# 3D interaction with scientific data : an experimental and perceptual approach

*Citation for published version (APA):*
Qi, W. (2008). *3D interaction with scientific data : an experimental and perceptual approach*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Mathematics and Computer Science]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR635519

*DOI:*
10.6100/IR635519

*Document status and date:*
Published: 01/01/2008

*Document Version:*
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

# 3D Interaction with Scientific Data

## An experimental and perceptual approach

The work in this thesis has been carried out:

- under the auspices of J. F. Schouten School for User-System Interaction Research, Technische Universiteit Eindhoven
- and sponsored by the SenterNovem under Dutch Minister of Economic Affairs



CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Qi, Wen

3D Interaction with Scientific Data / by Wen Qi.
- Eindhoven: Technische Universiteit Eindhoven, 2008.
Proefschrift. -
ISBN 978-90-386-1307-9
Keywords: 3D interaction / scientific visualization / virtual reality / tangible interface / volume rendering / transfer function.

# 3D Interaction with Scientific Data

## An experimental and perceptual approach

**PROEFSCHRIFT**

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op dinsdag 4 november 2008 om 14.00 uur

door

Wen Qi

geboren te Shandong, China

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr.ir. J.B.O.S Martens
en
prof.dr.ir. R. van Liere

# Acknowledgements

I owe my gratitude to all the people who have made this thesis possible and because of whom my PhD experience has been one that I will cherish forever.

First and foremost I would like to thank my promotor and advisor, Professor Jean-Bernard Martens for giving me an invaluable opportunity to work on this challenging and extremely interesting project over the past four years. He has always made himself available for help and advice in every perspective and there has never been an occasion when I have knocked on his door and he has not given me time. It has been a great pleasure to work with and learn from such an extraordinary individual.

I would also like to thank my co-promotor, Professor Robert van Liere. Without his extra-ordinary theoretical expertise and practical ideas, this thesis would have been a distant dream.

I am greatly indebted to Professor Russell Taylor from University of North Carolina at Chapel Hill (UNC). He had the great ideas in visualization and virtual reality. Without him, my trip to the Computer Science department at UNC would not have been so pleasant and fruitful.

I am grateful to Professor Christopher Healey for comprehensive discussions on the value of user studies in visualization research, together with Professor Russell Taylor. His expertise in perception always provided us with a unique perspective.

Thanks are due to Professor Don G. Bouwhuis, Professor Jack van Wijk and Professor Berry Eggen from TU/e, Professor Peter Werkhoven from TNO for agreeing to serve on my doctoral committee and for sparing his invaluable time reviewing the manuscript. I would like to express my gratitude to Professor Mary Whitton and Professor Frederick Brooks as

well for providing invaluable suggestions and help when I was at UNC.

My colleagues at the industrial design department have enriched my PhD life in many ways and deserve a special mention. Dima helped me start-off by introducing the first generation of the Visual Interaction Platform (VIP) in a user-friendly format. Andres Lucero provided help by showing his talent in design. My interaction with other members has been fruitful. I would like to thank all the colleagues who participated in my user studies. I would also like to acknowledge help and support from all staff members. Helen Maas and Nora van de Berg's administrational help was highly appreciated, as was the personnel support from Julma Braat and Jelmer Siben, and the hardware and software assistance from the ServiceDesk.

I owe my deepest thanks to my family, especially my parents, who have always stood by me and guided me through my career, and have pulled me through against impossible odds at times. Words cannot express the gratitude I owe them. Dr. Kathrin Burckhardt gave me lots of encouragement during this PhD, and will remain a good friend to me.

It is impossible to mention all, and I apologize to those I've inadvertently left out.

Lastly, thank you all!

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Scientific and professional activities have led us to a new stage at which huge amounts of data are created every day in different disciplines or domains through instrument measurement and computational simulation. In his review paper, Andries van Dam describes that the size of data sets in scientific research has been growing exponentially [vDLS02]. Understanding these data and reasoning about them poses a big challenge for scientists and professionals although it is essential to lead to more discoveries and push progress forward. Current computer systems are powerful tools that have become an indispensable part of scientific research or professional practice. However, viewing and manipulating data in order to reveal valuable information effectively and efficiently is still not an easy task. The main bottlenecks are the real-time processing and visualization of huge amounts of data and the human ability to understand and interact with these data. There is a pronounced asymmetry between observers and the data they observe, i.e., the bandwidth of information presented to an observer is much higher than the control (s)he has over the data representation. As shown in Figure 1.1, the scientists or professionals who are performing the data analysis need methods or tools to: a) represent the data in an effective form (be it visual, haptic or otherwise), b) interact with this representation in order to optimize it for subsequent analysis (e.g., creating and verifying hypotheses).

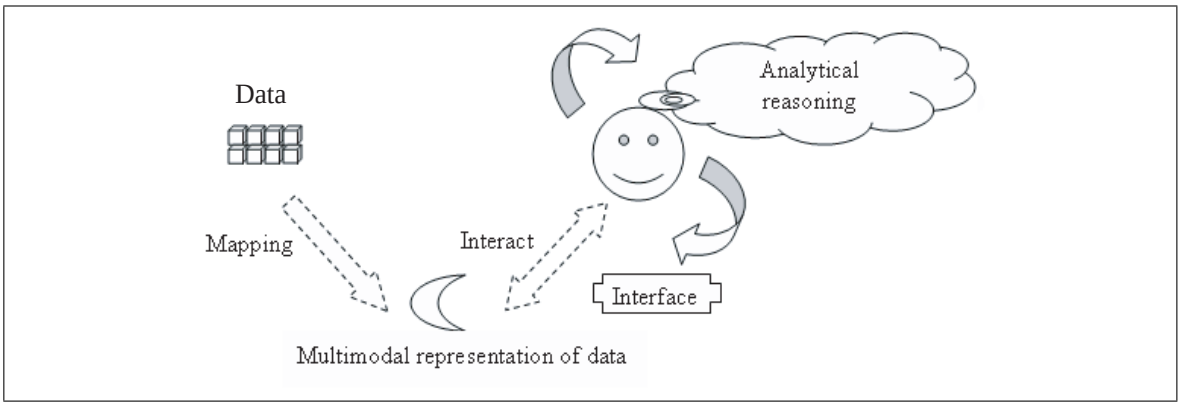



Figure 1.1: Interaction with scientific data including the analytical process.

## 1.1 What is 3D Interaction with Scientific Data

3D interaction with scientific data addresses several aspects, such as setting parameters that influence the mapping from data to image on the screen, performing manipulations in 3D space in order to reposition 3D data/objects or viewpoints and performing analysis, such as comparing or locating objects in a data set. This exploration enables the user to interact with the data to understand trends and anomalies, isolate and recognize information as appropriate, and engage in the analysis (analytical reasoning) process.

In this thesis, the athor classifies 3D interaction with scientific data based on the purpose of the interaction with data ([CM83] and [TC05]):

1. Interaction required for controlling visual mapping.
2. Interaction required for the modification of view transformation.
3. Interaction required for measuring data/object properties.

These three categories deal with different issues at different stages of 3D interaction. Interaction for visual mapping is concerned with mapping raw data into a multimodal representation. Although the visual representation is usually dominant, it may be complemented with other modalities, such as touch or sound. The transfer function (TF) specification discussed in the second chapter is an example of this kind of interaction. It uses a graphical user interface (GUI) to control the visual mapping such that structures of interest are rendered

more prominently than others. Interactions for modifying the view transformation allow users to manipulate or navigate the representation of a data set. The Virtual Reality (VR) systems presented in the third chapter provide alternative means for modifying the view transformation. Different VR systems offer different strategies of navigation/manipulation during a data analysis process. Especially the aspect of inside-out versus outside-in viewing of the data will prove to be important. The first two interactions can be regarded as generic interactions. The last type of interaction is usually task-specific, and in the fourth chapter the author studies an important example of such a specific interaction, i.e., creating intersection images. Such a specific interaction may for instance assist in the task of comparing the sizes of 3D objects. These three kinds of interactions together constitute a complete loop of 3D interaction with scientific data.

3D interaction can be intuitive and 3D interaction with scientific data can bring many benefits for scientists and professionals. The basic motivation for using 3D interaction is that human beings live and interact in a 3D space that is filled with 3D objects. Human beings develop physiological structures and practical skills to enable 3D interaction. For example, the anatomical structure of the human eyes enables stereoscopic vision. With stereoscopic vision, an observer can deduce depth from object disparities in both eyes. The added perception of depth makes stereoscopic vision rich and special. With stereo vision, an observer can understand where objects are in relation to his/her own body with greater precision, especially when those objects are moving towards or away from the observer. The benefits of stereoscopic displays in 3D interaction have also been established experimentally, for example in [WF96] and [WG98]. In addition, Marr and Biederman's 3D object perception theories indicate that if objects are represented in 3D forms, these objects will be easier to identify and memorise than 2D forms. Also, data structures will be better understood if they are mapped to object structure. The reason is that the human visual system can extract the object structure (and hence the data structure) using available perception mechanism [Bie87]. These arguments provide support for the position that scientific data analysis can profit from a representation and an interaction in 3D.

## 1.2 Tasks in 3D Interaction with Scientific Data

The ultimate goal for a scientist or professional is to discover valuable effects within a data set and to explore scientific or professional meanings. These general goals are pursued through single or multiple interaction loops in which specific subtasks are performed. Wickets et al summarized the prior work on 2D versus 3D interactions as "whether the benefits of 3D displays outweigh their costs turns out to be a complex issue, depending upon the particular 3D rendering chosen, **the nature of a task**, and the structure of the information to be displayed" [COAC97]. Therefore, understanding tasks should be emphasized while studying 3D interaction. In general, tasks discussed in this thesis can be categorized as supportive tasks or analysis tasks.

**Supportive task** Supportive (generic) tasks are those that assist a user to pursue further data analysis in an effective way. According to the author's classification, supportive tasks refer to those tasks that control the visual mapping and modify the view transformation. Specifying a TF is an example of controlling the visual mapping. View transformation tasks include manipulation (rotation, translation, zooming), selection, navigation, and etc.

Multiple individual tasks can be combined together to form a compound task, for instance, navigation. A compound task can be crucial to help a user pursue his analysis task smoothly, with the intention of making valuable measurements or drawing credible conclusions. For example, an angle measuring the angle of three adjacent atoms in a molecule structure can be done by performing several compound tasks in sequences (rotation, zooming, and selection, and etc).

**Analysis task** The goal of data analysis is to make judgements about a data set based on visual or other representations. In scientific data analysis, the analyst is usually a researcher or professional who typically adopts one or more analysis goals during the course of visual (or other forms of) exploration of the scientific data set. Several attempts may be undertaken to reach these goals.

Wehrend [WL90] comprehensively reviewed over 300 visual displays and produced a list of analysis tasks:

**identify** "To identify something is to establish the collective aspect of the characteristics by which it is distinctly recognizable or known".

**locate** "To locate something is to determine its specific position".

**distinguish** "To distinguish a thing is to recognize it as different or distinct from other things".

**categorize** "To categorize things is to place them in specifically defined divisions in a classification".

**cluster** "To cluster thing is to join them into groups of the same, or related type".

**rank** "To rank something is to assign it a particular order or position with respect to other things of similar type".

**compare** "To perform comparison between things is to examine them so as to note their likenesses and differences".

**associate** "To study and build up association is to link or join things in a relationship".

**correlate** "To correlate things is to establish a direct connection between them".

Using Wehrend's list to categorize a task can help an interaction/interface designer to understand its characteristic. The author can, for example, analyze the research questions studied in this thesis and summarize the tasks involved, as shown in Table 1.1. In this table it is specified what kinds of tasks and actions are required while a user pursues his goals.

Beddow [BB92] and Robertson [Rob90] categorized data analysis tasks within the field of scientific visualization using the following different characteristics:

- global level: implying the entire data set,
- group level: implying a subset of non-adjacent points,
- local level: implying corrected subregions in the data,
- point level: restricted to the data at a particular location in the data space.

The emphasis of different levels that are involved in a task are important because it indicates which kinds of features an interface should own. For instance, if a task is

Table 1.1: Tasks studied in this thesis, characterized using Wehrend's classification of analysis tasks

| **User goal** | **Tasks** |
|---|---|
| TF specification | **Identify** the correct parameters (the mapping between opacity and data value), **locate** the required object and **compare** the structure rendered with the target structure. |
| How many differently shaped objects are within a volume | **Identify** all existing shapes and **categorize** them. |
| How many differently sized spherical objects in a volume | **Identify** all existing sizes and **categorize** them. |
| Which region is the densest | **Compare** the densities of all regions, **distinguish** whether or not there are any differences and **rank** them. |
| How many curved tube in a volume | **Identify** the curved tubes and **distinguish** whether or not they are the same. |
| Where is the longest curved tube | **Compare** the lengths of different curved tubes, **rank** them and **locate** the longest one. |

performed at the global level, an interface should provide an overview functionality in order to provide access to the entire data set.

Casner gave another kind of classification in which analysis tasks are divided into two types: search and computation [Cas91]. For example, according to his classification, finding out whether or not there are any curved tubes in a volume is a search task. Computation tasks are regarded as those involving measurements/comparisons of objects in a data set. The measurement can be absolute, for example, measuring the coordinate difference between point $a$ and point $b$. It can be relative as well, such as comparing whether a selected point $a$ is closer to one point $b$ than to another $c$. Haimes and Darmofil described their understanding of user goals as belonging to three

categories: scanning through a complete data set, identifying features within regions of the data set, and probing at particular locations [HD91].

Each of these classifications highlights certain specific characteristics of a task from a particular point of view. Therefore, it is advisable to combine them while analyzing and characterizing a task. For example, judging whether or not there is a specific object shape in a volume can be regarded as a global level task since it is necessary to scan the whole volume according to Beddow and Robertson. At the same time, according to Wehrend, it is an identification task. Finding out which region is the densest is a kind of computation task since users need to judge and compare the densities of all regions. Again, it is is very important to differentiate and identify the category and property of a task in the sense that it can help to understand the requirements or demands of the task on both the human user and the system. Understanding these requirements then can assist in selecting the appropriate interaction devices or techniques.

## 1.3 Current Status of 3D Interaction with Scientific Data

Successful 3D interaction with scientific data requires the advance of both visualization and human computer interaction (HCI) research. Scientific visualization investigates possible methods to translate data into a 3D visible form that highlights important features, including commonalities and anomalies. At the same time, research activities that represent data with other sensory modalities, such as touch or hearing, are also emerging. A detailed discussion on the achieved progress in scientific visualization can be found in appendix A.

Progress in a single aspect, for example in visualization (modeling) techniques or user interfaces, does not by itself guarantee that users will be able to work with scientific data more efficiently and effectively. In his review paper,"Top Scientific Visualization Research Problems", Chris Johnson pointed out that one of the ten problems in scientific visualization research is HCI [Joh04]. HCI research has become more and more important for better data analysis. Therefore, the focus in this thesis is also on interaction issues, instead of on visualization issues.

Previous research in 3D interaction has addressed a wide variety of topics. These topics

include the design of novel 3D input or display devices (for example [FP00], [Sut68]), the experimental study of universal task performance with various input devices or interaction techniques (for example [PBWI96] and [FHSH06]), and the study of adding various tangible aids to devices (such as physical props [HPG94]). The design of interaction techniques and devices for supportive (manipulation) tasks has been the most central research topic. Their counterparts in traditional 2D interaction include devices such as the mouse and techniques such as the scrollbar for navigation, selection technique by point-and-click, drag-and-drop technique for manipulation [BCWea06], and etc.

Summarizing the results from those literatures, the author concludes that 3D interaction and user interfaces are not uniformly successful. There are contradictive evidences as to whether or not 3D interaction actually transforms into better efficiency and satisfaction, despite of obvious progress in each of the relevant subfields [BCWea06]. **The outstanding problem with 3D interaction for scientific data analysis is that, despite the broad investigation and extensive knowledge on 3D interaction devices and techniques, the usability of this approach in real-world applications still needs to be established.**

An imporant reason for the current status seems to be that previous studies on 3D interfaces and interaction techniques have been largely technology-driven [BCWea06] and that the tasks being studied have been mostly supportive tasks (such as travel, selection and manipulation). These generic 3D interaction tasks mainly relate to interactions for modifying the view transformation. They are essential building blocks for 3D interaction, but are far from complete. **At least two other important aspects are missing: 1) the interactions required for controlling the visual mapping and 2) the potential effects of 3D interaction techniques and interfaces on practical data analysis tasks.** As a result, knowledge from these available studies only partially contributes to improving the usability of 3D interaction in data analysis.

## 1.4 Research Topics in This Thesis

The discussion in the previous sections leads to the conclusion that there are two major problems with current understanding of 3D interaction research. First, there are very few

experimental studies that investigate user interfaces and interactions for controlling the visual mapping. Second, there are lots of studies that try to design or evaluate interaction devices and techniques for modifing the view transformations in generic 3D interaction tasks, such as travel and navigation, but few studies with specific data analysis tasks.

In this thesis, the author studies alternative interfaces for controlling the visual mapping. More specifically, the goal is to determine whether or not the proposed interfaces can lead to a successful rendering, i.e., one that supports the further analysis of the data set. Another goal is to design and test an experimental method for measuring performance in a TF specification task.

It should be noted that the author does not intend to design new types of user interfaces and interaction techniques, but focuses on investigating available interface solutions. VR and tangible user interfaces as two types of user interfaces that are of great interest today. The author is interested in studying the effects that these interfaces choices can have on different data analysis tasks, instead of concentrating on the effect on traditional navigation and manipulation tasks. In other words, the author questions whether or not these interfaces and interaction techniques, which were originally designed to better support the user in making modifications to the view transformation, can also support the user when performing data analysis tasks. The effects on modifying view transformations are not in the focus of attention since lots of research has already been done in this area.

Hence, the author formulates three individual research questions in this thesis:

1. **What are the usability issues with current user interfaces for TF specification ( in particular with the most frequently used method of trial-and-error)?**

2. **What are the performance differences between available VR systems when analyzing object properties within a volumetric data set, such as size, shape, density and connectivity?**

3. **What are the potential effects of tangible user interfaces on analyzing object properties within a volumetric data set? In particular, can using tangible objects for controlling a clipping plane operation provide help for data analysis tasks in 3D space?**

### 1.4.1 Research Topic 1: Transfer Function Specification

The first research question is about **how different elements in a dedicated graphical user interface (GUI) affect the efficiency of the interaction while specifying a TF using the trial-and-error method**. The TF that the author will study relates data values (density) to transparency and controls the visual mapping from the raw 3D volumetric data into the 2D visual representation. More concisely, this process will be referred to as "TF specification in direct volume rendering".

TFs are crucial for controlling the visual mapping in direct volume rendering. Most users of volume rendering are domain scientists/professionals who excel in domain knowledge, but who have very limited knowledge about TFs. The immature characteristic of this research area is that, although diverse user interface paradigms (for example trial-and-error, Design Gallery [MABea97]) have been proposed, there are very few experimental studies so far that provide concrete quantitative evidence about user performances with these methods. The author focused on the trial-and-error method because it is also the most widely used method today. The philosophy behind the method is to put complete control over the TF in the hands of the user. The study in chapter two adopts the trial-and-error method as the basic scheme and investigates whether or not data-dependent (histogram) information, data-independent (pre-defined TFs) information and limiting the degrees of freedom (DOF) of a TF, are useful additions to it. The user performance with the different interface alternatives are compared in a controlled experiment. Important usability issues in the specification process are identified partly through an analysis of the TF specification task. It is obviously only a first step in providing more experimental evidence from the experimental results as to what are the main user interface problems in TF specification. The author uses the expertise acquired in this study as a starting point to advocate that researchers need to pay more attention to these kinds of interactions that aim at controlling the visual mapping. Despite the fact that only the trial-and-error method is studied, it provides constructive guidelines for other researchers and designers on how to approach this problem in an experimental way.

### 1.4.2 Research Topic 2: Usability of VR systems

The second research question is about **the effects of different VR interfaces (systems) on selected data analysis tasks**. Previous VR research covers many different topics, from the development of technologies to studying related perceptual and cognitive issues (for example, presence and immersion). The focus is however mainly on creating new interaction styles and techniques and on developing domain applications (see table 1.2). In the research and development of new technologies and interaction techniques, researchers often concentrate on simple (generic) 3D interaction tasks, multimodality, etc. With respect to the perceptual and cognitive issues in VR system, there are frequent couplings with relevant research in psychology, regarding the effect of 3D interfaces on spatial reasoning and memory (such as [WP82] and [RS90]). For example, Peruch et al. tested the capability of an observer to learn spatial layouts of objects located in a wall-limited virtual space [PVG95]. The results indicated that spatial acquisition after active exploration was more accurate than after passive exploration, and that dynamic and static (passive) visual information yielded equivalent performance. Other studies investigated the presence and immersion aspects that are unique for VR interfaces (for example [PPW97], [MIWB02], [RW01]). Still other research projects promote the use of VR interfaces in specific domain applications, such as medical diagnosis, psychiatric treatment, flight simulation, entertainment and data visualisation [Bro99]. Data visualisation and analysis using VR systems is an important application domain that the athor focuses on here.

VR has been actively used as a tool for visualising and analysing scientific data. However, the decision of selecting a specific set-up is often based on the designer's subjective preference and available resources (see the literature in Chapter 3). Instead, it should be based on the understanding of the relationship between an inteded task and the properties of a proposed interface, i.e., on an informed estimate of the combined effects of different navigation and manipulation techniques with different display strategies. If the decision for a specific set-up can not be verified, it may well not prove to be suitable for the intended purpose. There are a few studies available regarding to the effects of different VR set-ups on generic interaction tasks, such as navigation and manipulation. For example, Werkhoven and Groen studied manipulation performance in a virtual environment using two types of interaction techniques: virtual hand and 3D mouse under both monoscopic and stereoscopic viewing

conditions [WG98]. There are no other studies that the author knows of that investigate the overall effect of different VR set-ups on data analysis tasks. Hence, there is very little knowledge on how an integrated VR set-up with visualisation capability helps or hinders the data analysis process. In other words, the advantages and disadvantages that a VR system has on performing specific data analysis tasks is not very clear currently.

Table 1.2: Research topics with VR systems

| **Development of technologies** | **VR related perceptual and cognitive issues** | **Interaction styles and techniques** | **Development of VR applications** |
|---|---|---|---|
| Hardware (Input and Output) | Spatial reasoning and memory | Navigation | Medicine (Therapy) |
| Software (Toolkit) | Presence | Manipulation and selection | Data visualization |
| Haptic and Tactile | Immersion | Multimodal interaction | ... |
| Auditory | Simulation sickness | | |

Three common VR set-ups (HMD based immersive VR, fish tank VR and fish tank VR with haptic feedback) were designed and implemented in order to carry out a user study aimed at investigating user performance in four generic but important 3D visualization (analysis) tasks. These tasks included judging the shape, size, density and connectivity of a priori specified object within a volume. They are derived and generalised from the research questions posed by domain specialists who study Cystic Fibrosis (CF). The study questions the effects of immersion and presence on those data analysis tasks within a HMD based VR system. The study also measures the effect of haptic force feedback on the same tasks within a fish tank VR system. The study does not test other potentially useful aspects, such as the possible effect of the auditory modality, partly because this modality is less commonly used for data analysis, and partly because resources are limited in term of experiment possibilities.

### 1.4.3 Research Topic 3: Tangible User Interfaces

**Whether or not tangible user interfaces, (i.e., physical objects as controls and representations within 3D manipulations) are useful for data analysis tasks** is

the third research question. In particular, the author studies whether or not the inclusion of a clipping plane, possibly controlled by a physical object, can assist in performing the data analysis tasks mentioned before.

Designing input devices with 6 (or more) DOFs is an active area of research within 3D interaction, despite that fact that only very limited knowledge is available on which properties a good 6 DOF device should have. The more general body of knowledge on human motor control and learning (see [SL98], for example) hardly provides useful design guidelines, although it offers valuable insights. Involving tangible user interfaces while interacting with scientific data seems a priori to be a promising approach. The rationale behind this is that when human beings interact with everyday objects in the real world, they do not consciously apply complex thought in order to manipulate or use them. Their "behavior" is inferred from their properties: shape, weight, size, etc. The functionality is also expressed through the object's physical form, i.e., the object has "affordances" [Nor93]. Seichter and Kvan introduced the concept of "augmented affordance" to indicate that tangible user interfaces can be seen as "offering a conduit between the real or perceived affordances implied by the physical properties of the interface and the affordances created by the digital behaviours in the virtualised interface" [SK04]. As proposed by Colin Ware, such coupling of input and output should also be achieved in interactive visualization for data analysis [WF96].

