Tank display, die elk samengevoegd danwel niet-samengevoegd zijn. De resultaten laten zien dat virtuele objecten met grotere nauwkeurigheid kunnen worden gemanipuleerd in Cubby dan in een enkelscherms Fish Tank systeem. Noch bij Cubby, noch bij het enkelscherms display kon experimenteel een significant verschil in prestatie worden gevonden tussen de samengevoegde en de niet-samengevoegde versie. Echter, hoewel er geen significant verschil in prestatie was, liet berekening van de gemiddelde volgorde van voorkeur zien, dat proefpersonen de voorkeur gaven aan de samengevoegde versies boven de niet-samengevoegde versies van zowel Cubby als het enkelscherms display.

J.P. Djajadiningrat, 1998

Curriculum Vitae

English

Johan Partomo Djajadiningrat was born in Rotterdam on 2 September 1968. After obtaining his Dutch 'A' levels at the C.S.G. Comenius in Capelle a/d IJssel in 1986, he enrolled on the Industrial Design course at Brunel University of Technology in Egham, England. With a first class BSc(Hons) degree, he continued his studies at a Master of Design course in Industrial Design Engineering, jointly organised by the Royal College of Art and the Imperial College for Science, Technology and Medicine, in London. In 1993 he returned to The Netherlands to take up a post as a research student at the Faculty of Industrial Design Engineering at Delft University of Technology. This thesis is the result of that period.

Nederlands

Johan Partomo Djajadiningrat werd geboren op 2 September 1968 te Rotterdam. In 1986 behaalde hij het VWO diploma aan de C.S.G. Comenius te Capelle a/d IJssel, waarna hij in 1987 startte met de opleiding Industrial Design aan Brunel University of Technology in Egham, Engeland. Na het behalen van zijn BSc diploma met first class honours in 1991, vervolgde hij zijn studie met de Master of Design opleiding in Industrial Design Engineering, verzorgd door het Royal College of Art en het Imperial College for Science, Technology and Medicine, in Londen. Na deze studie begon hij in 1993 als onderzoeker in opleiding aan de Faculteit van het Industrieel Ontwerpen van de Technische Universiteit Delft. Deze periode rondde hij af met dit proefschrift.