So far, several successful 3D tangible devices exist (the Cubic Mouse (CMouse) [FP00], ActiveCube [KIK01] and the Passive Interface Props (PassProps) [HPG94]) and their positive effects on generic tasks (modifying the view transformation) are partially confirmed (mainly through qualitative observations). For example, from detailed observations of user behaviors in 3D rotation tasks, Hinckley concluded that the physical form factor of a 3D input device significantly influenced user acceptance of identical input sensors. He indicates that if a device for rotation affords tactile cues, the user can feel its orientation without looking at it. In the absence of such cues, some users may be unsure of how to use the device [HTP+97]. However, those qualitative observations are not convincing enough to prove that tangible user interfaces can really support 3D manipulation tasks, let alone more complex data analysis tasks.

In our study, the potential of improving spatial reasoning in data analysis tasks is extensively explored. The data analysis (visualization) tasks are the same as the ones in the

previous VR study. The user performances with different tangible interaction devices (physical objects with specific shape) for controlling a clipping plane function on a 3D desktop VR environment are compared. Moreover, the study verifies whether or not these tangible interfaces have positive effects on modifying the view transformation.

## 1.5 Outline of The Thesis

The thesis consists of five chapters, which document the different steps taken during the research.

Chapter 1 has provided a brief introduction to relevant concepts, and has discussed the potential advantages of 3D interaction with scientific data. The tasks that are involved in 3D interaction with scientific data are classified within two categories: supportive tasks and analysis tasks. This classification is used to position and motivate the specific questions addressed within this thesis. The research questions are chosen in order to reflect different relevant aspects (TF specification, VR and tangible user interfaces). This chapter provides the basis for understanding the motivation for the specific user studies presented in the rest of the thesis.

In Chapter 2, empirical work is presented regarding usability issues of a GUI for TF specification in direct volume rendering. Various specification methods are discussed at the beginning of the chapter. With an emphasis on the trial-and-error method, the user experiment describes user performances and preferences for alternative interface choices.

Chapter 3 reviews the current research in VR and its applications. The value of VR for scientific visualization is discussed. A comprehensive experimental study is conducted to compare the user performance of three different VR set-ups for four specific data analysis tasks performed with visualizations of simulated data. The research problems are inspired by tasks that are considered to be important for domain researchers who study CF.

In Chapter 4, tangible user interfaces for scientific visualization and two-handed interaction are discussed based on state-of-the-art research. User performances on the same analysis tasks as in chapter 3 are investigated through an extensive user study with a focus on tangible interfaces for 3D manipulation, particularly for 3D clipping plane manipulation.

The design process of the physical objects involved is described as well.

Chapter 5 is the epilogue of this thesis. In this chapter, insights gained and lessons learned from the work in previous chapters are discussed. Design guidelines derived from the studies are proposed. Possible future research topics are identified, both within the context of scientific data analysis addressed in this thesis and within the broader area of 3D interaction.

# Chapter 2

# Graphical User Interfaces for Transfer Function Specification

Visualization via direct volume rendering is a powerful technique for exploring and manipulating large scientific data sets [BCE+92]. One problem that hinders effective use of it is the difficulty of understanding and specifying the correct transfer function (TF) for a specific data set, especially for non-expert users. The TF in a direct volume rendering system assigns optical properties, such as color and transparency, to the data values during the visualizing process. An appropriate TF can make a vast difference in quality and content of the rendered image. However, it is difficult to derive such a function automatically or manually as it is much dependent on the semantics of a specific data set. This chapter introduces important usability issues in TF specification, and analyzes the proposals that have been made in the literature to improve and optimize this interactive process. It summarizes the advantages and disadvantages of the current approaches in TF specification, and describes our visualization system prototype. Using this prototype, an experimental set-up has been realized to investigate the trial-and-error method. The author discusses the results of the usability test of a trial-and-error interface with varying additional information. The author draws conclusions about technical and psychological aspects of the experiment, and describe the lessons learned from this study for future interface design.

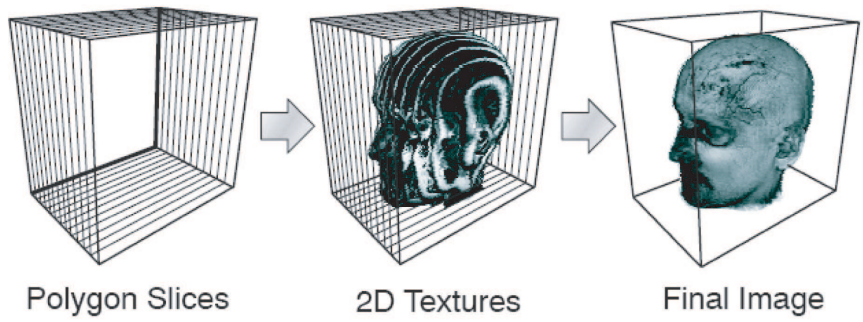



Figure 2.1: Volume rendering via 2D texture mapping (Photograph reprinted from [RS01]).

## 2.1 Hardware-accelerated Direct Volume Rendering

Due to the huge amount of data involved in 3D rendering, creating a visual representation of volumetric data often relies on hardware to improve rendering speed and decrease computational expense. There are two approaches while using hardware acceleration: customized hardware or general-purpose hardware. These different approaches lead to different methods to implement the TF specification. In this study, texture mapping with general-purpose hardware was selected as the rendering method. In the following section, both approaches are discussed briefly.

1. Customized hardware acceleration
   Researchers from the State University of New York at Stony Brook have designed and pioneered a series of hardware architecture prototypes, called *Cube-X*. The first generation Cube-1, was designed with a specially interleaved memory organization [KB88], which was also used in all following generations of the Cube architecture. The interleaving of the $n^3$ voxels makes conflict-free access to any ray parallel to a main axis of n voxels possible. A fully operational printed circuit board (PCB) implementation of Cube-1 can generate orthographic projections of $16^3$ data sets from a finite number of predetermined directions in realtime (30 frames per second). Several improvements have been made in the following series. For example, the second generation was a

single-chip Very Large-Scale Integration (VLSI) implementation of the first-generation prototype [BKPP92]. The third generation further reduced the critical memory access bottleneck to reach an estimated performance of 30 frames per second for data sets with the size of $512^3$. The fourth generation Cube-4 manipulates a group of rays at a time, rather than processing individual rays. It is easily scalable to very high resolution like $1024^3$ 16-bit voxels with true real-time performance implementations of 30 frames per second. Mitsubishi Electric has derived another system called EM-Cube (Enhanced Memory Cube-4). A system based on EM-Cube consists of a PCI card with four volume rendering chips, four 64Mbit SDRAMs to hold the volume data, and four SRAMs to capture the rendered image [OPL+97].

2. Texture mapping with general-purpose hardware
   Another approach for hardware-accelerated volume rendering utilizes texture memory on general-purpose graphics cards, and is called texture mapping. Texture mapping is an object space technique, since all calculations are done in object space. This means that the rendering is accomplished by projecting each element onto the viewing plane so as to approximate the visual stimulus of viewing the element based on the chosen optical model. The rendering speed of this approach depends only on image size instead of scene complexity, and geometric models are not required. After being loaded into texture memory, a data set is sampled, classified, rendered to proxy geometry, and composited. Classification typically occurs in hardware by means of a look-up table.

   Normally there are two ways to perform texture mapping: 2D texture mapping and 3D texture mapping.Volume rendering based on 2D textures is quite straightforward (Figure 2.1). As seen in Figure 2.2, 2D texture mapping interpolates two texture coordinates (s, t) across a polygon's interior. The pseudo code is like:

   o Render each xz slice in the volume as a texture-mapped polygon;
   o The texture contains color and opacity;
   o The polygons are drawn from back to front.

   The detailed algorithm description of 2D texture mapping is as follows:
   Turn off the z-buffer and enable blending,

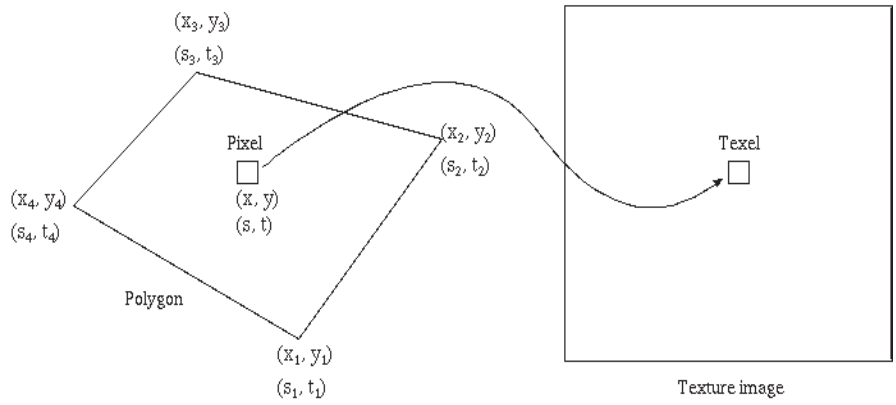



Figure 2.2: The texture coordinate transformation in 2D texture mapping (Photograph reprinted from [Kre00]).

For (each slice from back to front)
- Load the 2D slice of data into texture memory;
- Create a polygon corresponding to the slice;
- Assign texture coordinates to four corners of the polygon (Figure 2.2);
- Render and composite the polygon (use OpenGL alpha blending.)

However, there are several problems with 2D texture mapping [MB97]. Firstly, the difficulty with 2D textures is that the data slice polygons can't always be perpendicular to the view direction. Three sets of 2D texture maps must be created, with each set perpendicular to one of the major axes of the data set. Adjacent 2D slices of the original 3D volume data along a major axis are used to create these texture sets. The data slice polygons must be aligned with whichever set of 2D texture maps that is most parallel to them. The data slices can be slanted 45 degrees away from the view direction in the worst case. As the slices are more edge-on to the eye, the data sampling becomes worse. The extreme case for an edge-on slice is that the textured values on the slices aren't blended at all. At each edge pixel, all the other values are obscured except the sample that is from the line of texel values crossing the polygon slice. Secondly, the speed of rendering dramatically slows down when 2D texture mapping is used to render large data sets.

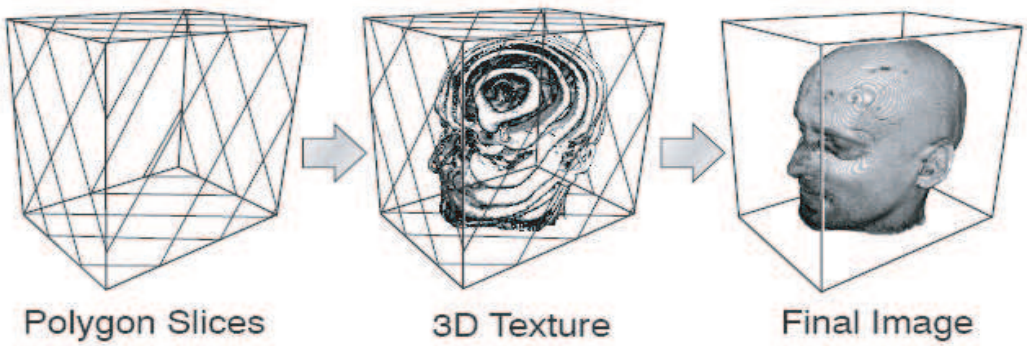



Figure 2.3: Volume rendering via 3D texture mapping (Photograph reprinted from [RS01]).

3D texture mapping has been developed to allow interactive generation of view-orthogonal slices, with a special hardware technique (Figure 2.3). In 3D texture mapping, three texture coordinates (s, t, r) are interpolated. For the calculation of a pixel's color and opacity, these three coordinates are used as indices into a 3D image, the 3D texture, as Figure 2.4 shows. Trilinear interpolation is the most frequently used method to reconstruct texture values. 3D textures enable direct treatment of volumetric data and hence avoid the generation of a set of 2D slices in a pre-processing step. The volumetric data set is loaded into the rendering hardware directly, and then used to determine color and opacity values for each pixel, which is covered by a rendered primitive. 3D texture-based volume rendering has the following advantages:

- Speed: Because available graphics hardware is optimized for texture mapping, this technique allows for interactive frame rates even on commodity graphics boards found in today's game market.
- Versatility: Due to its high rendering speed, 3D texture-based volume rendering can be used in many interactive applications, like radiology image pre-viewing and VR applications with direct volume rendering.

However, 3D textures mapping is not supported by all graphics cards. Different graphics card manufacturers have developed their own Application Programming Interfaces

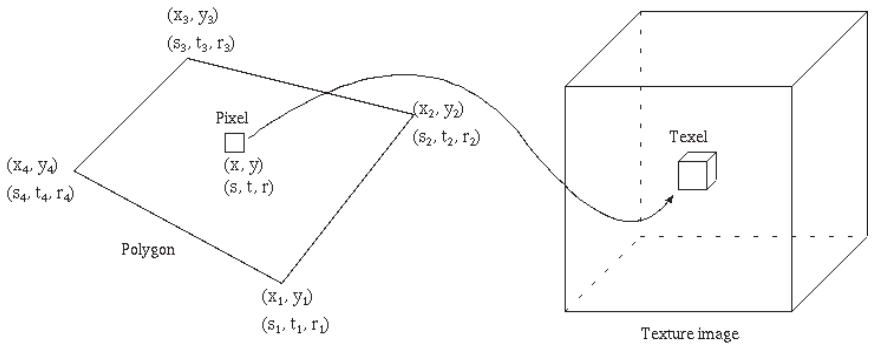



Figure 2.4: The texture coordinate transformation in 3D texture mapping (Photograph reprinted from [Kre00]).

(APIs) and implementation for 3D texture mapping, which causes compatibility problems.

## 2.2 Task Analysis

The TF is a critical component in the direct volume rendering process. It specifies the relation between scalar data (e.g. densities measured by CT or MRI imaging), and possibly also their first- or second-order derivative values, and optical characteristics, such as color and opacity [LCN98]. As discussed in the previous section, current graphics hardware-based algorithms provide the possibility to continually modify the TF so that the results of direct volume rendering can be updated in real time. There are several steps involved in this TF specification (Figure 2.5). Ideally, a user can hope that a system provides sufficient information in the initial stage to finish the specification in a single step, as is indicated by the dashed arrow in Figure 2.5. However, users usually need to go through multiple iterations of exploration and refinement before arriving at the final specification. During the initialization, a user is offered several inputs, such as derived data properties, like grey-value and/or gradient histograms, one or more initial TFs with correspondingly rendered images, etc. The user can explore the presented information and TF alternatives through a graphical or numerical user interface. He can assess the results of his operations based on the provided visual feedback. This visual feedback may not be restricted to the result

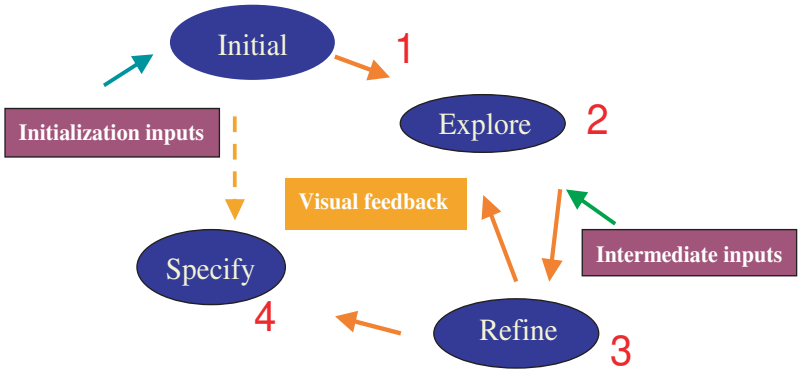



Figure 2.5: The iterative process of TF specification.

of the last operation, but may also include feedback of preceding iterations and/or of the initialization step. The user refines his previous actions until he reaches his final goal, i.e., obtains a transfer function that results in a rendered image that adequately portrays the structure(s) of interest.

The initial information that is presented by the system can consist of the following:

1. Data-dependent information such as histograms of grey or color values or (first- and second-order) derivatives of the input data, or a TF that is derived from the data through some sort of optimization algorithm;

2. Data-independent information that is based on prior knowledge or experience, such as standard or advised TFs (in medical applications, for instance, the TF may be determined by the kind of examination).

The intermediate feedback, in turn, can include the following:

1. Information from the initialization stage;

2. Visual feedback from the last operation of the user;

3. Feedbacks from one or more previous operations, that can assist in assessing the progress, without having to rely on memory.

## 2.3 Related Work

The TF specification in volumetric visualization is a fairly unique and complex interaction compared to the elementary interactions, such as selection and positioning, that occur in most 3D graphics applications. It is only recently that this interactive process has become feasible in real time, since it relies on the use of hardware graphics accelerators. Several alternative proposals have been made for answering the question of how this interaction can be performed best. They range from completely manual to completely automated, and differ in the amount and kind of feedback that is provided (see [HHKP96], [MABea97], [KD98], [BP02], [KG01], [Ma99], [JKM00], [TL03], [RSKK06]).

The most common method is the trial-and-error method. It involves manually editing the TF by modifying a graphical curve and/or by adjusting numerical parameters, and visually inspecting the resulting image (Figure 2.6 left) [PLB+01]. This method is primitive and problematic because it requires the users to go through all specification steps without intermediate feedback. Even with high-end facilities, this method can be very inefficient and time-consuming, because of the complexity of understanding the non-trivial relationship between a TF and the correspondingly rendered image. It also requires a reasonably accurate understanding of the visualization process by the user. However it is still the dominant method because it puts the user in control.

A method that tries to avoid the reliance on the user's visualization expertise is the Design Gallery approach [MABea97] (Figure 2.6 right). It involves creating and displaying a large number (hundreds) of rendered images that correspond to a range of predefined TFs. Design Gallery is an example of an image-centric method. Ma's image graph [Ma99] and Kelly's spreadsheet [JKM00] are related techniques. The image-centric methods do not focus on how to assist the user in finding a good TF by providing adequate feedbacks on relevant data-set properties, but instead focus on the design of the user interface. In the Design Gallery, all the user has to do is pick the rendered image icon that is most satisfactory, which implicitly selects the most suitable TF. The major challenge for this method is that possibly hundreds of volume rendering results have to be created for a user to choose from. These random TFs need to be generated by the system such that they result in the widest spread of dissimilar output renderings. This implies that an automated way of judging dissimilarity is available, and the Design Gallery method hence has data-dependent

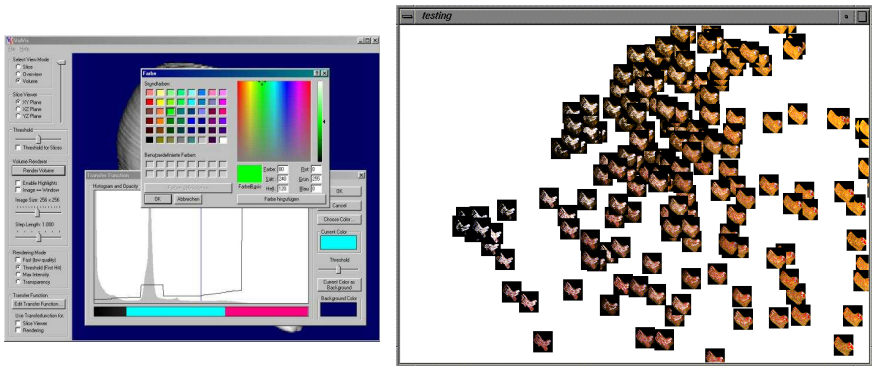



Figure 2.6: Two user interfaces for TF specification. Left: a trial-and-error interface; Right: the Design Gallery (Photograph reprinted from [MABea97]).

characteristics through this dissimilarity measure. As far as the author knows, there is little or no experimental information on how reliably the user can judge the results of the alternative renderings based on the relatively small image icons and how effective users can search using this method. Because a large number of image renderings are required, Design Gallery also relies on real-time volume rendering functionality to be feasible.

Kindlmann's semi-automatic method uses data-dependent properties to generate an optimized transfer function. It makes the reasonable assumption that the features of interest in a data set are often the boundaries between different materials [KD98]. By making use of the relationship between the data values and their first and second derivatives along the gradient direction, Kindlmann's method can generate one solution for the TF from the multi-dimensional scatter plot of data values. It tries to remove the user from the interaction process and does not provide any intuitive interface. This method is very sensitive to noise and could not generate desired results for data with noise [PLB+01]. This automatic method is obviously data-dependent, and cannot be guaranteed to provide results that agree with user expectations. It may however be useful in the initialization stage. The automated method of Tzeng [TL03], on the other hand, uses a more intuitive interface and combines user input through a neural network in order to select and adjust the TF. The user can for instance indicate areas in the rendered image that he finds interesting or not. It is a data-dependent method and achieved good results for one MRI data set. It is however not clear how their results extrapolate to other data sets. Their results can also not be reproduced, since the implementation details of their neural network are unknown.

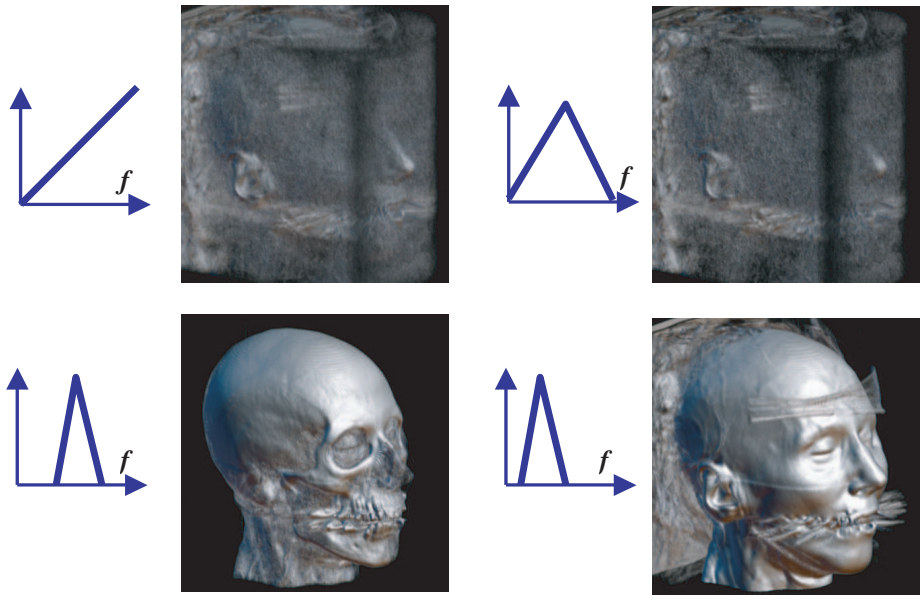



Figure 2.7: Results of TFs versus rendered images for CT scan data of a head.

Recent work from Rezk-Salama et al. introduces a high-level semantic model with a simple user interface for TF specification [RSKK06]. It borrows the concept of technical director in computer animation. A technical director compiles combinations of the low-level parameters required for each motion into high-level parameters and hides the complex set-up of low-level parameters from the animator. Rezk-Salama proposes that the visualization expert who is familiar with all the parameters involved in the image generation may play the role of a technical director. However, this method is still not successful in overcoming the major difficulty in a TF specification process. Firstly, the proposed method is only tested with CT angiography data. The effectiveness for other more complex data, for example MRI data, is not clear. Secondly, it still asks for the cooperation between a visualization expert and a non-expert user, which is often impossible in practice. Thirdly, it is still a technique that is mature from a technical point of view, but not practical from a HCI point of view, because it does not answer the fundamental interaction question behind the TF specification.

In summary, finding an appropriate TF can be described as a time-consuming and unintuitive interaction task with all available methods. As Rezk-Salama describes, although many existing techniques are mature in terms of technical implementation, the complexity

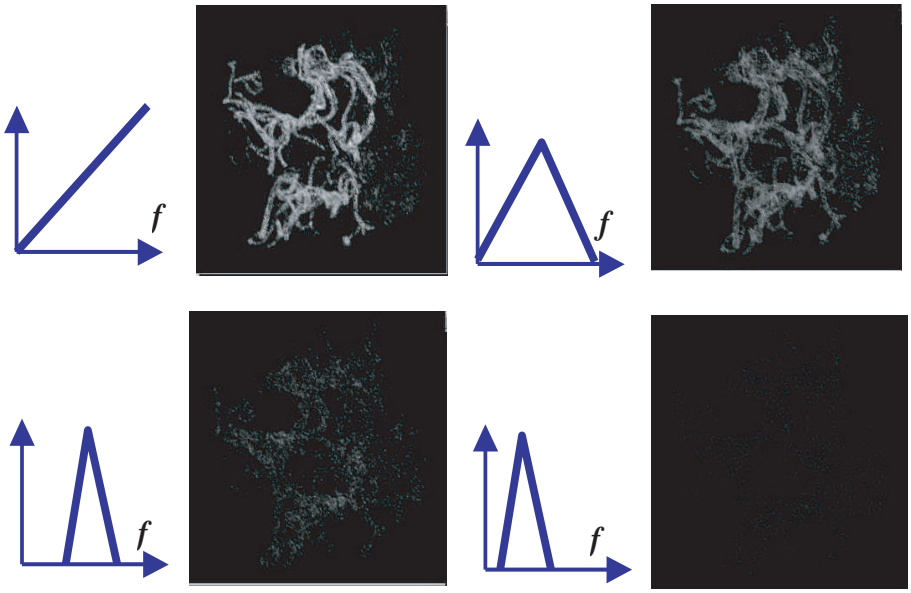



Figure 2.8: Results of TFs versus rendered images for a data set containing an aneurism.

of managing too many view parameters is still the same [RSKK06]. In this thesis, the main problems in TF specification that the author has identified are the following:

1. *Too many degrees of freedom (DOF) in the interaction process.* As simple and direct as the principle of a TF is, it is extremely flexible as well, because of the immense variety of possible TFs. With the trial-and-error method, a user arbitrarily and repeatedly manipulates the coefficients of the mathematical representation of the TF, in order to adjust the visualization outcome. Common forms of such mathematical representations are piecewise linear functions or higher-order splines. Each control point in the graphical representation of a TF has two DOFs, because all control points are located in a 2D plane, so that even with a limited number of control points the number of DOF can be substantial. In the case of a TF specification that assigns opacity to grey values, and provided these grey values range from 0 to 255, the total number of DOFs is theoretically equal to 256.

   A user is usually guided in his interactions with the TF by how closely the rendered image matches his interaction goal, which is most often to adequately reveal specific

structures in the data set. Slight changes in TF can however change the result dramatically (as seen in Figure 2.7). Moreover, the specification of the TF is data-set dependent. For different data sets, the same TF will not achieve similar results, as is illustrated by comparing (Figure 2.7) and (Figure 2.8). A TF that is well suited for one data set may be completely useless for another one. Therefore, even for a visualization expert it is a hard task. Most of the targeted users of visualization tools are moreover domain experts, that cannot be expected to have a deep understanding of the relationship between the TF and the rendered image.

2. *Inappropriate design of the user interface and inadequate information for TF specification.* The amount of control that a user has over the TF is determined by the interface. Besides the fundamental problem of the large number of DOFs, a poor design or arrangement of the user interface can make the TF specification more difficult and less efficient, especially when useful information is not available. This may be a contributing factor to why many of the available interfaces frustrate the user and fail. So far, there have been several suggestions for solving this problem by creating more intuitive interfaces (for example, the Image Graph [Ma99]). Most of these suggestions have however never been evaluated in a formal user study. Therefore, it is not clear which information is really useful to the end user, and which one is not.

## 2.4 Empirical Work

### 2.4.1 Experiment Design

TF specification, which is the interaction for controlling visual mapping, has been listed among the top ten problems in volume rendering [PLB+01]. The author proposes to use empirical research to get a better grip on the problem. The analysis of the interaction process, summarized in Figure 2.5, has indicated that a proper method for TF specification should provide useful information for starting the process and should support continuous feedback during exploration.

The proposed interface solutions for TF specification from previous research can be classified into three categories:

1. Trial-and-error method
2. Image-based methods
3. Data-based methods

Among those methods, the trial-and-error method simply asks a user to modify the TF curve directly. Imaged-based methods focus on presenting the rendering results of multiple TF set-ups at the same time so that users do not need to go through the iteration loop. But they do not consider how to design user interfaces for modifying and interacting with TFs. Even though these methods save users from adjusting the TF curve repetitiously, they create another problem, which is how to arrange **a large amount** of rendered images effectively in a limited display space. Data-based methods utilize the edge detection principles from image processing research and search for salient boundaries to render the desired structure. There exist many methods to perform edge detection, most of which belong to one of two categories: gradient and Laplacian. However, even with these methods, the problem of edge detection is not completely solved. First, it is difficult to design a totally automatic method. Although there are some automatic methods available, they do not work effectively and are not robust enough to create the results a user wants. Second, completely removing users from the process does not seem to be desirable because users are involved in making the final judgement of the result. Edge detection continues to be one of the most complicated research problems in image processing and computer vision. Therefore, data-based methods suffer from the problems of incompetent algorithm design and failing to deliver satisfying results consistently.

Improved solutions for TF specification should be those that can minimize the efforts for a user from exploration to refinement. The trial-and-error method is the easiest to understand and still the most widely used approach in volume rendering applications. It provides minimal visual feedback, and the author therefore adopts it as the baseline system. The author devised an experiment in which he explored the usefulness of additional information for the trial-and-error method. Extra information is based on either previous relevant experiences of TF or data-dependent information. More specifically, the author aims to assess the effects of the following additional feedback: 1) data-dependent information, such as the histogram of a data set; 2) data-independent information, such as suggested or standard

TFs. The author also wanted to investigate the effect of a GUI with a limited number of DOFs.

The author a priori formulated three hypotheses. All hypotheses are based on the assumptions that additional information that relates to the TF or the data set will help a user in finding the correct TF.

1. *Data-dependent (more specifically, histogram) information assists the users in TF specification.* Most available interfaces offer such information, so that it seems to be generally accepted that it can help a user in his/her search for a proper TF. The most frequently provided information is the grey-value histogram. It graphically depicts the frequency distribution of grey values in the data. In a standard histogram, the horizontal axis represents the range of grey values from 0 (shadows) on the left to 255 (highlights) on the right. The vertical axis represents the number (or percentage) of pixels that have each one of the 256 grey values. The higher the line coming up from the horizontal axis, the more pixels there are with that grey value in the data. In a cumulative histogram, the vertical axis represents the number (or percentage) of pixels that have a value smaller than and equal to each of the 256 grey values. The cumulative histogram integrates the standard histogram, and therefore has a more regular shape. The author uses the cumulative rather than the standard histogram as the data-dependent feedback in the experiment.

2. *Data-independent (more specifically, suggested TF) information assists the users in TF specification.* It is supposed that data-independent information comprises suggestions for a user to narrow down his/her search. These suggestions can for instance be in the form of geometric shapes of TFs, such as triangular, rectangular, hat-shaped, level-up, up-level, and step-like functions. These suggestions are derived from many rendering experiences. Although the TF is data dependent, there are simple TFs, such as piecewise-linear TFs, that often produce reasonable results. Higher-order spline representations of the TF are more difficult to control and seem to have only limited added value in most cases. Moreover, within the class of piecewise-linear TFs, not all shapes are equally likely to produce meaningful results, and the shapes that are expected to be most useful a priori can be suggested. Decreasing functions will for instance be absent from these suggestions, because they do not often create useful

```
0   #if defined GL_SGI_texture_color_table
1
2               glEnable(GL_TEXTURE_COLOR_TABLE_SGI);
3
4               glColorTableSGI(
5                   GL_TEXTURE_COLOR_TABLE_SGI, // GLenum  target,
6                   GL_RGBA,                    // GLenum  internal format,
7                   m_nColorTableSize,          // GLsizei width,
8                   GL_RGBA,                    // GLenum  external format,
9                   GL_UNSIGNED_BYTE,           // GLenum  data type,
10                  m_pColorTable);             // const GLvoid *table
11
12  #endif // GL_SGI_texture_color_table
```

Figure 2.9: OpenGL set-up for the texture color table extension.

results.

3. *A GUI with limited DOFs in the TF control assists the users in TF specification.* A user manipulates the TF via a GUI. Because the main difficulty for TF specification is too many DOFs, it is reasonable to think that users might have less difficulty if they are presented with an interface with limited DOFs. The author therefore will also test the case where the piecewise-linear TFs that the author describes above are not only provided as suggestions, but are actually the only shapes available to the user. In order to properly evaluate this case, the author will obviously not only have to look at the time that people take to realize a TF setting, but also at the quality of the result that they produce.

### 2.4.2 Apparatus

In order to enable us to experimentally investigate the usability aspects of TF specification, the author created a volume visualization prototype. Our experimental hardware set-up consists of a DELL graphics workstation (Pentium IV, 2.4 GHz, 512 MB RAM, ATI *FireGL 4* graphic card); a 17' CRT monitor; a 14' CRT monitor; a keyboard and a mouse; and loudspeakers (stereo). The key software component of the system is a volume-rendering engine that visualizes volumetric data with the help of hardware-accelerated 3D texture

mapping [SWWL01]. This implies that the TF specification is implemented by means of a texture look-up table. The OpenGL extension to access this feature is called SGI texture color table. The extension must be simply enabled and a color table must be set-up as described in Figure 2.9. The color table lookup is performed after the texture interpolation. Usually the texture color table, which is enabled by this extension, is not restricted to a specific texture object, so it can be efficiently shared among multiple texture images.

### 2.4.3 Interfaces

In the experiment, the visual feedback at any time consists of a single TF, with its available controls and feedback, and the correspondingly rendered image of a scientific data set. The subjects interact with the TF via a GUI. The experiment involves five interface conditions.

- The baseline interface with free-style control, referred to as condition 1, consists of parts 1a and 1b in Figure 2.10. With this free-style interface, a user has full control over the TF. The panel 1b controls the course of the experiment, i.e., starting and stopping a single TF control trial, saving the rendered image, or loading a new data set. The part 1a allows the user to manipulate the TF by creating and moving control points of a piecewise linear function along the horizontal and vertical direction within a 2D interaction area. There is no movement limitation for the control points except that the grey values for the first and last point have to remain at 0 and 255, respectively. The user can create TFs with as many control points as he wants.

- Experimental condition 2 includes data-dependent information, and consists of parts 1 and 2 in Figure 2.10. A cumulative histogram and free-style TF interface are presented at the same time.

- Experimental condition 3 includes data-independent information, and consists of parts 1 and 3 in Figure 2.10.

- In condition 4, both data-dependent and data-independent information are offered, so that all parts in Figure 2.10 are presented.

- The interface for the final condition 5 is shown in Figure 2.11. It is a user interface

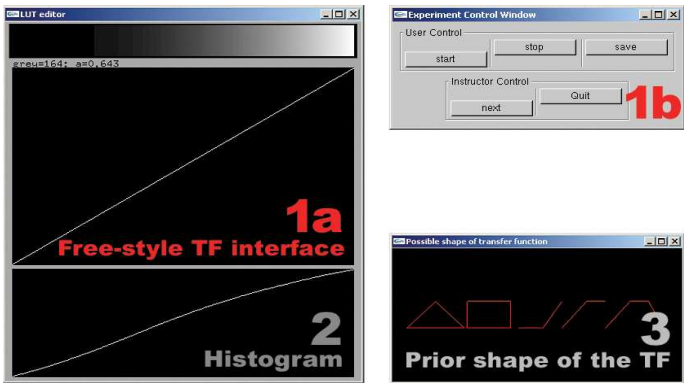



Figure 2.10: The user interfaces for experimental conditions 1 (part 1), 2 (part 1+2), 3 (part 1+3) and 4 (part 1+2+3).

that allows for a number of piecewise-linear TFs, and that does not provide (data-dependent) histogram information. Each kind of TF is represented by a graphical icon. The available TF graphical representations have only few control points and limited DOFs. The shape of a graphical representation cannot be altered.

### 2.4.4 Experimental Procedure

There were 14 subjects in the experiment, six female and eight male persons between 19 and 50 years old. All of them had university education in engineering or science. Each subject participated in all 5 conditions (within subject design). The order in which conditions were presented to the subjects was randomized. The subjects were given a consent form to read and sign.

Upon entrance, subjects were given an experiment instruction sheet that described the system and tasks to be performed. These written instructions remained available during the entire experiment. The subjects were introduced to all five user interfaces and could interact with them, using a data set that was not part of the actual experiment for the training purpose. Afterwards, each participant performed four tasks in each of the five interface conditions. Each task involved a different data set, and required the subjects to visualize a pre-described structure within the data as well as possible, as shown in Figure 2.12. For example, one task was to ask the subjects to find a TF that can create a 3D image that

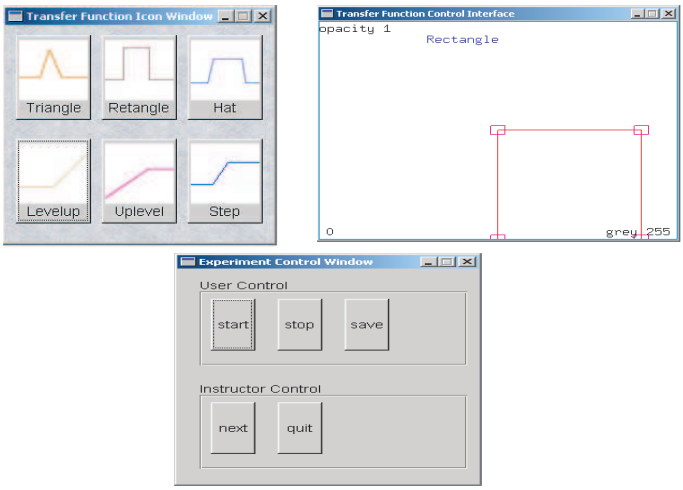



Figure 2.11: The user interface for experimental condition 5.

contains only the bone structure of the skull similar to Figure 2.12, upper left. The order in which the four tasks were executed with an interface was also randomized across subjects.

Table 2.1: Variables lists

| **Independent Variable** | **Dependent Variables** |
|---|---|
| User interfaces<br>(trial-and-error baseline interface,<br>baseline with histogram,<br>baseline with predefined TF,<br>baseline with predefined TF and histogram,<br>baseline with limited-DOF control) | Time,<br>quality,<br>number of mouse clicks,<br>mouse clicks for each icon in condition 5 |

The five different interface set-ups constitute the independent variables and represent different conditions that correspond to the hypotheses (Table 2.1). The experimental data recorded during the experiment as dependent variables were the follows:

- the time needed to finish a task;
- the difference between the rendered image produced at the end of a task and the target image as the quality of the performance;

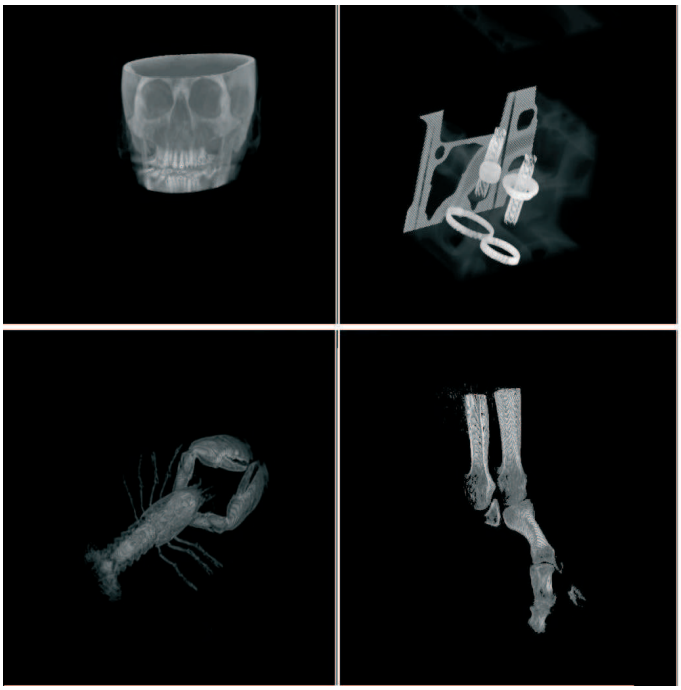


Figure 2.12: The required structure being rendered with each data set.

- the total number of mouse clicks during a task;
- the number of mouse clicks for each icon in condition 5;

The subjects could stop the interaction process when they thought they got the most satisfying rendering result for a task. The subjects were given several questions to answer after the experiment in order to collect their subjective impression of the interfaces and the produced images. All the subjective ratings of the subjects were performed on a 7-point scale. The design of the questionnaire was based on an available usability questionnaire [GHD99], [Lew95]. More specifically, the questionnaires contained the following parts:

- Personal data, such as age, education, and former experience or knowledge about visualization systems and the problem of TF control;
- Their agreement or disagreement with general usability statements about the system, such as: "It was easy to use the system";

- Questions addressing the usability of the individual interfaces. More specifically, subjects were asked to rate effectiveness, efficiency, satisfaction and overall quality (the detailed questions are specified in Appendix B);
- In order to assess the task performance, which was defined as the degree to which the produced images matched the subjects' goal, they were asked both to rate the output images individually and to express their preferences for all pairwise combinations of output images [Ahu93], [EF95];
- Follow-up questions with an open answering format, in order to collect additional (a priori unexpected) comments from subjects.

## 2.5 Results

In the following sections, the author discusses the quantitative results of the experiment, as well as the results from the subjective evaluation by means of the questionnaires.

### 2.5.1 Quantitative Data

The left part of Figure 2.13 shows the mean response time for all five conditions and all four tasks (or data sets). Table 2.2 shows the absolute values for five conditions. "Analysis of Variance"($ANOVA$) with repeated measures and significance level of $\alpha$ =0.05 was carried out on the logarithm of the response time. There were no significant differences on the time spent among the five conditions for the "lobster", "head", and "foot" data sets, respectively. For the "engine" data set, the order of the conditions with respect to response time was one to five from shorter to longer, except that the response time in condition 2 was shorter than in condition 1. The detailed results with each data set are the following:

**Engine** There were significant differences among the five interface conditions, F(4, 52)=7.466, $p = 0.000 < 0.05$. The post-hoc test shows that between condition 1 and 5 ($p = 0.001 < 0.05$), condition 2 and 5 ($p = 0.001 < 0.05$), and condition 3 and 5 ($p = 0.002 < 0.05$), there were significant differences on response time. Condition 4 and 5

also demonstrated a significant difference ($p = 0.007 < 0.05$). In absolute terms, the mean time spent in condition 5 was the longest.

**Foot** There was no significant difference among the five conditions, F(4, 52)=0.357, $p = 0.838 > 0.05$. The results of post-hoc tests indicate there was no difference between pairs of conditions.

**Head** There was no significant differences among the five conditions, F(4, 52)=0.682, $p = 0.607 > 0.05$. The results of post-hoc tests indicate there was no difference between pairs of conditions.

**Lobster** There was no significant differences among the five conditions, F(4, 52)=0.662, $p = 0.661 > 0.05$. The results of post-hoc tests indicate there was no difference between pairs of conditions.

Table 2.2: Performance Time in different tasks (in second)

| **Trial No** | **Condition 1** | **Condition 2** | **Condition 3** | **Condition 4** | **Condition 5** |
|---|---|---|---|---|---|
| NO. 1 (Engine) | 116.3 | 104.5 | 138.6 | 177.9 | 313.7 |
| NO. 2 (Foot) | 185.6 | 162.3 | 168.5 | 192.2 | 209.9 |
| NO. 3 (Head) | 204.4 | 196.3 | 290.4 | 290.9 | 262.5 |
| NO.4 (Lobster) | 253.3 | 228.6 | 249.2 | 228.1 | 208.7 |

The right part of Figure 2.13 shows the average number of mouse clicks for all five conditions and all four data sets. Table 2.3 shows the absolute values for five conditions. There was no significant effect in either of the four data sets among the five conditions. However, with the "lobster", "head" and "foot" data sets, there was a tendency that the number of mouse clicks becomes lower in condition 5 than in the other four conditions. *ANOVA* with repeated measures on the number of mouse clicks for all four tasks shows the following statistical details:

**Engine** There was no significant effect, F(4, 52)=2.411, $p = 0.061 > 0.05$. The results of post-hoc tests indicate there was no difference between pairs of conditions.

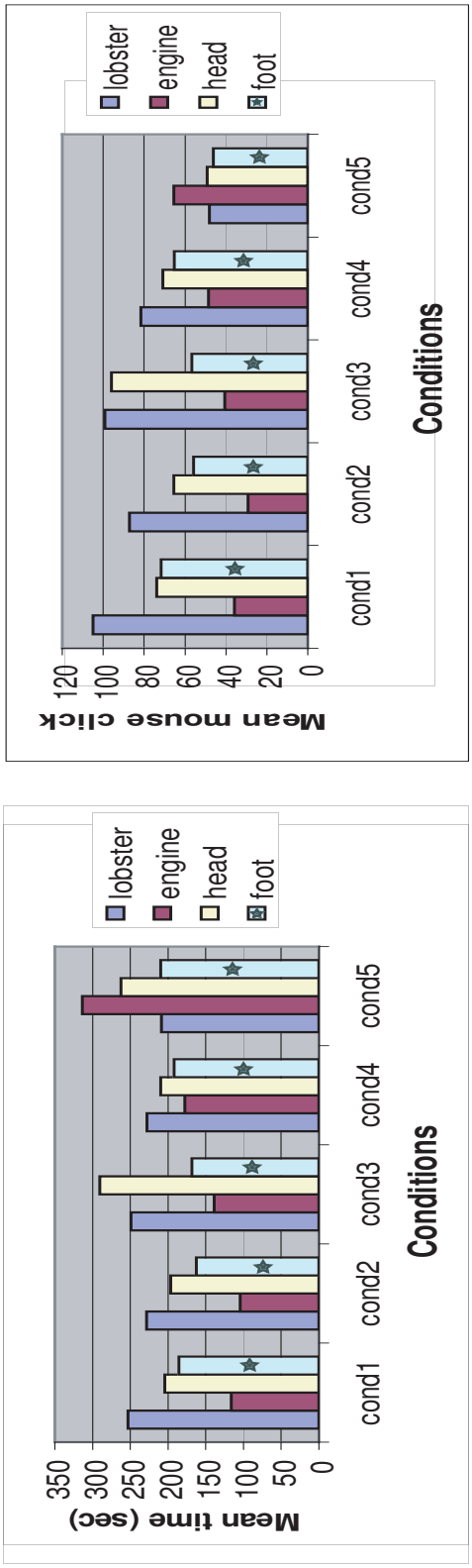



Figure 2.13: Mean time and number of mouse clicks for five conditions with four data sets.

**Foot** There was no significant effect, F(4, 52)=0.774, $p = 0.547 > 0.05$. The results of post-hoc tests indicate there was no difference between pairs of conditions.

**Head** There was no significant effect, F(4, 52)=1.841, $p = 0.135 > 0.05$. However, the results of post-hoc tests indicate there was a significant difference between condition 3 and condition 5 ($p = 0.04 < 0.05$).

**Lobster** There was no significant effect, F(4, 52)=2.202, $p = 0.082 > 0.05$. However, the results of post-hoc tests indicate there was a significant difference between condition 1 and condition 5 ($p = 0.035 < 0.05$), between condition 2 and condition 5 ($p = 0.042 < 0.05$), and between condition 3 and condition 5 ($p = 0.024 < 0.05$).

Table 2.3: Number of mouse clicks in different tasks

| **Trial No** | **Condition 1** | **Condition 2** | **Condition 3** | **Condition 4** | **Condition 5** |
|---|---|---|---|---|---|
| No. 1 (Engine) | 35 | 29 | 40 | 48 | 65 |
| No. 2 (Foot) | 71 | 56 | 57 | 65 | 46 |
| No. 3 (Head) | 74 | 66 | 96 | 71 | 49 |
| No.4 (Lobster) | 105 | 87 | 99 | 82 | 48 |

### 2.5.2 Subjective Evaluation

This subjective evaluation is based on the answers of the questionnaire (see Appendix B). Firstly, the four tasks were evaluated in terms of difficulty in order to find out whether or not there is a correlation between the difficulty of a task and the user performance. Then the results are mainly summarized along four important characteristics, i.e., overall image quality, effectiveness, efficiency, and satisfaction. These subjective evaluations were also compared against the quantitative results in order to test our hypotheses.

**The difficulty of the tasks** The tasks with these four data sets presented different levels of difficulty for the subjects, as is shown in the Figure 2.14a. The tasks with "head" and

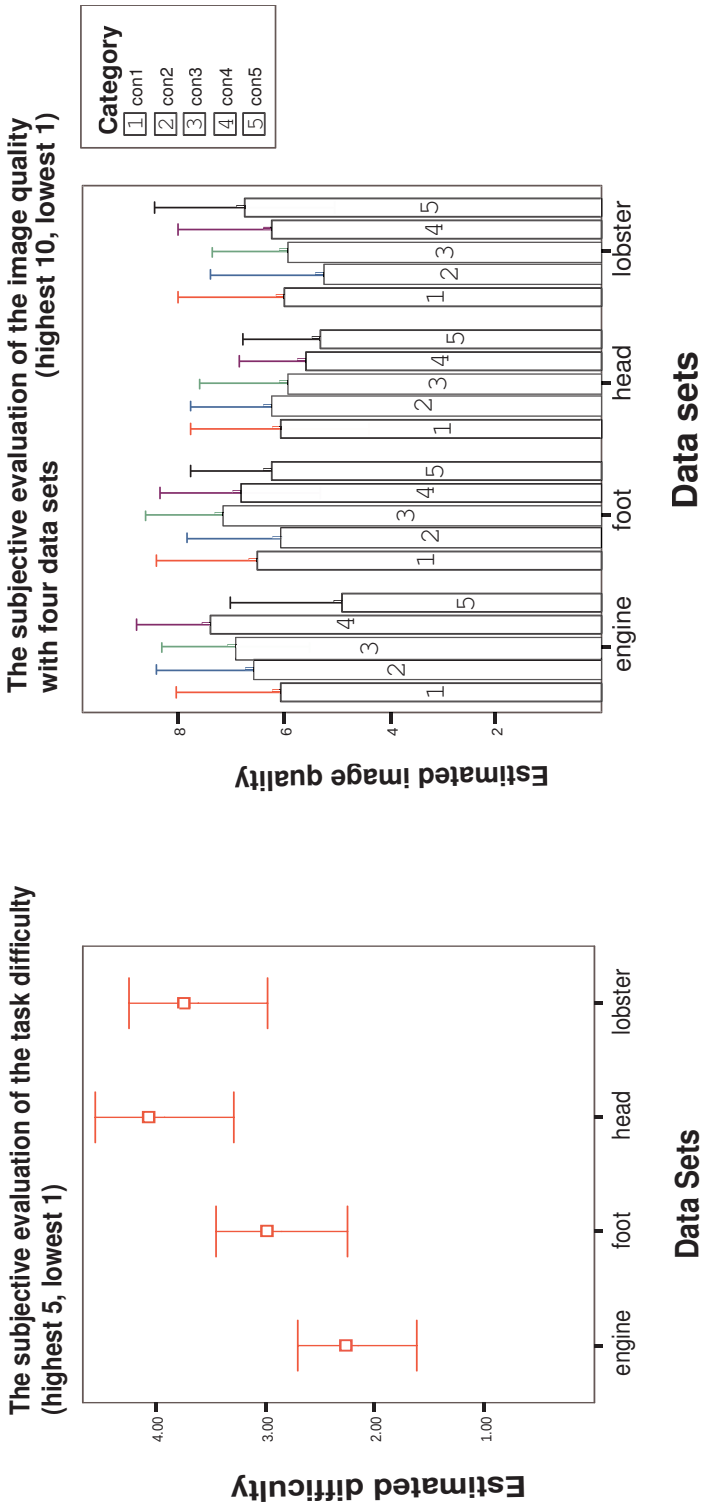



Figure 2.14: The estimated difficulty and performance (image quality) for the four tasks (i.e., data sets). Bars show means; error bars show 95% confidence interval of the mean. (a): difficulty; (b): performance.

"lobster" data sets were significantly more difficult than the "engine" data set. The task with the "engine" data set was recognized as the easiest one. The most difficult task is the one with the "head" data set. The feedback from the subjects indicated some of the reasons. The task with the "engine" data set could be finished satisfactorily with a simple TF. Most of the subjects found that the task with the "head" data set required too many details to be rendered simultaneously, which required detailed adjustment of a TF. The differences in difficulty between the four tasks are assumed to reflect the real situations encountered in common practice of TF specification and should therefore be taken into account in the experiment.

**Task performance Versus image quality** To evaluate the effectiveness of the five different interfaces, task performance in terms of the image quality of the final result is taken into account and measured. The image quality was scaled by how close the produced image is compared with the original goal. Figure 2.14b illustrates that with the "engine" and "head" data sets, the task performance in condition 5 is the worst. However, the best performance was achieved in condition 5 with the "lobster" data set.

**Effectiveness** In terms of effectiveness, the subjects were asked to evaluate "which condition or interface provided more control over the TF". Condition 4 got the highest rank, with condition 2 as a close second. Condition 5 was considered the least effective. Figure 2.15 upper left shows the details. *ANOVA* with repeated measures on the effectiveness for all five conditions shows that there is a significant difference between condition 1 and condition 4 ($p = 0.005 < 0.05$), and between condition 4 and condition 5 ($p = 0.015 < 0.05$).

**Efficiency** Efficiency was defined as "how fast the user feels that (s)he can finish the task" with each of those interfaces. Figure 2.15 upper right illustrates that condition 4 again scored best, while condition 5 was the worst. *ANOVA* with repeated measures on the effectiveness for all five conditions shows that there was a significant difference between condition 1 and condition 4 ($p = 0.044 < 0.05$), between condition 4 and condition 5 ($p = 0.009 < 0.05$). There was also significant difference between condition 2 and condition 5 ($p = 0.013 < 0.05$).

**Satisfaction** Look and feel is a very important factor in the design of an interface. The

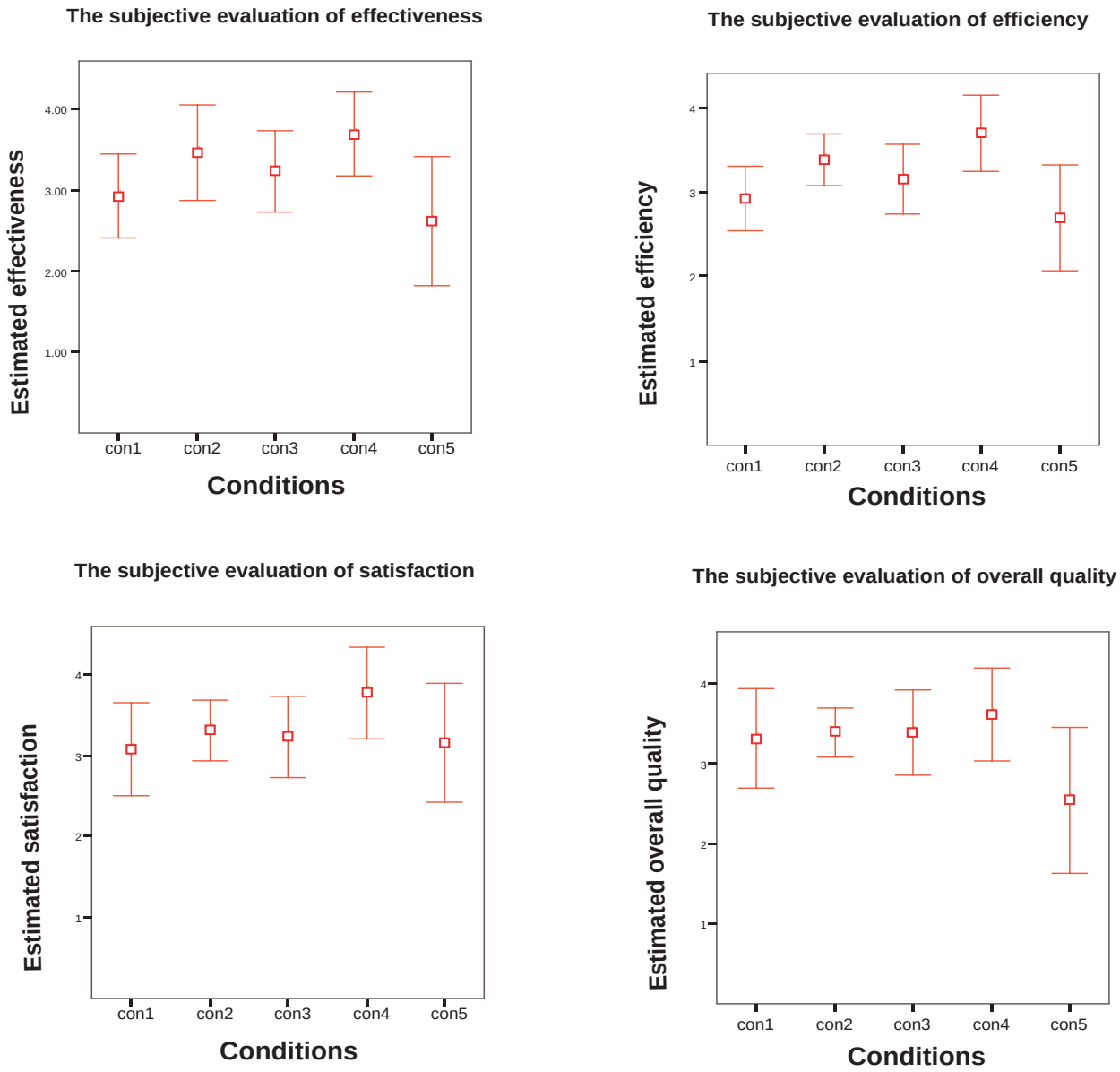



Figure 2.15: The subjective evaluation of five TF interfaces on four attributes.
Bars show means; error bars show 95% confidence interval of the mean.
Upper Left: effectiveness; Upper Right: efficiency.
Lower Left: satisfaction; Lower Right: overall preference.

subjects were requested to give an evaluation on "the arrangement of each interface". The results are illustrated in Figure 2.15 lower left. Surprisingly, condition 4 still got the highest appreciation. Condition 5 scored higher in this attribute than in the previous attributes.

**Overall quality** Considering all three factors above (effectiveness, efficiency and satisfaction), subjects gave their evaluations of the overall quality of the five interfaces. The results are shown in Figure 2.15 lower right. Condition 4 was ranked as the highest. Condition 1, 2, 3 were very close to each other, while condition 5 was least appreciated. None of the difference is however significant.

## 2.6 Discussions

The experimental results do not provide strong evidence to support the three hypotheses described before. In this section, the author discusses the results and the reasons individually.

### 2.6.1 With and without integrated histogram

Neither quantitative nor qualitative analyses of the results support the hypothesis that the histogram can assist a user during the TF specification process. Although most of the subjects could understand the principle and the purpose of a histogram, it was still hard to apply it during the specification process. The usage of a histogram requires a user with more mathematical background or knowledge, which hindered the subjects from deriving useful information from it easily during the course of the experiment. Only two out of thirteen subjects who had experiences with Photoshop and the histogram feature within it, gave positive feedback on the use of a histogram. These subjects also used the information about grey value distribution that they derived from the histogram in the construction of their TFs. Therefore, it seems that, for untrained subjects, a histogram is not very useful and does not assist in the TF specification. It indicates that the way that the histogram has been presented is not straightforward enough for a user to derive useful information for the interaction process.

### 2.6.2 With and without Additional Transfer Function Information

Contrary to the hypothesis, no statistically significant differences have been found between these two conditions. The inclusion of additional TF information did not help a user to speed up his/her exploration and refinement process. However, several users who had less background knowledge in direct volume rendering or TF did appreciate this information, and they preferred it to be present, because it helped them to get started. Since the information about possible TF configurations does not present any active mechanism to shorten the search time, it is not unexpected that it did not improve the user performance in terms of execution time.

### 2.6.3 Free-style versus limited DOFs

Surprisingly, the GUI with limited DOFs in condition 5 did not help subjects in terms of both response time and the number of mouse clicks. For the "engine" data set, which represented the easiest task, the mean interaction time spent even increased. However, the task performance in terms of image quality was best for the "lobster" data set. This implies that the GUI with limited DOFs in condition 5 might not be suitable for simple TF specification tasks, nor for difficult tasks, such as in case of the "head" data set. The comments from the subjects suggest that posing limitations in DOF were mostly useful for exploring the data and accumulating experience. They soon became an obstacle in finishing a task that requires a user to perform more subtle adjustments to a TF, such as in the case of the "head" data set. But it can be suitable and useful for tasks that have intermediate difficulty.

### 2.6.4 Working Memory for Transfer Function Specification

Based on the experimental results and using knowledge from psychology and perception, a further analysis can explain why the specification task is so difficult. Human beings have two different storage systems with different durations: working memory and long term memory [WH99]. Working memory is the temporary, attention-demanding store that a human user uses to retain new information until he or she uses it [Bad86]. A human user uses working memory as a kind of "workbench" of consciousness where he or she examines,

evaluates, transforms, and compares different mental representations. A human user might use working memory, for example, to carry his or her mental arithmetic or predict what will happen if he or she set a TF one way instead of another. Working memory is used to hold new information (for instance the resulting image with one TF setting) until a human user gives it a more permanent status in memory, that is, encodes it into long term memory.

Several experiments have demonstrated the transient character of working memory [Bro59]. Estimates generally suggest that without continuous rehearsal, little information is retained beyond 10 to 15 seconds. This transient character of working memory presents a serious problem for those work domains/tasks when information can not be rehearsed because of time constraint [Mor86].

Working memory is also limited in its capacity (the amount of information it can hold) [Bad90]. And this limit interacts with time. Experiments show that faster decay is observed when more items are held in working memory, mainly because rehearsal itself is not instantaneous [Mel63]. The limiting case occurs when a number of items can not successfully be recalled even immediately after their presentation and with full attention allocated to their rehearsal. The limiting number is sometimes referred to as the memory span. In a classical paper, George Miller identifies the limits of memory span as the magical number seven plus or minus two [Mil56]. Thus, somewhere between five and nine items defines the maximum capacity of working memory when full attention is deployed.

Task analysis suggests that TF specification is a task that puts high demands on working memory from a user. When a user uses a TF interface to search for required results or structures, he or she continuously inputs different parameters for the TF through the interface and judges whether or not the corresponding rendering results are the ones he or she needs. Often he or she needs to retrieve previous settings that are better after comparison. A user needs to perform so many interactions (modifying the TF parameters and observing the corresponding visual feedback) and has to hold mapping information between TF setting and visual feedback with respect to each data set in working memory, which introduces more possibilities for error. The loss of mapping information often leads to unnecessary repetitive work. Clearly, the limited capability of working memory has a major impact on the effectiveness and efficiency of a TF interface. On the other hand, it also indicates that a user interface will be more efficient and effective if it can relieve the workload of a user's

working memory during the specification process.

## 2.7 Conclusion

The described experiment is the first step towards more quantitative investigations on the usability issues of user interfaces for TF specification in direct volume rendering. More specifically, the author has compared five interface prototypes in order to find out whether or not specific instances of data-dependent and data-independent feedback can assist a users in this specification task. The obtained results can be summarized as followings:

- There is no evidence that histograms can help to improve a user's performance in terms of response time and quality of rendering results.
- Additional information about possible/suggested TFs may be useful to novel users in the beginning of an interaction process, but do not assist a user in finding the required TF.
- Interfaces that restrict the number of DOF of a TF also do not improve the chance to find a suitable TF, and are moreover not better appreciated by users.

The trial-and-error method is a basic and important scheme to help users interact with the TF because it assigns a user the central role in the interaction dialog. The data-dependent and data-independent feedback mechanisms, which were proposed and used in this study and other interface prototypes from other researchers, did not substantially improve the effectiveness and efficiency of the trial-and-error interface. The working memory theory provides a possible reason that subjects failed to finish the tasks effectively and efficiently: the experimental interfaces do not provide any mechanisms to relieve the workload of working memory. This suggests that a designer of TF user interfaces should take the limitation of working memory into account, besides the look-and-feel of an interface. A well-designed user interface for TF specification should alleviate the memory workload and facilitate the interaction process by presenting means to supply mapping information between TF and corresponding visual feedback whenever necessary, such as in the "Design Gallery" interface [MABea97].

It is not the author's intention to study TF specification extensively and to investigate all the available methods in this thesis, but to introduce the usability issues unsolved. It is the author's intention to present the methodology necessary for studying this question through investigating a GUI for the trial-and-error method. The author also advocates that future research should focus on usability testing and user evaluation on available methods, which has not been regarded as an important part of the design process, instead of developing techniques and interfaces solely. For example, a possible next step in this research would be to investigate how image-based and data-based automated methods (which are most useful to improve the initialization phase of the interaction), compare to the trial-and-error method. Further experiments, which can profit from the experimental methodology presented in this paper, should help to address this parameter optimization question more quantitatively.

# Chapter 3

# User Performance with Three Virtual Reality Systems

Although Jaron Lanier initially coined the term "Virtual Reality" (VR) in 1989, the history of VR can be traced back as early as 1960. Ivan Sutherland described his vision of the "ultimate display" in his presentation to the International Federation for Information Processing (IFIP) Congress, which can be regarded as the earliest proposal for a 3D simulated environment. Other researchers since have proposed several related concepts, which include "Artificial Reality" [KBF+95], "Cyberspace" [Gib]. More recently, "Virtual World" and "Virtual Environment" are also used to refer to a VR-like system or interface. Initially, the term VR referred to "Immersive VR with head mounted display". In immersive VR, a user becomes fully immersed in an artificial 3D world that is completely simulated by a computer. Currently, VR exists in different forms and refers to different designs, sometimes in a confusing and misleading way. Although it is difficult to categorize all VR systems, this thesis separates them based on their display technology:

- projection-based VR systems (e.g. CAVE [CNSD93] in Figure 3.1 or workbench [KBF+95] in Figure 3.2).
- Head mounted display (HMD) VR systems [Sut68].
- Monitor-based desktop VR systems, (e.g. Fish tank VR [WAB93] ).

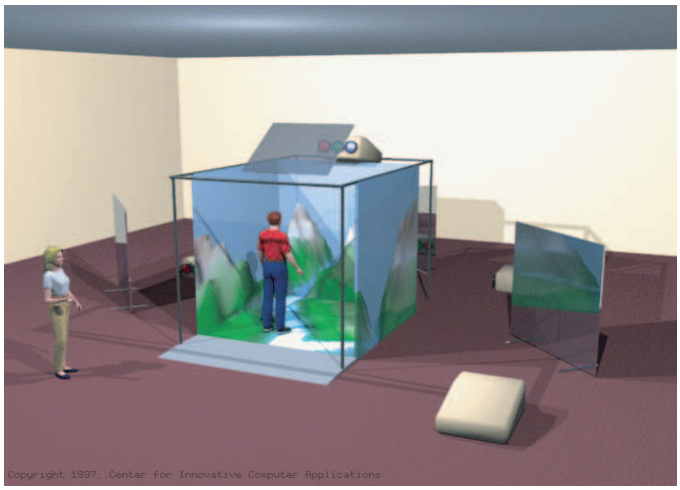



Figure 3.1: The immersive VR-CAVE (Photograph courtesy of Advanced Visualization Laboratory, University Information Technology Services at Indiana University).

Frederick Brooks defines a VR experience as "any in which the user is effectively immersed in a responsive virtual world. This implies user dynamic control of viewpoint" [Bro99]. As an emerging tool for scientific visualization, VR makes multisensorial, 3D modeling of scientific data possible. While the emphasis of scientific visualization is mainly on visual representation, other senses, such as touch or hearing are used as complements to enhance what the scientists or professionals can experience.

Visualization researchers increasingly use VR interfaces to build applications for professionals or domain scientists to display their data in 3D form [HJ04]. However, there are currently few guidelines regarding which type of VR system (interface) should be used for an intended task. Most existing guidelines are also not based on empirical evidence, i.e., supported by qualitative and quantitative analysis. This leads to a situation where applications may not use the most effective design to solve a domain scientist or professional's problem. In this chapter, the difference in user performance between alternative VR systems has been investigated through a controlled user study for four generic analysis tasks.

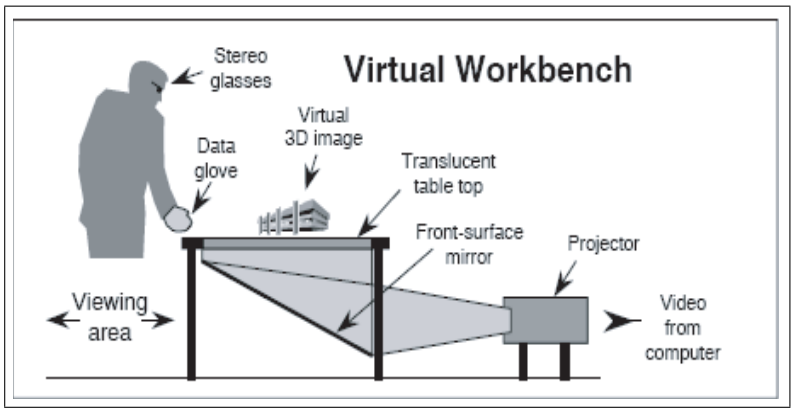



Figure 3.2: The Responsive Workbench (Photograph reprinted from [KBF+95]).

## 3.1 Scientific Problem and Tasks

The domain specialists that were involved are studying the structure of human lung mucus in both normal "wild-type" lungs and in the lungs of CF patients. This mucus is made up of a number of long polysaccharide molecules called mucins. It is known that there are a number of different types of mucin present in the mucus, and that the mucus is denser for CF patients than wild-type mucus. What is not known is how the different types of mucin are distributed in the mucus, and how particles can diffuse through it.

The mucins may be uniformly distributed, or form distinct domains. There may be web-like superstructures formed by a subset of the mucins which contain clumps of other mucins. There may be large, small, or a variety of differently sized water pockets surrounded by thin membranes. There may be a continuous water path with thin webs of mucins forming a lattice. The domain specialists are probing this by developing fluorescent dyes that attach specifically to each different mucin type; they subsequently scan the mucus with a confocal microscope to produce multiple 3D scalar fields, one for each dye. The author wishes to display the resulting scalar fields in 3D to help them estimate sizes, distributions, and shapes of any resulting voids and structural elements.

A virus, bacteria, or bacterial colony would traverse the mucus differently depending on its structure. The motion of such pathogens is of great interest to the study of CF, because

lung infections are the cause of many CF deaths. The domain specialists are probing this by placing small beads of various radii into the mucus and tracking the Brownian-driven motion of these beads over time to understand how they move through the mucus matrix. The author wishes to display the resulting motion paths in the presence of the above mesh structure to help them correlate structure and density with bead motion paths.

With these problems in mind, the data sets and questions of this user study have been designed to help determine which display and interaction system best supports the types of queries the domain specialists and users are asking without requiring our subjects to be experts in CF. The domain specialists are performing a number of tasks within a dense volumetric scalar field (for example, connectivity and relative density are of interest in addition to counting, shape, and size analysis). The author tried to provide tasks in the user study that were similar to these needs and yet as generic as possible, so that the results of the user study could apply to other applications that explore dense 3D scalar fields looking for structure and pathways. The author thinks that studying oil fields and blood vessels within tumor tissue might have similar needs for understanding dense data and for studying connectivity between areas within the data.

## 3.2 Related Work

There has been a great deal of effort in the VR and scientific visualization research community that aims at developing and integrating new devices (technologies) and interaction techniques and understanding and improving the usability of VR systems. User studies have been performed to investigate the usability and effectiveness of VR systems for interaction tasks inspired by real-world scenarios.

### 3.2.1 Stereoscopic Displays in VR and Scientific Visualization

The case for stereo in VR and scientific visualization is clear. Ware has shown that stereo combined with motion parallax enables improved user performance in the 3D visualization of graphs, which argues for using VR rather than a traditional desktop display [WF96]. Arthur's study demonstrated the advantages of a fish tank VR system over a desktop display for 3D tasks such as ....[ABW93]. Of interest to us is which type of stereo VR system is

most effective for scientific visualization of dense volume scalar fields.

### 3.2.2 VR Related Perceptual and Cognitive Issues

The Effective Virtual Environments (EVE) group at UNC Chapel Hill has conducted presence, locomotion and redirected walking studies within immersive HMD VR systems [MIWB02], [RW01].

Other researchers have studied the potential effects of VR or 3D interfaces on spatial memory and reasoning. For example, Peruch et al. tested the capability of an observer to learn spatial layouts of objects located in a graphically displayed wall-limited space. The results indicated that spatial acquisition resulting after active exploration was more accurate than after passive exploration and that dynamic and static passive visual information yielded equivalent performance [PVG95]. Bakker et al. studied the effects of head-slaved navigation and the use of teleports on spatial orientation in a virtual environment (VE) [BPW03]. They found that head-slaved navigation had an advantage over indirect navigation for the acquisition of spatial knowledge in a VE, at least for small-scale environments. However, the benefit of head-slaved navigation can only be achieved in applications in which a new spatial layout needed to be learned each time or when the primary users are novices. They also found that displacement through an environment is fastest when using a teleport, but may cause temporary disorientation.

There is additional evidence for the claim that spatial memory is improved in 3D. For example, Robertson et al. showed that task times and error rates were lower when retrieving web pages using their 3D user interface ("Data Mountain") than using the standard 2D "Favorites" mechanism of Internet Explorer [RCL+98]. In the comparison of effects of 2D and 3D interfaces on spatial memory, many of the findings are however dependent on the precise tasks under analysis ([WLPO95], [WF96]).

### 3.2.3 Interaction Styles and Techniques for Different Tasks

Werkhoven and Groen studied manipulation performance in a VE using two types of controllers: virtual hand control and 3D mouse control under both monoscopic and stereoscopic

viewing conditions [WG98]. The results showed that virtual hand control proved to be significantly faster and more accurate than 3D mouse cursor control. It was also shown that the speed and accuracy of manipulations are much improved under stereoscopic conditions.

Navigation in HMD versus CAVE has been studied by Bowman [BAR+02]. He presented a preliminary experiment comparing human behavior and performance between a HMD and a four-sided spatially immersive display (SID). In particular, he studied users' preferences for real versus virtual turns in the VE. The results indicated that subjects have a significant preference for real turns in the HMD over virtual turns in the SID. The experiment also found that females were more likely to choose real turns than males. This suggests that HMDs are an appropriate choice when users perform frequent turns and require spatial orientation.

Immersive versus fish tank VR for searching and labeling has been studied by Cagatay [DLJ+03], who compared fish tank VR and CAVE displays for a visual search task. The results of their qualitative study showed that users preferred a fish tank display to the CAVE system for a scientific visualization application because of a perceived higher resolution, brightness, crispness and comfort of use. The results showed that users perform an abstract visual search task significantly faster and more accurately in a fish tank environment, compared to the CAVE.

Schulze [SFKL05] presented a user study comparing performance across multiple immersive environments for a counting task. He tested three VR displays: a CAVE-like environment, a single-wall display, and a desktop system (fish tank VR). Data he collected led to four significant findings: (1) in the CAVE the subjects preferred medium sized or large spheres over small spheres; (2) when only a few targets have to be marked, larger spheres were marked faster than smaller spheres; (3) large spheres are marked most accurately; and (4) performance for the wall display was not comparable to the fish tank VR display when the spheres were small. Additionally, occlusion and a larger field of view inhibited performance in the CAVE more than in the fish tank display when the task was dominated by visual search.

### 3.2.4 Multimodal Interaction (Haptic Feedback)

The scientific visualization and VR communities are continually developing better techniques and technologies to represent data in a form suitable for comprehension. Traditional visualization schemes are entirely visually dependent. More and more VR systems for visualization applications take the multimodal interaction approach that is to incorporate other kinds of sensor modalities, for instance haptic feedback, into the interface. This section mainly describes the available work in haptic rendering for scientific data. Related work in auditory rendering can be found in Appendix A.

An early example of haptic representation of scientific data is found in the work of Brooks [BOYBK90]. Users are assisted by a force reflective master manipulator during a complex molecular docking task. In this work, a force display is used to drive the system towards a local minimum and indicate tightness of fit. The nanoManipulator (nM) [TCO+97] is a VR system that provides an improved, natural interface to data from scanning probe microscopy, including scanning tunneling microscopes and atomic force microscopes. The nM couples the microscope to a haptic VR interface that gives the scientist virtual telepresence on the surface, scaled by a factor of up to a million to one. The Visual Haptic Workbench [BIJH00] is another testbed system for conducting research on the synergistic benefits of haptic displays using an integrated, semi-immersive VE.

Several studies have shown the effects of a haptic display on human perception. Studies from Ernst have shown a clear influence of haptics on vision, demonstrating that vision does not necessarily dominate haptics [EB02]. The human central nervous system seems to combine visual and haptic information in a fashion that is similar to a maximum-likelihood integrator. Visual dominance occurs only when the variance associated with visual estimation is lower than that associated with haptic estimation. Sarter [Sar06] reviewed the current research in multimodal interaction and summarized several design guidelines as follows:

- Use multiple modalities only when necessary, i.e., when it supports users' preference, needs and abilities, increases bandwidth or assists disambiguation.
- Map the appropriate modalities (vision, touch, hearing and olfaction) to tasks and corresponding types of information.

- Combining, synchronizing and integrating different modalities must be compatible to the resulting spatial and temporal combination and synchronization of these channels.
- Multimodal interaction need to be flexible and should be able to adapt to changes in the user needs and abilities, his/her task and workload.

User studies should be designed to evaluate visualization methods, as Kosara [KHI+03] suggested. This also applies to VR systems with visualization capabilities. Previous user studies have mainly offered insight into the appropriate selection of VR systems for universal and relatively simple manipulation tasks, such as rotation, navigation and sparse visual search. The study in this chapter extends this work to include several tasks specific to the visualization of dense volumetric data sets.

## 3.3 Empirical Work

### 3.3.1 Experiment Design

For the scientific problems described in section 3.1 , researchers need tools to perform relevant tasks and to provide answers in which they have confidence. Currently available VR interfaces can be classified into five categories:

1. Immersive HMD-based VR system
2. Desktop visual fish tank VR system
3. Multimodal fish tank VR system (visual and haptic)
4. Immersive CAVE
5. Projector-based Workbench

Those VR systems have different navigation and manipulation styles. The navigation style in a HMD based VR system is similar to our actual experiences in the 3D world. It provides a user with fully immersive experience and presence feeling. A user navigates within the data space as if he/she were inside the data set. The data set is also rendered

in human body size. As the study from Bakker [BPW03] indicated, head-slaved navigation had an advantage over indirect navigation for the acquisition of spatial knowledge in a VE, at least for small-scale environments.

A fish tank VR system is different from the traditional 3D desktop system in that it is equipped with both a head-tracking device and a near-field six DOF manipulation device. It is a non-immersive type of VR. The fish tank set-up allows a user to manipulate the data set with the hand and observe it from the outside. No flying through can be done and the viewpoint is often from the outside of a data set. Therefore, the inside structure of the data set is difficult to observe. For a fish tank VR system with haptic force feedback, a user can feel and touch the objects even when they are located inside the data set. Such additional touch modality often increases the presence feeling of a user.

The stereoscopic images are created differently in those system. In a HMD-based VR system stereoscopic images are produced through a time-parallel stereoscopic display method, which means each eye observes a different screen and the optical system directs each eye to the correct view. Shutter glasses are used in a fish tank VR system in combination with a time-multiplexed display, which means that the left and right-eye views of a stereoscopic image are computed and alternatively displayed on the screen. Imperfect separation of both views by the shutter glasses can have a negative effect on depth perception.

The CAVE provides a similar immersive experience as a HMD-based VR does, although the stereoscopic images are generated by different approaches (the CAVE often uses the same shutter glasses as a fish tank VR system has). A projector-based VR workbench is another option for viewing and manipulating 3D scientific data, and is regarded as a partially immersive VR system. These two kinds of systems were not tested in this study because they are less widely used in real applications.

The user study compared three kinds of VR systems: HMD-based VR, fish tank VR, and fish tank VR with haptic feedback. The rendering paradigms were only tested in their most common configurations: inside-out for HMD-based VR and outside-in for fish tank VR. Relative performance of these systems was compared in four generic tasks involving the visualization of volumetric data. The study did not test other potentially useful aspects, such as the possible effect of the auditory modality on data analysis tasks, partly because this modality is less commonly used for such tasks, and partly because resources were limited

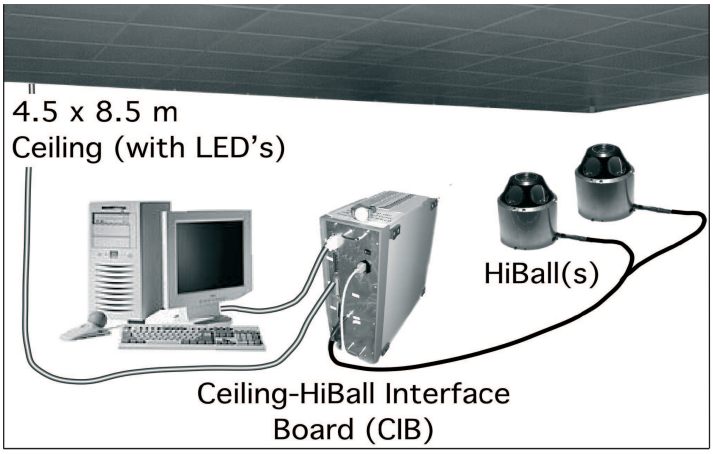



Figure 3.3: The diagram of HiBall tracking system (Photograph reprinted from [WBV+01]).

in term of experimental possibilities.

The author hypothesizes the following results based on the above characteristics of all experimental systems (navigation and manipulation styles, display and etc):

- A HMD-based VR system will have advantages on all data analysis tasks in this study. The user preference will be positive as well.
- A fish tank VR set-up will not help users much with most of the tasks. For a tracking path task, this set-up is particularly difficult to use.
- The additional inside-out feature introduced by haptic force feedback will help a user answer questions regarding object properties, such as size and shape, compared to the case where no haptic feedback is provided. It will also make tracking of a path easier.

### 3.3.2 Apparatus

All three systems display the volumetric data using the Visualization Toolkit (VTK), an open-source library that provides different rendering algorithms (ray-casting, isosurface and 2D texture mapping) [SAH00]. To enable real-time interaction, the author chose Marching

Cubes as the primary algorithm and rendered isosurfaces of the volumetric data. The standard structure of VTK does not provide any mechanism for integration with VR input devices, so the author combined the VTK library with VRPN and UNC's virtual-world library toolkit (Vlib) to enable access to the visualization capabilities of VTK from our VR set-ups.

### Immersive HMD VR system

The immersive VR system uses a V8 HMD from Virtual Research System. Each LCD provides a color VGA pixel resolution of 640×480 at a refresh rate of 60Hz. Head tracking is performed via a 3rdTech HiBall tracking system, a high-performance wide-area optical tracker that incorporates a six DOF sensor. The HMD/head tracking system consists of three main components as shown in Figure 3.3. The outward-looking HiBall sensor is mounted on the back of the HMD (Figure 3.4b). The HiBall observes a subset of fixed-location infrared LEDs embedded in the ceiling. A tracking server coordinates communication and synchronization between the host computer and the HiBall and ceiling LEDs. Tracking data are transmitted through network switched Ethernet from the tracking server to a rendering computer via VRPN. The author used a DELL Precision 530 (dual 2.8-GHz Xeon with 2GB RDRAM) and an NVidia *Quadro FX 1000* graphics card. Two VGA outputs from the graphics card are connected to the LCDs for each eye in the HMD via a video splitter to provide stereo-offset images.

The working space for a user in this VR system is about 4.5 meters wide by 7 meters long by 4 meters tall (15 feet × 23 feet × 13 feet) as shown in Figure 3.4a. A calibration procedure is used to calculate a precise transformation matrix between the sensor and the eyes. An additional hand sensor is also available for hand input, but it was not used during the experiments.

### Fish tank VR

The second VR system is based on the concept of fish tank VR introduced by Colin Ware [WAB93]. The central computing platform of this VR system is identical to the HMD system with the following additional components:

Figure 3.4: HMD based VR system: (a) A user in the immersive VR system; (b) HMD with head tracking sensor.

- A 17-inch CRT monitor with resolution of $1024 \times 768$ and a refresh rate of 100 Hz to support stereo display, an infrared emitter and shutter stereo glasses from Stereo-Graphics.
- A PHANTOM Desktop$^{TM}$ haptic device for precise 6-DOF positioning and high fidelity 3-DOF force feedback output at 1 kHz. In fish tank VR mode, the PHANTOM was used to rotate the volume around its center only (additional operations were available during fish tank VR with haptics, as described below).
- A DynaSight 3D optical tracker for measuring the 3D position of a target (reflective disc) attached to the front of the stereo glasses. When dynamic perspective is combined with stereoscopic viewing, a real-time 3D display appears that provides a virtual window into the computer-generated environment. Dynamic perspective eliminates the perceived image warping associated with static stereoscopic displays. An additional benefit of using the head to tune the perspective is that the hands are free to control the object being visualized, in this case with the PHANTOM.

The hardware components are organized to enable accurate and easy calibration. The tracker's control box is placed above the monitor on a metal plate supported by an arm (Figure 3.5). The arm's height guarantees continuous detection of the tracking and stereo

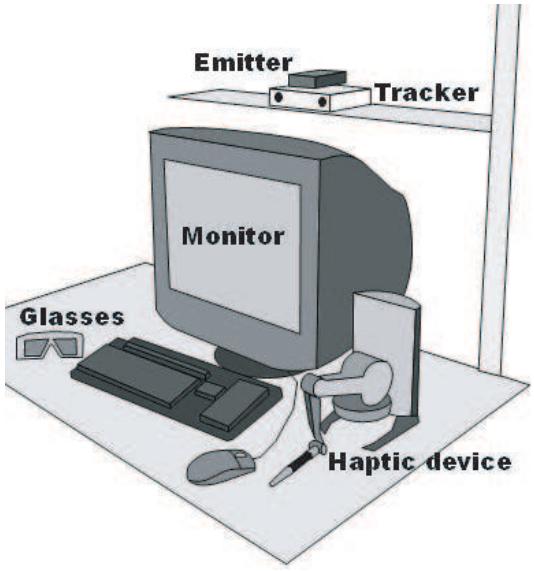



Figure 3.5: A diagram of the fish tank VR system.

signals. A cable between the infrared emitter for the stereo glasses and the control box for the head tracker synchronizes the two devices. The real set-up is shown in Figure 3.6.

### Fish tank with Haptics

Haptic visualization techniques have been developed for force feedback systems such as the PHANTOM. The fish tank VR with haptics prototype uses the same hardware set-up as the fish tank VR system, except that the PHANTOM also provides force feedback, specifically a single point of haptic response, which is sufficient for our tasks. Although the stylus where force is applied is not visually located within the display volume (as compared to the Visual Haptic Workbench or the *ReachIn* systems), no users complained about the cognitive effort required to move the hand in one location while viewing another. An axis-aligned on-screen icon follows the stylus's motion in 3D, producing an effect similar to using a mouse to control the on-screen cursor. The haptic representation of volumetric data employed different force models for different objects within the volume: viewers felt the outside surface of spheres and ellipsoids, but the inside of long curved tubes and cylinders.

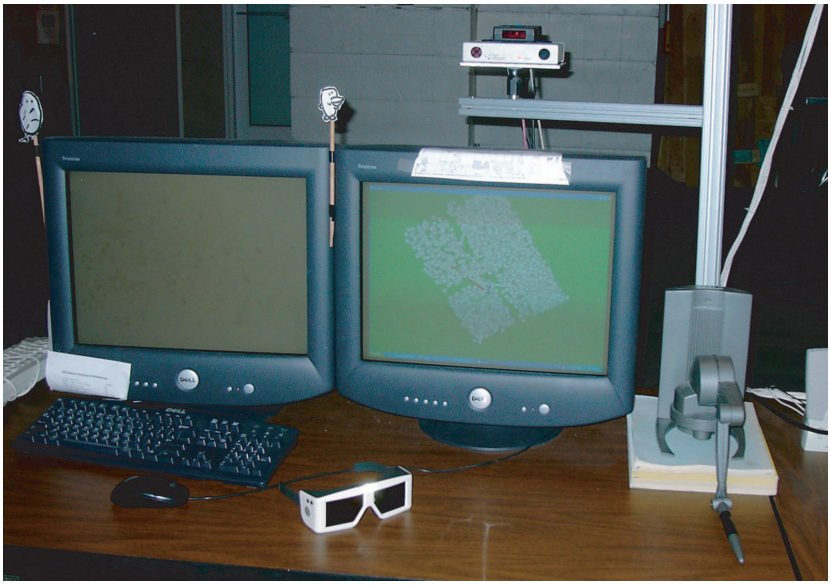


Figure 3.6: The snap shot of the fish tank VR system.

### 3.3.3 Data and Task

Simulated volumetric data are generated to act as trials during the study. A random number of two to four types of differently-shaped objects (sphere, ellipsoid, cylinder, and curved tube) are inserted at random positions (Figure 3.7). These objects may overlap with each other to become connected. The objects' properties (size, shape) and the density of each volume form experimental conditions that vary between trials. The bounding box of the volume is subdivided into eight equally-sized regions (a $2 \times 2 \times 2$ array in the $x$, $y$, and $z$ directions) within which object density may differ. Regions are labeled with unique numbers (1 through 8) to enable subjects to describe the paths of curved tubes within a volume.

There are always spheres and at least one curved tube within every volume. Trials may also contain ellipsoids, cylinders, and up to two additional curved tubes. Sphere size may vary between four possible radii ranging from six to twelve OpenGL units. The density of objects within each region is controlled to be sparse, medium, or dense. A single densest region exists within each volume. Sparse regions contain between 10%-60% of the number of objects of the densest region, while medium regions contain between 60%-90% of this number.

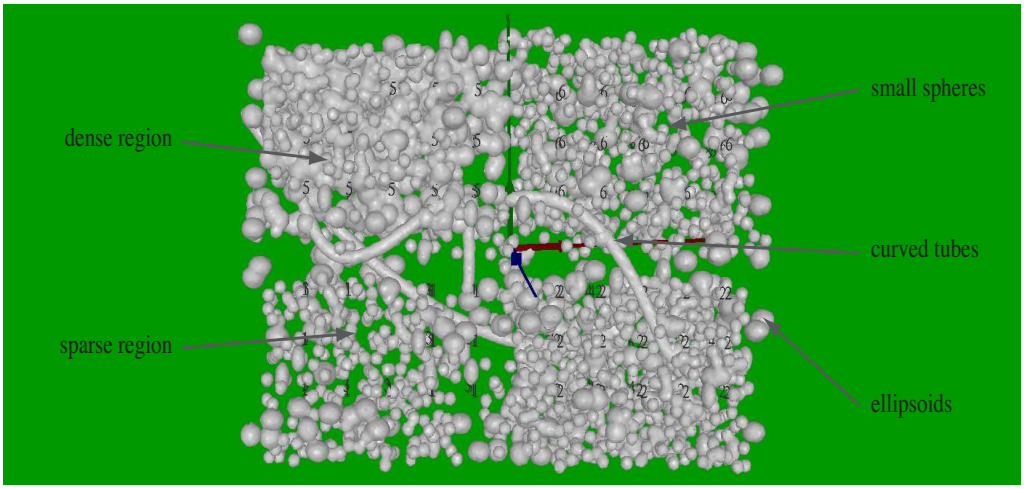



Figure 3.7: An example trial from the experiment, showing a top-down view on a simulated volume with different experiment conditions like shape, size, density, and connectivity highlighted.

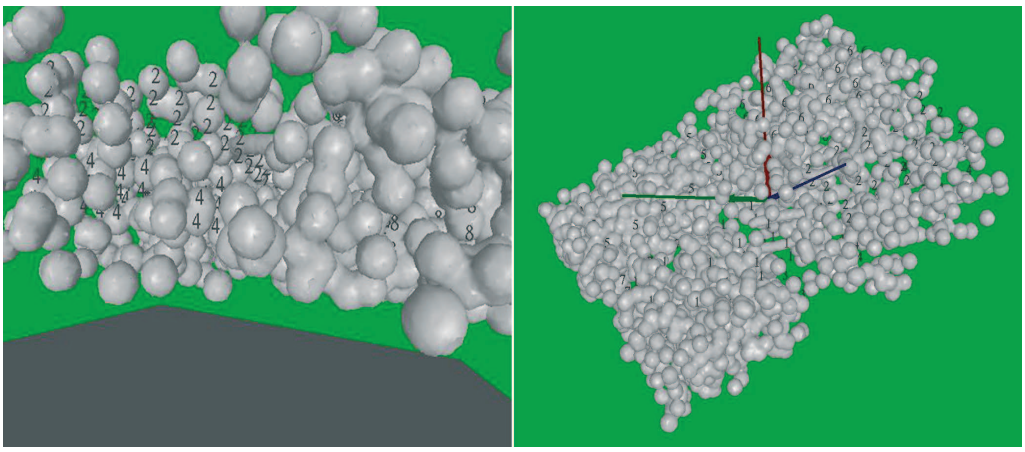


Figure 3.8: Two views of a volumetric data set from an example trial, as seen in the HMD system on the left, and as seen in the fish tank and fish tank with haptics systems on the right.

Subjects were asked to complete fours tasks within each trial. Each task involved judging the properties of a specific object or of the overall volume, specifically:

- *Shape task:* Subjects identified the number of differently-shaped objects within the volume and name the objects.
- *Size task:* Subjects reported how many differently-sized spheres exist.
- *Density task:* Subjects identified the densest region in the volume.
- *Connectivity task:* Subjects reported how many curved tubes exist in the volume (numerosity question), and then determined which region(s) the longest curved tube passes through (spatial region question). For example, Figure 3.7 shows two curved tubes.

Subjects were asked to give their answers as accurately as possible and to minimize response time. The size, density, and curve counting questions were presented in a multiple choice format. Subjects were asked to describe the name of each kind of object for the shape question and all the region numbers for the tube tracking question.

### 3.3.4 Experimental Procedure

A between-subject design was used, with VR system type as an independent variable: HMD-based VR, fish tank VR, and fish tank VR with haptics (Table 3.1). Subjects were randomly assigned into one of three groups. The HMD group wore the HMD and walked around within the tracked environment to observe the volumetric data as seen in Figure 3.8 left. The fish tank group used the fish tank VR system and wore stereo shutter glasses to interact with volumetric data through the stylus of the PHANTOM as seen in Figure 3.8 right. Although the stylus was tracked and displayed as an icon on the monitor, no force feedback was provided to this group. The haptics group added force feedback to the basic fish tank VR system.

Subjects completed several steps during the experiment. As part of an initial interview session, they signed a consent form, answered basic demographic questions (age, gender, and occupation or major field of study), and identified their frequency of computer use and prior

Table 3.1: Variables lists

| **Independent Variable** | **Dependent Variables** |
|---|---|
| VR user interfaces<br>(HMD based VR,<br>fish tank VR,<br>fish tank VR with haptics) | Time,<br>accuracy of the tasks<br>(shape, size, density,<br>connectivity) |

experience with any kind of VR system (see Appendix C). A training session introduced the equipment and described the tasks to be performed. Next, the formal experiment session was conducted. Each experiment included 20 trials, with each trial containing a single volumetric data set. These twenty data sets were different from one another, and varied by object property (type, size, position, and density). However, the same set of trials (20 data sets) in the same order were used for all three groups (HMD, fish tank, and fish tank with haptics).

Two dependent variables, the total time taken to respond to a trial and the subject's answers all four tasks within a trial, were recorded by the experimenter. A short break was provided every half hour or whenever a subject asked for one. After completing the last trial in the formal experiment session, subjects filled out a questionnaire describing their opinions about the system, any suggestions they had on how to improve the system, and so on (see Appendix C). The study ended with a short debriefing during which the experimenter summarized the study goals. The subjects were paid $9 for their participation.

## 3.4 Results

Forty subjects volunteered for this experiment, 33 males and 7 females. The subjects were randomly assigned into one of the three display system groups: 14 subjects (12 males and 2 females) for the HMD group, 13 subjects (11 males and 2 females) for the fish tank group, and 13 subjects (10 males and 3 females) for the haptic group.

The age of each subject and the frequency of computer use were recorded before the

experiment began. The measurement of computer experience questions used a standard seven point rating scale. Average ages and frequencies of computer use were 23.2, 23, and 23.7, and 6.3, 6.0, and 5.6 for the HMD, fish tank and haptic groups, respectively. These data suggested we had similar user population within each group.

### 3.4.1 Summary

Two types of measures of performance were derived for each trial a subject completed: total response time $rt$ and error rates $\widehat{P_e}$ on four tasks. A single $rt$ value representing the total time in seconds needed to complete all four tasks was captured for each trial. The author did not record the individual $rt$ for each task since it was difficult to record separately. Four separate $\widehat{P_e}$ values for the four tasks subjects completed were also generated.

- For the shape, size, density and numerosity questions, subjects' answers were coded as 1 for correct and 0 for incorrect. Then error rate $\widehat{P_e}$ is defined as the proportion of wrong answers among all the answers. Because the answer sheet for each task is presented in a multiple choice format, the answer from subjects is regarded as categorial data.
- For the spatial region question, subjects answers were coded as two parameters: the *false negative* and the *false positive* used in a Receiver Operating Characteristic curve (ROC). Further analyses are based on these two parameters.

For $rt$ statistics, trials were divided by VR system (HMD, fish tank, or fish tank with haptics). For $\widehat{P_e}$ statistics, trials were divided by VR system (HMD, fish tank, or fish tank with haptics) and task (shape, size, density, or connectivity). At times, more in-depth analyses on the data were performed when results obviously depended on other task parameters, such as in the case of counting sphere sizes, where performance obviously depended on the number of sizes present. The shifts across conditions on average values of the logarithm of $rt$ were studied using Analysis of Variance ($ANOVA$). A discussion on why analysis of lg(rt) should be preferred over analysis of $rt$ itself can be found in a recent publication [MLK07]. The differences in error rates $\widehat{P_e}$ were studied using Fisher's exact test statistics [FLP03].

In summary, the following significant differences in performance were identified:

1. The HMD group had longer $rt$, compared to both the fish tank and the haptic groups. Using haptics in the fish tank VR system also resulted in longer $rt$.

2. For most tasks (counting number of different shapes and number of curved tubes and finding the densest region), the HMD group had higher $\widehat{P_e}$, compared to both the fish tank and haptic groups.

3. In counting number of different sizes of sphere objects, none of the three groups was very accurate. The HMD group made significantly more errors than the fish tank and the haptics group in case only one size of sphere was present (Figure 3.12). In case more than one size was present, subjects from all three groups mainly underestimated the number of sizes (Figure 3.13).

4. For identifying the regions that the longest curved tube passes through during the connectivity task, the HMD group had higher chances in both missing the right regions and misjudging the wrong regions, compared to both the fish tank and the haptic groups.

In addition to statistical results, a number of anecdotal findings were made, pointing to: (1) the desire for an overview display in the HMD system; (2) the desire for immersion in the fish tank VR systems; (3) fatigue in the HMD system; and (4) the preference for including touch in the haptic system.

### 3.4.2 Detailed Analysis of Quantitative Results

#### Performance times

The response time $rt$ needed to complete all four tasks during a trial was recorded during the formal experiment session. Subjects in the HMD group had significantly longer $rt$ compared to the fish tank and the haptic groups. The $ANOVA$ for the logarithm of $rt$ was significant, $F(2, 165) = 40.058, p < 0.001$ (Figure 3.9). Post-hoc paired comparisons showed that the fish tank group was significantly faster than the haptic group ($p < 0.001$). Overall, the HMD group spent 43% more time compared to the fish tank group. And the haptic group spent 23% more time compared to the fish tank group. Because of the high $rt$ for the

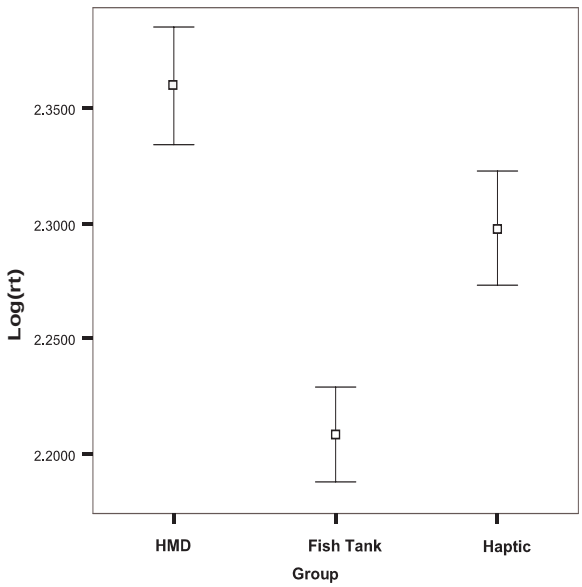



Figure 3.9: ANOVA of lg($rt$) for the different experiment conditions, all results are divided by VR system (HMD, fish tank, fish tank with haptics), error bars represent the 95% confidence interval.

HMD group, the author was forced to reduce the total number of trials for this system to 16. Because each trial tests all four tasks, this did not unbalance the experiment to favor certain conditions. Although subjects in the other two groups were able to finish all 20 trials within reasonable time, to maintain consistency the author analyzed only the first 16 trials completed by each group.

Table 3.2: Fisher's Exact Test for the density task

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | **0.00** | - |
| **Haptic** | **0.00** | 0.318 |

### Accuracy in the density task

For the density task, the answers for every combination of two groups are compared through a Fisher's exact Test to find out whether or not there is an association between the error rate

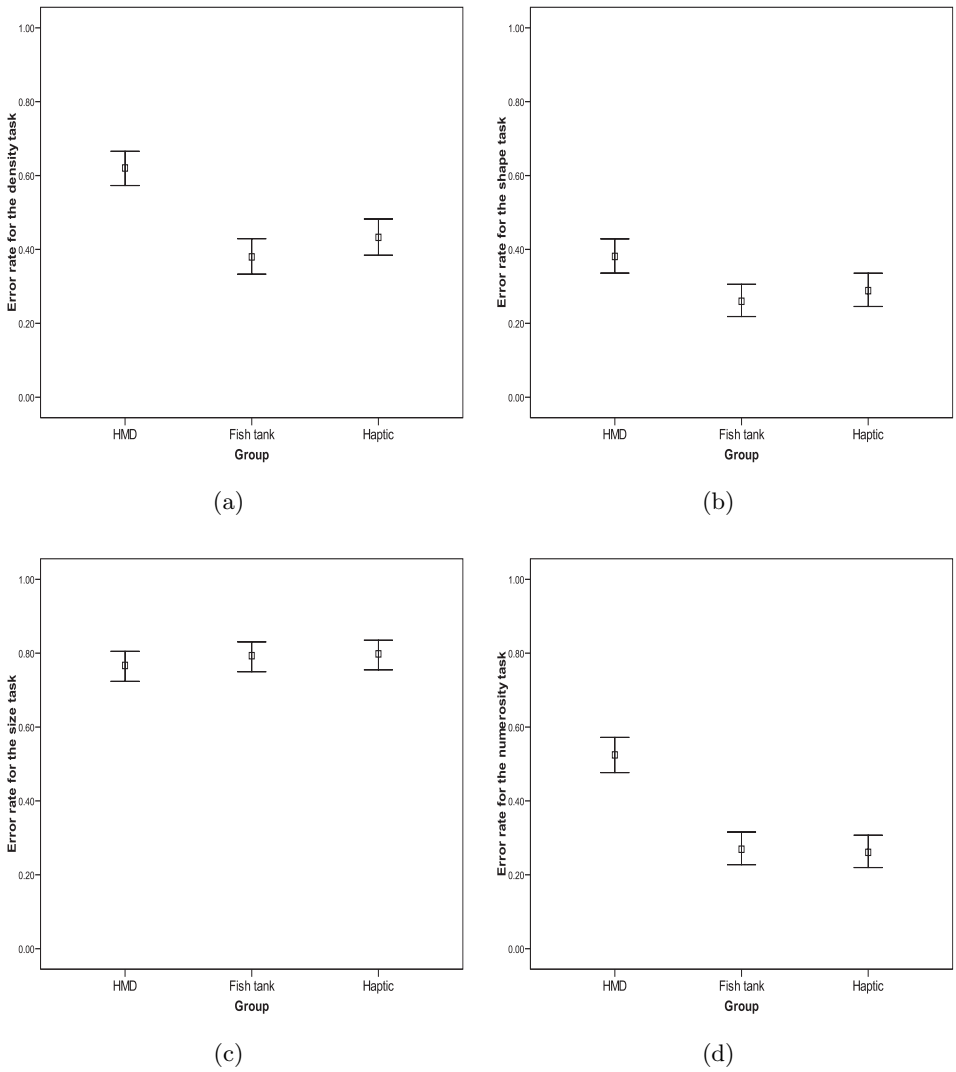



Figure 3.10: $\widehat{P_e}$ values for the different experiment conditions, all results are divided by VR system (HMD, fish tank, fish tank with haptics), error bars represent the 95% confidence interval: (a) $\widehat{P_e}$ for the density task; (b) $\widehat{P_e}$ for the shape task; (c) $\widehat{P_e}$ for the size task; (d) $\widehat{P_e}$ for counting the number of curved tubes in the connectivity task.

of finding the densest region within a volume and the choice between two VR systems. The results are shown in Table 3.2 (**The significant results where $p < 0.05$ are displayed in boldface, a convention that will also be used in the following tables**) and summarized as follows.

The users in the HMD group produced significantly more errors than the users in both fish tank groups, while there was no significant difference between the two latter groups. In absolute terms, none of the three groups had very high accuracy, with $\widehat{P_e} = 0.62, 0.38$ and 0.43 for the HMD, fish tank, and haptic groups, respectively (Figure 3.10a).

Table 3.3: Results of Fisher's Exact Test for the shape task

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | **0.007** | - |
| **Haptic** | **0.042** | 0.583 |

### Accuracy in the shape task

The results of the Fisher's exact analysis for the shape task are shown in Table 3.3 and the conclusions as to the relative performance of all three systems are identical as in the case of the density task, i.e., the HMD group was performing significantly worse than both fish tank groups.

In absolute terms, all three groups had reasonable accuracy, with $\widehat{P_e} = 0.38, 0.26$ and 0.29 for the HMD, fish tank, and haptic groups, respectively (Figure 3.10b). Further analysis indicates that the user performances of the three groups depended on the experimental condition in terms of the number of shapes (Figure 3.11). When there were only two (sphere and curved tube) or all four kinds of shapes, the haptic group was more accurate than the other two groups (significant difference between HMD and haptic for the four shapes situation). The fish tank group was the most accurate when three kinds of shapes were presented (sphere, ellipsoid and curved tube or sphere, ellipsoid and cylinder or sphere, curved tube and cylinder). Irrespective of the number of the shapes, the error rate was always the highest for the HMD group.

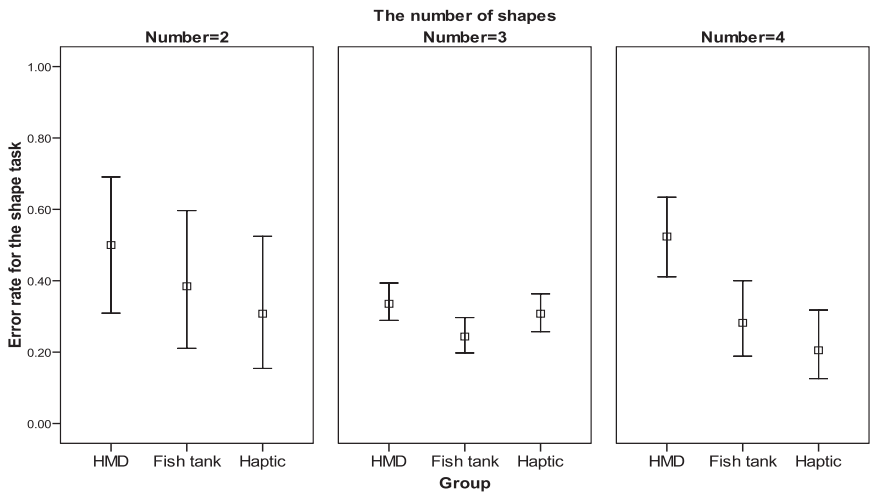



Figure 3.11: $\widehat{P_e}$ for the shape task based on the number of shapes.

### Accuracy in the size task

The results of the Fisher's exact analysis for the size task are shown in Table 3.4. No significant differences could be observed between the performances in the three groups.

Table 3.4: Results of Fisher's Exact Test for the size task in terms of overall error rate

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | 0.561 | - |
| **Haptic** | 0.484 | 1.0 |

In absolute terms, none of the three groups was accurate, with $\widehat{P_e} = 0.76, 0.79$ and 0.80 for the HMD, fish tank, and haptic groups, respectively. The error rates were all above 70%, although fewer errors were made in the HMD group (Figure 3.10c). Further analysis based on the task condition in terms of the number of sizes shows that performance differences between systems varied (Figure 3.12). When there was only one size or two sizes of spheres, the haptic group was more accurate than the other two groups (although only the difference in case of one size between the HMD group and the haptics group was statistically significant). When there were three or four sphere sizes, the HMD group was somewhat more accurate than the other two groups (although this was only statistically

significant in the case of four sizes). The only case where the error rate was below 50% (for all three groups) was when there was only one size of sphere present. The chance of estimating the number of sizes correctly was even lower than guessing in case of the three or four sized condition, which indicates that subjects significantly underestimate the number of different sizes in these latter cases.

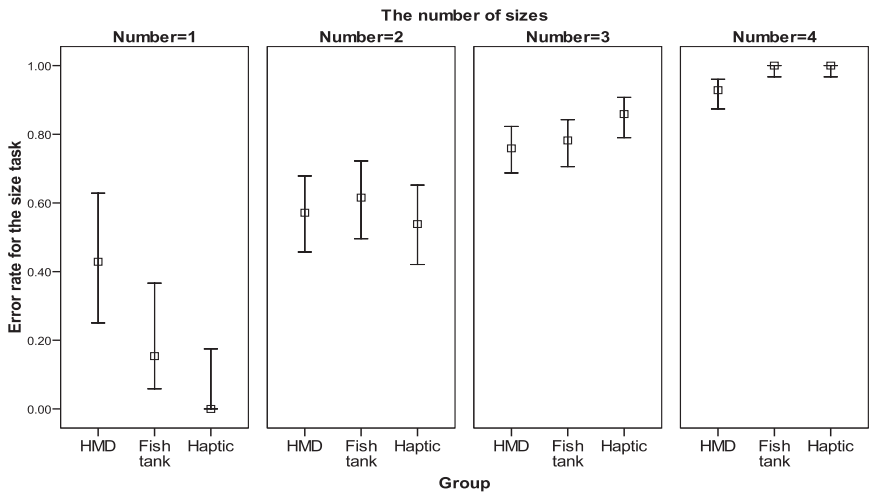



Figure 3.12: $\widehat{P_e}$ for the size task based on the number of sizes.

In the previous analysis, user performance is only based on the error rate, that is, the proportion of completely wrong answers among all the answers. When performing the size task, three cases can arise. The number of sizes can be estimated correctly, over-estimated or underestimated. The resulting Fisher's exact Test analyses, based on three instead of two (right or wrong) categories, are reported in Table 3.5.

This more refined analysis reveals a significant difference between the HMD condition and the two fish tank conditions. In absolute terms, the proportion of underestimation was above 65% in all cases, although the proportion was lower during the HMD trials, with 0.68, 0.77 and 0.79 for the HMD, fish tank, and haptic groups, respectively. This reflects the fact that mistakes mainly originate from an underestimation of the number of different sizes for sphere objects, which is illustrated graphically in Figure 3.13.

Table 3.5: Results of Fisher's Exact Test for the size task in terms of estimation difference

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | **0.013** | - |
| **Haptic** | **0.000** | 0.599 |

### Accuracy in the connectivity task

In the connectivity task, subjects answered two questions: the total number of curved tubes in a volume (numberosity question), and which regions of the volume the longest tube passed through (spatial region question). The results of the accuracy analysis in case of the numerosity question are reported in Table 3.6. Similarly as in the above tasks, the HMD condition differed significantly in terms of accuracy from the two fish tank conditions, while the two latter conditions performed similarly.

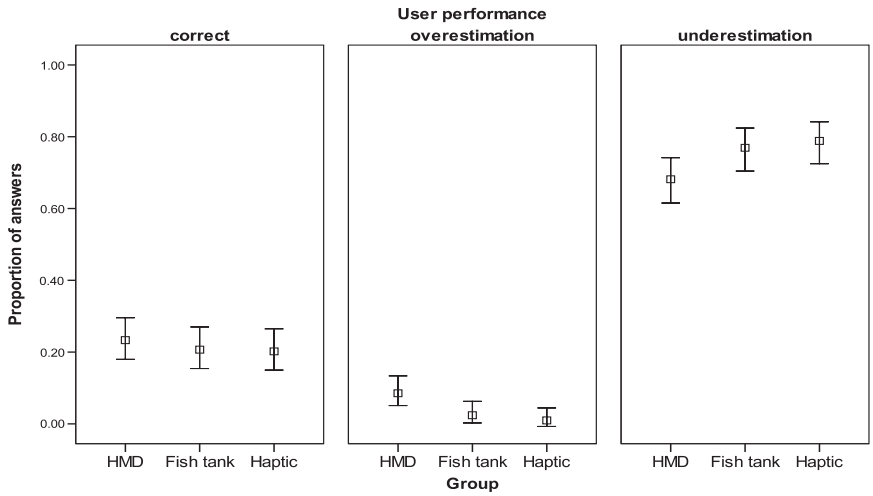



Figure 3.13: $\widehat{P_e}$ for the size task based on the estimation difference.

In absolute terms, the haptic group was somewhat more accurate than the fish tank group, $\widehat{P_e} = 0.52, 0.27$ and 0.26 for the HMD, fish tank, and haptic groups, respectively (Figure 3.10d). Further analysis based on the task condition, i.e., the number of curved tubes present, provides more insight into the performance differences (Figure 3.14). When there are two or three tubes, both the fish tank and haptic groups were significantly more

Table 3.6: Results of Fisher's Exact Test for counting the total number of curved tubes

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | **0.000** | - |
| **Haptic** | **0.000** | 0.912 |

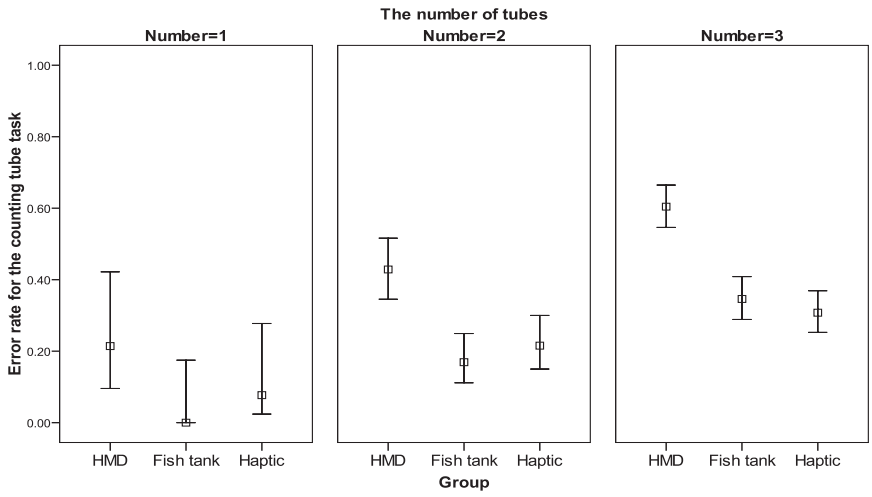



Figure 3.14: $\widehat{P_e}$ for the counting task based on the number of curved tubes.

accurate than the HMD group. The users in the HMD group obviously had more problems in distinguishing whether or not the perceived tubes were connected. This was a priori unexpected, since these users were able to observe the data from the inside.

For the spatial region question, the answers of all three groups are analyzed based on the Receiver Operating Characteristic Curve (ROC). ROC is a graphical plot of the *sensitivity* versus *1-specificity* for a binary classifier system. The ROC can also be represented equivalently by plotting the fraction of *true positive* (TP) versus the fraction of *false positive* (FP). In the context of this user study, there are four possible combinations of whether or not a region is passed through by the longest curved tube in a trial, on the one hand, and a subject's answer based on his/her own judgment, on the other hand (see in Table 3.7). A *true positive* situation is that the longest curved tube in a trial passes through a region, and a subject does recognize this fact correctly. A *false positive* situation is that the longest curved tube in a trial does not actually passes through one region, but a subject thinks it

Table 3.7: Four situations in judging whether the longest tube passes through a region

| | **PASS (REAL SITUATION)** | **NO PASS (REAL SITUATION)** |
|---|---|---|
| **pass (subjects's answer)** | *true positive* (TP) | *false positive* (FP) |
| **no pass (subjects's answer)** | *false negative* (FN) | *true negative* (TN) |
| **sum** | *number_of_pass* (NP=TP+FN) | *number_of_empty* (NE=FP+TN) |

does by mistake. The other two situations can be described similarly. From these situations, the author derives several statistics to describe the user performance. The error rate $\widehat{P_e}$ consists of two parameters $\widehat{P_{FN}}$ and $\widehat{P_{FP}}$. The fraction of false negative $\widehat{P_{FN}}$ (also known as the chance of missing $\widehat{P_M}$) is defined as $\widehat{P_{FN}} = FN \div NP$, while the fraction of false positive $\widehat{P_{FP}}$ (also known as the fraction of false alarm ) is defined as $\widehat{P_{FP}} = FP \div NE$. Both probabilities can be analyzed as a function of the experimental conditions using Fisher's exact test statistics.

Table 3.8: Results of Fisher's Exact Test for $\widehat{P_{FN}}$

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | **0.001** | - |
| **Haptic** | 0.921 | **0.001** |

The analyses for the *false negative* are summarized in Table 3.8. The fish tank system performed significantly different from the other two systems. In absolute term, the fish tank group was more accurate than the other two groups, with $\widehat{P_{FN}} = 0.39$, 0.31, 0.39 for the HMD, fish tank, and haptic groups, respectively. This is reflected in Figure 3.15.

The analyses for the *false positives* are summarized in Table 3.9. In this case, only the difference between the HMD condition and the fish tank condition without haptics was shown to be statistically significant. In absolute terms, $\widehat{P_{FP}}$ is equal to 0.20, 0.17 and 0.17 for the HMD, fish tank, and haptic groups, respectively. This is reflected in Figure 3.15 as

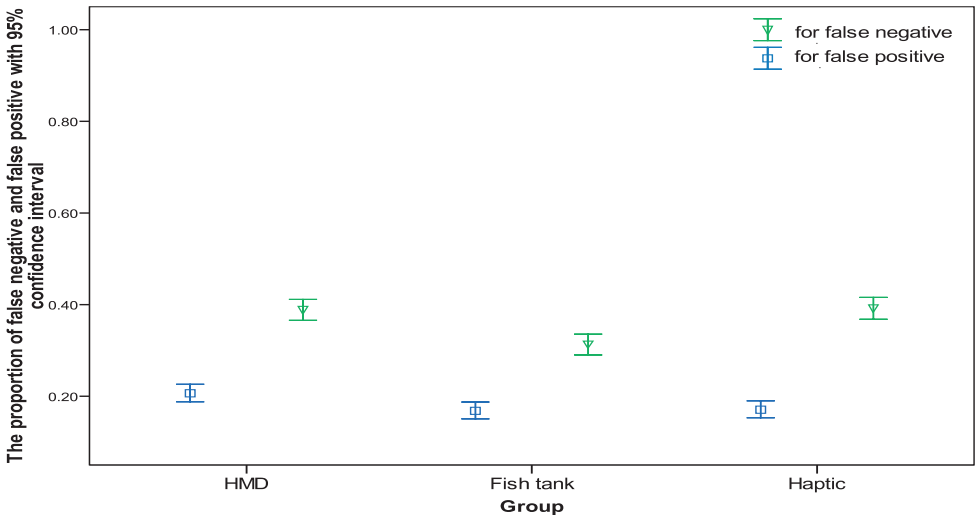



Figure 3.15: $\widehat{P_{FN}}$ and $\widehat{P_{FP}}$ for the different experiment conditions in locating the longest curved tube during the connectivity task, all results are divided by VR system (HMD, fish tank, fish tank with haptics), error bars represent 95% confidence interval.

well.

Table 3.9: Results of Fisher's Exact Test for $\widehat{P_{FP}}$

| | **HMD** | **Fish tank** |
|---|---|---|
| **Fish tank** | **0.049** | - |
| **Haptic** | 0.065 | 0.948 |

The overall performance can be quantified by attributing costs to both the *false negative* $\widehat{P_{FN}}$ ($\widehat{P_M}$ misses) and the *false positive* $\widehat{P_{FP}}$ (false alarm). In summary, the fish tank (without haptics) group seems to have the best performance. The HMD group was least accurate in finding all the regions the longest curved tube passes through and had highest probability of misjudging some regions as those that the longest curved tube passes through. The haptic group had intermediate performance, with almost the same frequency of *false negative* as the HMD group but lower frequency of *false positive*.

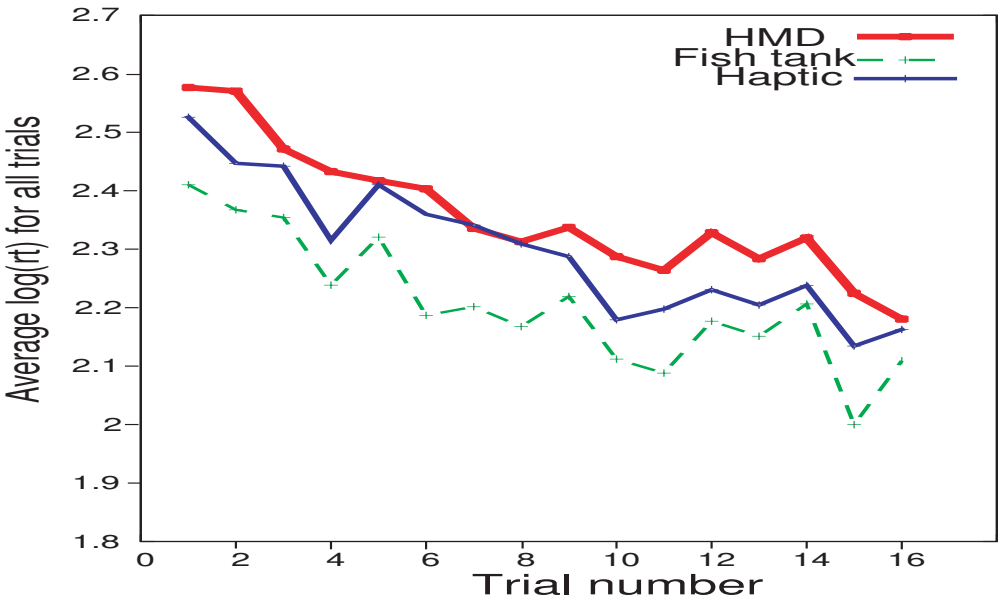



Figure 3.16: Time curves for each VR system, in the order that subjects completed the trials.

### 3.4.3 Interpretation of Results

The time needed to complete tasks in the HMD system was significantly longer, compared to both the fish tank and the fish tank with haptics systems. One explanation is that the HMD system requires subjects to walk around within the tracking space, which takes more time to explore compared to moving hands and head in the fish tank systems. Another critical issue was the reported inability of HMD subjects to remember where they had previously seen target items within the volume because of the high density of data sets. They would often have to re-search the volume for objects they had previous located, but had "lost"as they walked into a different region. Finally, subjects may simply be more familiar with a standard desktop system.

The fish tank group was also significantly faster than the haptic group. When touch was available, subjects often spent more time "feeling" inside the volume to confirm their decisions, even when a correct answer could be derived from visual evidence alone.

A curve of the time spent on each trial indicates a similar learning effect for all three

groups. Time decreases as the subjects complete more tasks (Figure 3.16). The first five trials showed the strongest learning tendency. After the first five trials, time spent on each trials still varied mainly due to the varying difficulty in each trial, which was caused largely by the density of the data set for that trial. The learning effect did not affect the credibility to draw conclusions because the three groups shared the same learning pattern. For users who have become familiar with the task and equipment, it is believed that performance will stabilize to times similar to those seen during the later trials.

The HMD group was significantly less accurate than the fish tank and haptic groups in the shape task. Error results showed that subjects from all three groups found it relatively easy to identify the sphere object. The HMD group made more mistakes in identifying cylinders than the other two groups. It was difficult to judge the number of shapes for the HMD group when all four shapes existed. The accuracy of the haptic group consistently increased as the number of shapes increased, and got the best performance among the three groups when all four shapes of objects were present. It indicates that touch does help the subjects in the shape task. With only vision, the HMD and the fish tank groups did not show similar behavior (Figure 3.11). Finally, subjects from all three groups sometimes misjudged a curved tube as a cylinder, which is indicated from the feedback of the subjects.

Although there were no significant differences in accuracy in the size task, absolute performance was poor across all three groups. Not surprisingly, when there was only one size of sphere the responses were quite accurate ($\widehat{P_e} = 0.43, 0.15$, and 0.0 for HMD, fish tank, and haptic, respectively). When two sizes of sphere with a large difference in radii were presented, the subjects from all groups also did well. The haptic group performed best in case of one or two sizes. However, when the radii difference between the two spheres was small, or when there were three or four different sizes of sphere, all subjects had difficulty determining how many different sizes they saw ($\widehat{P_e} = 0.76, 0.78$, and 0.96 with three sizes of sphere, and $\widehat{P_e} = 0.93, 1.0$, and 1.0 with four sizes of sphere for HMD, fish tank, and haptic, respectively as shown in Figure 3.12). The average accuracy of each group was even lower than the probability of guessing for three or four sizes. However, the HMD system tended to help the subjects get the best performance among the three groups when there were more than two sizes. This suggests that: (1) some of the radii differences are too small to be easily distinguished by the visual system; (2) touch does not help much in the size task with more than two sizes; and (3) asking subjects to compare between more than two

objects (e.g. three or more differently sized spheres) may lead to lower accuracy.

For the density task, the HMD subjects were significantly less accurate than the fish tank and haptic subjects. None of the three groups had high accuracy, however. The reason might be both the characteristics of high density and slight density differences between adjacent regions. There was no significant difference between the fish tank and haptic groups, implying that haptic feedback did not help with finding spatial regions with different density of objects, particularly for high density situations. The errors for all three groups were spread out across the trials, and showed no learning effects. This suggests that identifying regions of varying density (especially small differences) within a 3D volume is a difficult task that none of these three display systems supports sufficiently.

For the connectivity task, subjects were asked to count the number of curved tubes in a volume, then locate the longest curved tube and identify which regions of the volume it passed through.

1. For the numerosity question, the HMD subjects were significantly less accurate than the fish tank and haptic subjects. The situation where only one curved tube existed did not present a big challenge to the subjects. When more than one curved tube was present, the average accuracy for all three groups decreased as the number of curved tubes increased (Figure 3.14). The subjects in the fish tank group got the best performance when only one or two curved tubes existed. However, the ability to "feel" along the inside of the tubes (both the curved and straight tube) helped the haptic group provide slightly more accurate counts of the number of curved tubes contained within a volume when there were three or more of them. It indicates that touch becomes more useful. The lack of an overview of the volume for the HMD group and the absence of clear complete view of the path of every tube for all three groups created major difficulties for all subjects to judge whether or not different segments they saw belonged to the same tube, no matter how many tubes existed.

2. For the connectivity task's spatial region question, the HMD and haptic groups was significantly less accurate than the fish tank group in finding the right regions that the longest curved tube passes through. The HMD group was significantly less accurate than the fish tank group in identifying wrong regions as the correct ones that the longest curved tube passes through. Task analysis for this question indicates that

subjects first have to identify which curved tube is the longest by comparing the lengths of all curved tubes, then determine which regions contain this target tube. If the wrong curved tube is identified as the longest, the final answer will also be wrong (there are possibilities that the wrong tube may cross the same regions as the longest one does). But for all the trials in this study, there were no two curved tubes that cross exactly all the same regions. In the HMD system, subjects often misjudged a curved tube to be the longest one because of the wrong judgement of the total number of curved tubes and the difficulty of measuring the length of each curved tube precisely. For the fish tank and haptic systems, when the length differences among the curved tubes were large, haptic feedback helped subjects locate the longest tube by touch. They could then correctly identify the regions containing the tube. When the length differences were small, however, the haptic system showed no better performance than the fish tank system. This indicates that, in these cases, the inclusion of touch can not help to identify the length difference and sometimes intervenes with a subject's judgement. Visual feedback acts as the main determinant in locating the longest tube. This explains the slightly different cost of *false positive* between the fish tank and haptic systems. Our results match the findings of Ernst and Banks [EB02]: when visual and haptic feedback are present and haptic feedback can add definite assistance for a task or judgment, it will be used. Otherwise, visual feedback is still the dominant sensory input.

In addition to statistical results, a number of anecdotal findings were made, pointing to: (1) the desire for an overview display in the HMD system; (2) the desire for immersion in the fish tank VR systems; (3) fatigue in the HMD system; and (4) the preference for including touch in the haptic system.

Several HMD users spontaneously suggested adding the ability to see a high-level overview (which might be provided through a button press, Mine's head-butt zoom, or a worlds-in-miniature interface). One casual user was tall enough that he could stand above the data, enabling him to get an overview in the HMD-based VR system, which he reported to be useful. This matches the author's later analysis as well as issues related to the effects of memory on subjects' results.

Some subjects in the fish tank and haptic groups wanted to zoom in and see the volume

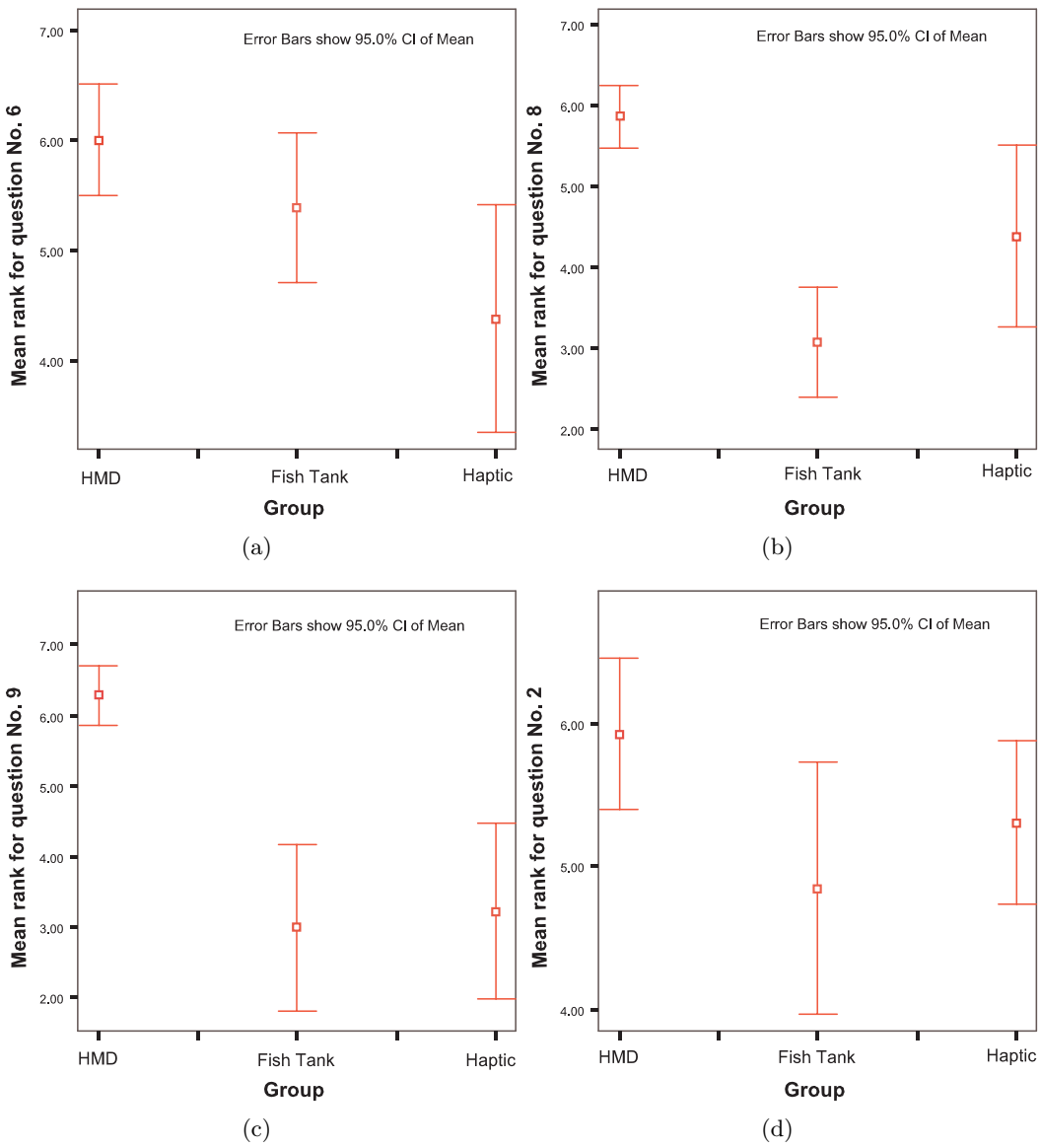



Figure 3.17: Mean values for the different questions regarding the perception of a VR system (see Appendix C), all results are divided by VR system (HMD, fish tank, fish tank with haptics), error bars represent 95% confidence interval: (a) mean rank for the presence question; (b) mean rank for the question of acting inside VR space; (c) mean rank for the question of the degree of surrounding the subject in a VR system; (d) mean rank for the immersion question.

from the inside (some tried to do this by moving their head near the screen). The author concludes that both overview and immersion are helpful for performing the tasks in the experiment. Anecdotal and formal results indicate that a system designed for the study of dense volumes should include both capabilities.

Most subjects said that the HMD and haptic systems were "cool" or "neat" upon initial exposure. Several subjects mentioned without being asked that they liked the HMD VR system or the haptic system. However, subjects in the HMD group requested more breaks after five trials and sometimes asked "How many trials do I still have?" after around ten trials, indicating heavy workload and the dissatisfaction with the system. The author believes this is due to physical or mental fatigue. The increased number of breaks requested did not happen in the fish tank or the haptic cases.

### 3.4.4 Subjective Results

Subjective measurements were obtained through analysis of the post-experiment questionnaires (see Appendix C). Most questions used a standard seven point rating scale (some used a five point rating scale). The answers indicated that overall, subjects preferred the haptic and HMD VR systems due to perceived ease of use, presence, and immersion. The author summarizes the findings over the following categories of questions asked.

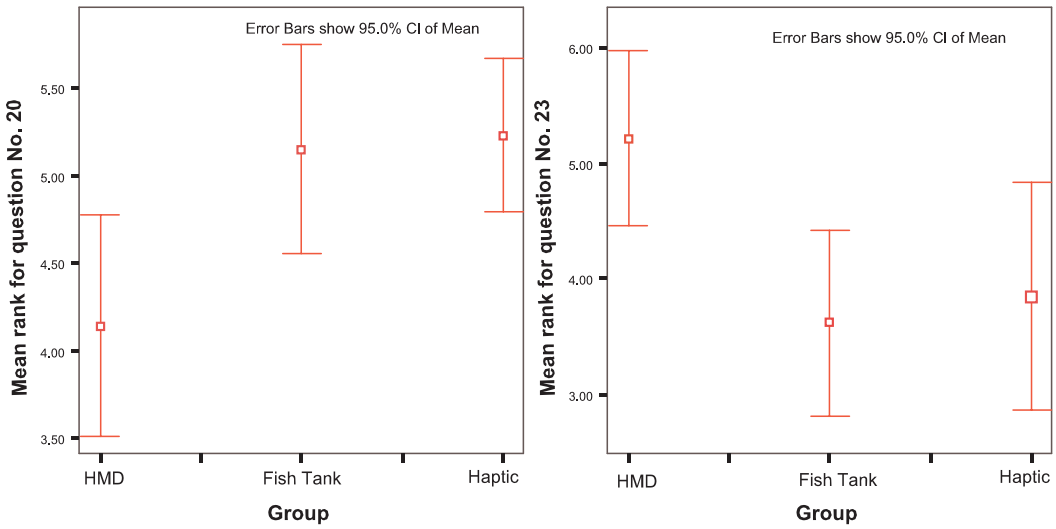
**Overall perception of the VR systems.** The first category of questions targeted the perceived properties and characteristics of a VR system, including the immersion, presence, depth, and spatial relationships (question 1 to question 12). For the question: "the extent that you felt you were within a VE", the HMD system ranked significantly higher than the fish tank and haptics systems, $F(2, 37) = 5.481, p = 0.008$, with a post-hoc comparison between HMD and haptic of $p = 0.006$, and absolute rankings of 6.0, 5.4, and 4.4 for HMD, fish tank, and haptic, respectively (Figure 3.17a). There was also a significant difference on the question: "the extent you had a sense of acting in the virtual space, rather than operating something from outside". The HMD system ranked significantly higher than the other two systems, $F(2, 37) = 15.666, p = 0.001$, with scores of 5.9, 3.1, and 4.4 for HMD, fish tank, and haptic, respectively (Figure 3.17b). Further post-hoc comparison showed the fish tank with haptics system ranked significantly higher than fish tank alone due to the existence of touch ($p = 0.03$), indicating that haptic feedback does add an inside-out property to a

fish tank display. For the question: "the extent you felt that the VE surrounded you", the HMD group again ranked higher than the other two groups, $F(2, 37) = 16.464, p = 0.001$, with scores of 6.3, 3.0, and 3.2 for HMD, fish tank, and haptics, respectively (Figure 3.17c). This suggests that HMD subjects felt more strongly that they were acting within a VE. The author found no significant differences on the questions: "a sense of being there", "a sense of immersion", "difficulty of understanding the spatial relationships", "the quality of multiple view points", or "the quality of depth cues", although the HMD system did rank slightly higher in absolute terms in the immersion (Figure 3.17d), presence, multiple viewpoints and depth cues questions.

**Usability of a VR system.** The ease of learning and using a VR system is the main focus of this category (question 13 to question 25). Answers to the question: "how much consistency did you experience in the VR system compared with a real world experience" were similar for subjects from each group, indicating the act of moving from place to place was judged to be relatively natural and easy. There were no obvious differences on the question about system delay, although HMD subjects reported a slightly shorter perceived delay. No subject from any group complained about the resolution, frame rate or delay; these parameters did not seem to bother them. The haptic system ranked higher than the other two systems for identifying the shape and location of individual objects, and the global topology of the volume data sets. Although subjects from all three groups felt their system was easy to use, the HMD group ranked highest for the perceived difficulty in carrying out their tasks. Moreover, HMD subjects reported a significantly higher demand for memorizing than the other two groups, $F(2, 37) = 5.534, p = 0.008$, with scores of 5.2, 3.6, and 3.7 for HMD, fish tank, and haptics, respectively (Figure 3.18a). Finally, HMD subjects were less confident about the accuracy of their answers, $F(2, 37) = 5.521, p = 0.008$, with scores of 4.1, 5.2, and 5.2 for HMD, fish tank, and haptics, respectively (Figure 3.18b).

**The added value of haptics.** The use of haptics requires subjects to employ multiple sensory modalities to perform tasks. Most subjects in the haptic group were excited about the additional functionality, and claimed that haptic feedback did help in some way. Subjects in the haptics group were asked four questions that related to their experiences:

- consistency of the information from multiple senses;
- easiness of searching within the VE through touch;




(a) (b)

Figure 3.18: Mean values for the different questions regarding usability issues of VR systems (see Appendix C), all results are divided by VR system (HMD, fish tank, fish tank with haptics), error bars represent 95% confidence interval: (a) mean rank for the level of confidence in the answers; (b) mean rank for the level of demand on the subjects's memory.

- the effects of multimodal sensory on understanding the space;
- the effects of multomodal sensory on understanding the structure of the data set.

The first two questions used a standard seven point rating scale. The last two questions used a standard five point rating scale. 80% of the subjects from the haptic group thought the visual and haptic information was consistent, and that searching the VE through touch was easy. 75% of the subjects thought touch helped them better understand the space, and 80% thought it helped understand global structure. Subjects reported that haptics was especially helpful for the connectivity questions: "How many curved tube are there?" and "Please name all the regions the longest tube crosses," since the tubes are hidden behind other objects.

## 3.5 Conclusion

This chapter presents results of an empirical evaluation that compares the human performance of using different VR systems for four generic volume visualization tasks. The tasks are derived from data and questions that are asked by domain researchers studying mucuciliary clearance in CF. Results showed that the haptic fish tank VR system offered subjects both an inside-out and an outside-in perspective on a volume, a property that was identified as important to complete the tasks. Subjects using the HMD VR system were significantly slower than subjects from the other two systems, and were less accurate for the shape, density, counting the curved tube, and spatial region questions. Finally, none of the systems allowed for accurate judgment of different sizes of sphere objects, or of which regions of a volume had the densest spatial packing.

The speed difference for the HMD system was not unexpected, but the inferior task performance was quite surprising. Subjects' responses to questionnaires and anecdotal comments reveal that memory load was a significant factor. In the absence of an overview capability, subjects were forced to make an internal representation of the total volume; the dense nature of the data removed visible landmarks that can normally provide such a reference frame. It is believed that a future system that includes both an overview and an inside-out capability within the HMD would produce a system whose performance is at or above the level of the haptic-enabled system for some tasks. Furthermore, the poor performance of the HMD VR system for these data visualization tasks does not mean it is not appropriate for other tasks or applications. The lack of reference of the frame does not exist in other applications, for example, 3D gaming or architecture.

Some of the experimental results about the haptic system lead us to rethink the multimodal interaction scheme in current 3D interaction research as well. The design of a multimodal interface asks the designer to consider how the brain combines and integrates different sources of information in order to make the interface with additional sensory modality truly helpful. Correct combination and integration of multiple sources of sensory information, for example vision and touch, is the key to creating robust perception and judgment for tasks in a multimodal interaction situation. Combination does not only mean the presence of two modalities, but an integration and coordination that matches the user's available sensories. The author observes from the experimental results that the haptic system has

different effects on the user's performance for different tasks or different conditions of the same task. For some tasks, the presence of haptics maximizes the information received from both modalities (vision and touch). It also reduces the variance of the sensory estimation to increase its reliability. For other tasks, it does not. Clearly, the inclusion of touch should be guided by the criterion that it creates robust information combination for a multimodal interaction system.

# Chapter 4

# Tangible User Interfaces for A Clipping Plane in Visualization

Developing user interfaces for portraying and interacting with large quantities of data, with the intention of facilitating data analysis, is rapidly becoming one of the most challenging areas in both HCI and visualization research ([Sne92], [BCE+92]). In the previous chapter, a user study was designed to establish the knowledge of which kind of VR interface is most suitable for the visualization analysis tasks performed by domain researchers who study CF. Those results indicated that an immersive VR environment without overview capability did not help users with most of the tasks (identifying data structure and properties). The dense nature of the data sets used may be one of the main factors that contribute to the low task performance in immersive VR. The desktop VR environment enabled better performance in terms of accuracy and response time. But still the absolute performance is disappointing. Naturally, the author asks whether or not there are other means that could further improve user performance within such an environment. In this chapter, the author mainly investigates whether or not the inclusion of a tangible interface, and performing two-handed interaction with such interface, can help. The author presents an in-depth investigation into one specific interface aspect, i.e., the use of a clipping plane for exploring volumetric data. The proposed interface prototypes of a clipping plane were realized with the help of wireless vision-based tracking ([LM03], [ML02], [LR03]). These prototypes combine aspects of 3D input devices with tangible interaction elements. The clipping plane is included

in order to help a user explore the inside of a dense volumetric data set that can be rendered in either a 2D intersection image or a 3D view. By varying the design, these prototypes enable the comparison of different user interface strategies for performing the clipping plane interaction task. A user evaluation is carried out with these prototypes to measure their effectiveness.

## 4.1 Tangible User Interfaces and Two-handed Interaction

The ideas proposed in this chapter are based on existing work within tangible user interfaces (TUIs) and two-handed interaction. TUIs and two-handed interaction are human-centered approaches to potentially improve the interaction process by focusing on user behaviors that are well developed, due to continuously practice in daily life.

### 4.1.1 Tangible User Interfaces

To discuss tangible interaction or TUIs , it is necessary to describe the related concept of "Graspable User Interface" and a brief history of the term "tangible" first. George W. Fitzmaurice et al. designed Bricks in 1995, which were input devices that allow direct control of electronic or virtual objects through physical handles. In his dissertation, Fitzmaurice introduces and explores Graspable User Interfaces, putting emphasis on the fact that input control can be "space-multiplexed", which means different devices can be attached to different functions, each independently (but possibly simultaneously) accessible. Graspable User Interfaces enable gestures and grasping behaviors within a larger expressive range, and leverage a user's innate spatial reasoning skills and everyday knowledge of object manipulations. As to the introduction of TUIs, a paper by Hiroshi Ishii and Brygg Ullmer of MIT Media Lab, published in 1997, is often referred [IU97]. They coined the phrase "tangible bits" as: "an attempt to bridge the gap between cyberspace and the physical environment by making digital information (bits) tangible." They deliberately play with words in using the term "bits" - "we use the term bits to refer to physical things, but in computer science, the term bits refers to also digital things (i.e., binary digits)". Therefore, the phrase "tangible bits" attempts to reflect to not only digital, but also physical entities in the same manner.

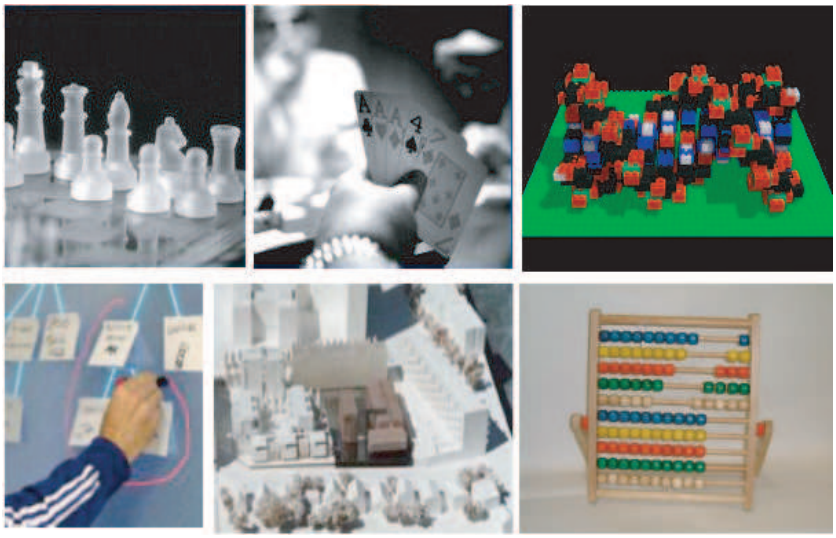


Figure 4.1: The tangible objects in normal life. (Photograph reprinted from [Kea90]).

Traditional input devices, for example the keyboard and mouse, are tools to manipulate digital spaces. These input devices are used to control (usually visual) representations displayed on output devices such as monitors, whiteboards or touch screens. TUIs are introduced to remove or at least decrease this distinction between input and output and try to establish new possibilities for interaction that merge the physical and digital worlds [UI01]. The ideal of TUIs will be those for which the distinction between input and output becomes blurred to a large extent. There are many examples in our daily life that use physical objects to integrate input and output, as shown in Figure 4.1. For example, when using an abacus, there is no distinction between "inputting" numbers and its representation of calculation results. This sort of blending is what is envisaged by the computing technology with TUIs. Touch and tangibility in both input and output are regarded as particularly important characteristics of TUIs. The physical representations of actual objects are often linked to TUIs (such as buildings in an architecture planning application, or bricks in a children's block building toys). The key issue to differentiate a typical desktop computer system and a tangible computing system is that in the first system so-called graphical user interfaces (GUIs) are used for control and manipulation. With such interface strategy, there is no direct mapping or tight coupling between the manipulation of the physical interface (e.g., the point and click of the mouse) and the resulting digital representation on the output

device (the screen).

### 4.1.2 Two-handed Interaction

Many interaction tasks undertaken in the real world are performed subconsciously by two hands (Figure 4.2). For example, Hauptmann observed users spontaneously using two hands to perform object rotation, translation, and scaling whenever possible [Hau89]. Moreover, users are likely to attempt two-handed interaction, even when such interaction style might not be well supported in applications or systems. It is believed that systems restricting users to one-handed interaction will be inadequate for most common manipulations. Therefore, many interface designers try to implement two-handed interaction techniques that can provide more powerful and natural interaction, particularly for complex manipulations tasks, such as rotating, which are frequently required in 3D interaction [MM95]. The two-handed interaction model widely accepted within the HCI community is the Kinematic Chain (KC) model proposed by Guiard [Gui87]. This model distinguishes the following principles for right-handed users:

- The right hand lags behind the left hand.
- Spatial reference of the right hand in motion can be found typically in the results of motion of the left hand.
- The motions of right and left hand are asymmetric in temporal and spatial scales. The right hand is good at rapid, small-scale movements; the left at slower, larger-scale movements.

There are some experiments show that two-handed interaction, when guided by the KC model, can be faster and more accurate than one-handed interaction, both in the case of 2D input devices [LZB98], [KBS94] and 3D input devices [GH98]. Other researchers have studied the difference between two-handed and one-handed interaction on real tasks, to circumvent the possible effects of computer-mediated interaction devices. Hinckley [Hin96] discusses some experiments that analyze manipulation of physical objects by both hands. Those experiments include an analysis of real-world tasks, such as handwriting and sketching, as well as a formal experimental task that involves using a tool in one hand to point at

Figure 4.2: Two-handed interaction in 3D space.

a target object held in the other hand. Hinckley's study contributed to the following pieces of formal knowledge:

- The experimental task used, which represents a general class of 3D manipulative tasks involving a tool and a reference object, requires an asymmetric contribution of both hands.
- For such tasks, performance is best when the right hand operates relative to the left one. Reversing the roles of the two hands significantly reduces performance in terms of both time and accuracy.
- Specializing the roles of the hands is significant only for skilled manipulation. This does not imply that two-handed input will be ineffective for tasks that afford symmetric manipulation, but instead restricts the scope of tasks where asymmetry factors will have important design implications.

Gribnau [Gri99] has studied the advantages of two-handed interaction for 3D conceptual modeling. His first evaluation of two-handed operations consisted of a comparison of one- and two-handed operations for two 3D assembly tasks. The first task was a stacking task that proved to have some difficulties. The second task was a puzzle task. The important finding was that working with two hands was faster than working with one hand. This result

was found after approximately an hour of practice by subjects who were not experienced in operating the computer with two hands. This short period suggests that two-handed 3D interaction was easy to learn, especially when subjects were required to learn a lot of new skills in a short time. In the second evaluation, Gribnau studied two specific aspects of two-handed operation. The first aspect was the clutch mechanism of the 3D system that determines how objects in the scene were selected. The second was the selection mechanism that determines how the manipulation space was coupled to the display space. The experimental results showed that the clutch mechanism did influence workload. The results presented also showed that no significant differences exist in completion times achieved by subjects with the more generally applicable selection method and with the more specialized selection method.

## 4.2 Tangible User Interfaces for Data Visualization

Interacting with volumetric data in 3D space requires spatial reasoning and 3D perception skills. Currently, the dominant interface for 3D manipulation with volumetric data is the desktop computer with a GUI that is controlled by a mouse and keyboard. Besides using VR systems, researchers are also trying to tackle this 3D interaction problem from the perspective of interface design, which refers to adding tangible elements into 3D input devices. 3D interfaces that are based on more advanced tracking technologies potentially provide the possibility of improving the 3D interaction process, and several studies have already been undertaken to develop alternative interfaces for this purpose. A large portion of these studies ([HTP+97], [ZM93], [CMS88]) have focused on generic 3D manipulation tasks, such as selection, positioning, and etc. Some others looked into the specific task of creating 3D intersections.

The Passive Interface Props (PassProps) [HPG94] was one of the first 3D interfaces that support continuous clipping in 3D space (Figure 4.3). The PassProps was developed to allow surgeons to explore a patient's anatomy data by interactively generating cross-sections through the 3D data. The PassProps contains a head prop, a cutting-plane prop for creating intersections, and a pen-like prop for planning trajectories. The six DOFs that specify the pose (i.e., translation and orientation) of each individual prop are tracked using

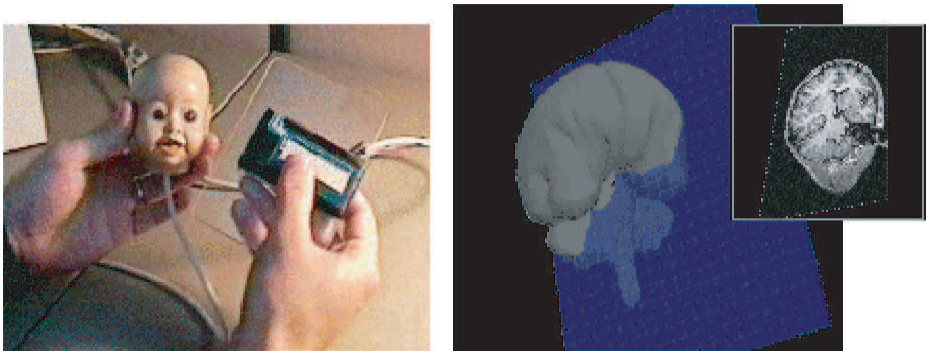


Figure 4.3: The Passive Interface Props from Ken Hinckley. (Photograph reprinted from [HTP+97]).

(wired) magnetic trackers from Ascension-Tech Inc. Visual feedback of the user's actions is provided on a computer display in front of the user. The head prop is used to manipulate the orientation of the patient's anatomy. The rendering of the volumetric data on the screen follows the rotation of the head prop. The rendering is always positioned in the center of the screen, i.e., it does not follow the translations of the head prop. The rendering scale (i.e., the zoom factor) is determined by the observer-to-object rendering distance, and is controlled by moving the head prop closer to or further away from the body. The user holds the cutting plane prop relative to the head prop to specify the location and orientation of the slice through the 3D data. The generated intersection image is presented on the display, next to a volume rendering of the 3D model.

De Guzman et al. presented two tangible devices for navigating a slice through a human body model [DGWlHCM+03]. Interface A consisted of a 30-inch 2D model of a human body, together with a U-shaped fork at the end of an adjustable arm that could be rotated 180 degrees along the device's baseboard (see Figure 4.4 upper left). Interface B consisted of a transparent 3D model of a human body and a free-moving hand-held fork. The fork in each case represented the intersection plane (window), and its position and orientation was used to generate an intersection image on a separate display (see Figure 4.4 upper right).

The Cubic Mouse (CMouse) [FP00] was developed to support exploration of 3D geological data (seismic data) and car crash analysis data (Figure 4.5). The CMouse allows users to specify three orthogonal cutting planes and to perform so-called "chair cuts" through the data. The prop is a cube-shaped case with three perpendicular rods passing approximately

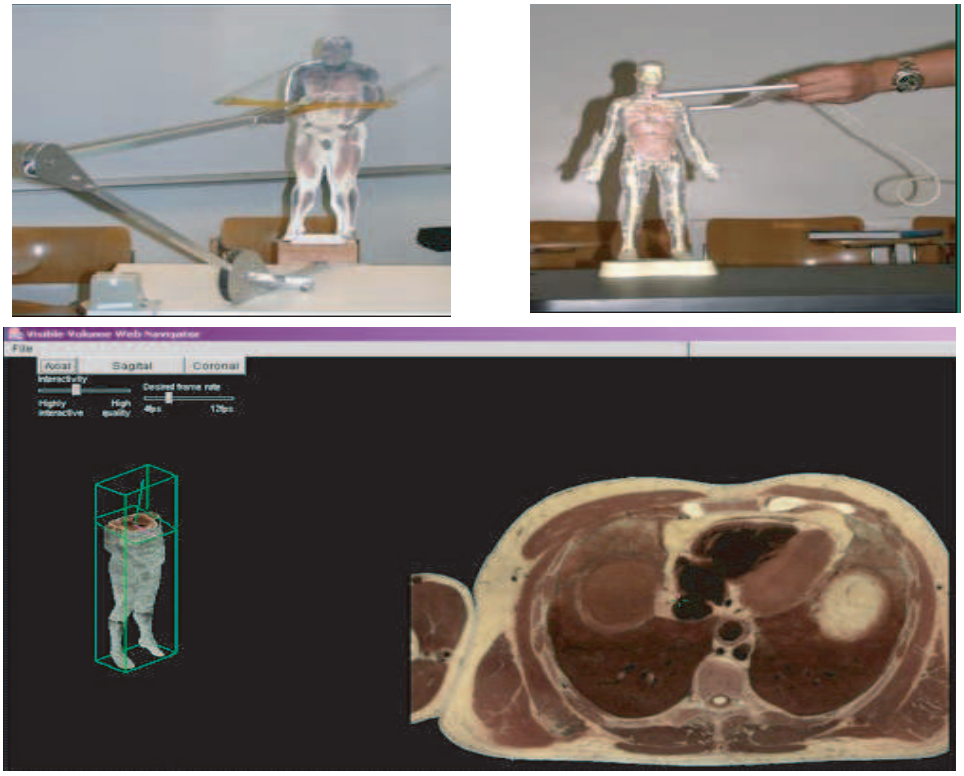



Figure 4.4: The tangible devices for navigation. (Photograph reprinted from [DGWlHCM+03]).

through the centers of two parallel faces of the case. It is usually held in the non-dominant hand. The rods are used to control three orthogonal slices through the 3D data set, i.e., by pushing or pulling a rod, usually with the dominant hand, the corresponding intersection plane moves back and forth. The movement of a slice is hence constrained to the direction orthogonal to the slice. There is also a (wired) magnetic tracker embedded in the cube-shaped case. The tracked six DOFs are used to translate and orient the data set in the virtual world, relative to the observer. The 3D data set and the orthogonal slices are visualized on a large stereo display in front of the observer.

There are some limitations in the above systems that are likely to have an impact on the usability and the user acceptance. First, because active tracking technology is used in these systems, the interaction elements need to be wired. Such wires obviously will pose some constraints on the freedom of the movement, an issue that is seldom mentioned, let alone

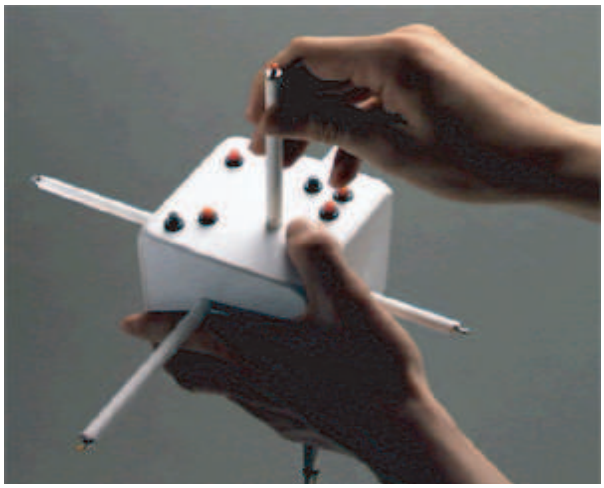


Figure 4.5: The Cubic Mouse. (Photograph reprinted from [FP00]).

evaluated. Alternative techniques, such as optical tracking, enable interaction elements to be passive and unwired, and are therefore likely to ameliorate this problem. Second, there is currently little insight into how different aspects of tangible interfaces, such as passive haptic feedback and two-handedness, assist users in their data analysis tasks in 3D space.

## 4.3 Design Practice of Tangible User Interfaces

The above discussion provides theoretical arguments and practical examples of designing tangible interfaces for two-handed use in order to take advantage of our physical body and improve the interaction. It stimulates us to have the idea of applying two-handed interaction in 3D space, particularly for visualization tasks. The study in this chapter is an attempt to explore the potential of two-handed interaction with the assistance of tangible user interfaces. In this section, the author explores the challenge of designing alternative tangible interfaces for controlling a clipping plane.

Positioning a clipping plane is a complex and frequent operation in volume visualization applications for improved data analysis. Such a plane is used to cut through a 3D data set in order to explore its interior structure. The common method of controlling the 5 (or 6) DOF of the virtual clipping plane, i.e., its position (3 DOF) and orientation (2 DOF in case of a plane, or 3 DOF in case of a window), is by means of a 2 DOF control device

such as a mouse. To accomplish this, the positioning task needs to be decomposed in at least 3 subtasks that require at most 2 DOF at a time. Despite the fact that such a 2D interface (in principle) enables the task, it is difficult for the user to obtain enough awareness of the spatial relationships and to manipulate effectively. This is due to the fact that the 2D interaction is unrelated to the natural interaction process in 3D space. The author therefore proposes alternative interface designs for the clipping plane task that make use of 3D (tangible) intersection devices. The design principles that the author adheres to in the design of these interfaces are the following:

- Easy to use. The interface should not distract the user from the actual clipping-plane task.
- Easy to learn. The interaction should be natural and intuitive, requiring little explanation and training.
- Adequate perceptual feedback. The interface should provide (passive) tactile and visual cues that assist in the interaction.
- Real-time interaction. The interface should work in real-time despite of the use of volume rendering of realistic high-resolution data.

### 4.3.1 Hardware and Software

Several different interaction prototypes are implemented on a desktop visualization system that contains the following functional components:

1. Tangible devices. The currently available tangible interaction devices include a wooden cube, a metal frame and a metal pen. The frame and pen are made of lightweight metal (aluminum). All devices are painted black in order to reduce unwanted reflections. Every interaction device is characterized by one or more unique dot patterns. The patterns consist of small dots that are created with infrared-reflecting tape. In case of the cube, each of the six sides contains a unique pattern. The interaction devices are tracked in 3D using available vision-based stereoscopic tracking software [LM03], [LR03], [MQA+04], [ML02].

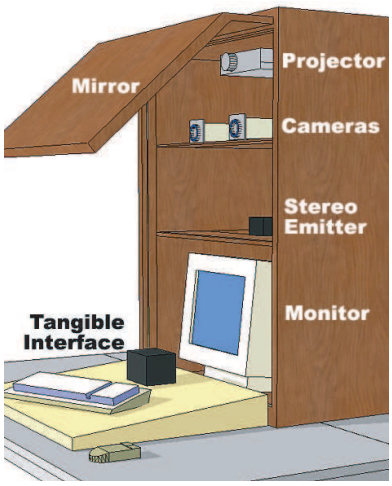



Figure 4.6: The diagram of the system set-up.

2. A GUI enables users to interact with the underlying application;
3. A control unit that takes care of the following functions:
   - Data input and output;
   - 3D manipulation (the position and orientation) of the data, interaction elements, and viewing camera;
   - Parameters control, such as TF specification and etc.
4. A direct volume rendering engine that converts a 3D data set into one or more displayable 2D images on a computer screen;
5. A display in which the resulting 2D images are rendered.

The hardware set-up is organized around a DELL graphics workstation Precision 530 (Pentium IV, 2.4 GHz, 512 MB RAM) with the following specific interface components:

- An ATI *FireGL 4* graphics card coupled to an infrared emitter from StereoGraphics Inc.
- Two analog Leutron Vision LV-8500 progressive scan CCD cameras (720x576 pixels, 50Hz frame rate) with COSMICAR/PENTAX lenses with a focal length of 12mm and

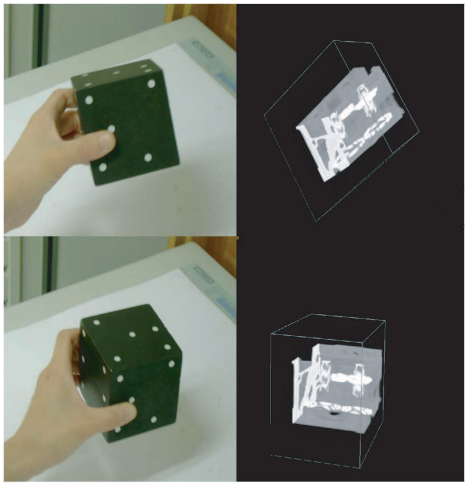
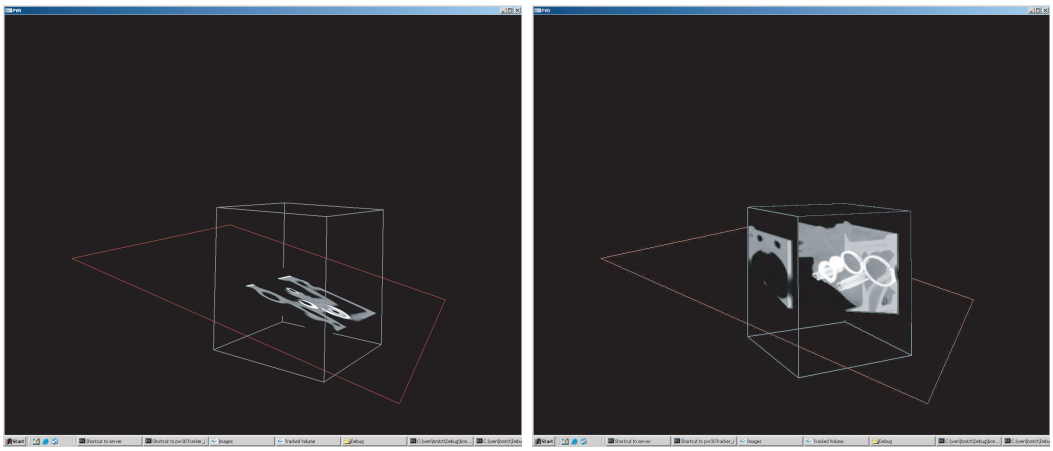


Figure 4.7: 3D manipulation of a volumetric data set with a tangible cube.

infrared transparent filters that block visible light; these cameras are connected to two synchronized Leutron Vision PictPort H4D frame grabbers.

- A 14' CRT monitor with a vertical refresh rate up to 120Hz, so that stereoscopic images can be viewed with the help of a pair of active liquid crystal shutter glasses (CrystalEyes 3).

A wooden chassis has been constructed for integrating the different components and creating a workspace for the users. The two infrared cameras are mounted on the upper layer of the wooden chassis as shown in Figure 4.6. A silver mirror mounted on a wooden slab is hung in front of the chassis under an angle of 45 degrees to reflect the image of the user's hands with the interaction devices to the cameras. The use of the wooden cabinet for the cameras makes the system set-up stable and allows for easy transportation. In the current prototype, there are no provisions for tracking the user's head (which can be useful for also providing motion parallax feedback in the displayed image). The rendering engine uses hardware-supported 3D texture mapping in OpenGL to speed up the volume rendering. The 3D texture mapping algorithm supported by the ATI *FireGL 4* graphics card assumes that every volumetric data is contained within a texture memory that is associated with a virtual cube. In the prototype, the user is provided with a wooden cube that is tracked by the cameras. This cube can be rotated to control the orientation of the virtual cube,



(a) (b)

Figure 4.8: (a) the slice mode for the clipping interaction; (b) the opaque mode for the clipping interaction.

with its associated volumetric data, and moved towards or away from the user's body to change the zoom factor (see Figure 4.7). The image on the display, which is the result of volume rendering for a fixed virtual camera position that is 6 centimeters behind the actual observer's viewpoint, hence changes in accordance with the movement of the tangible cube. The cube can also be placed on a small (physical) pedestal in case a user prefers to perform clipping operations with a single hand. Although the data set remains in a fixed position in such a case, its orientation can still be varied discretely in a very simple way, i.e., by changing the side of the cube that is resting on the pedestal.

The virtual clipping plane is controlled by means of the devices proposed below. It can operate in two modes, i.e., slice mode and opaque clipping mode. In slice mode, only the planar intersection image is displayed in 3D space, as shown in Figure 4.9a. In opaque clipping mode, the part of the volumetric data that is in front of the clipping plane is made transparent, as shown in Figure 4.8b.

### 4.3.2 Tangible Frame Prototype for Controlling A Clipping Plane

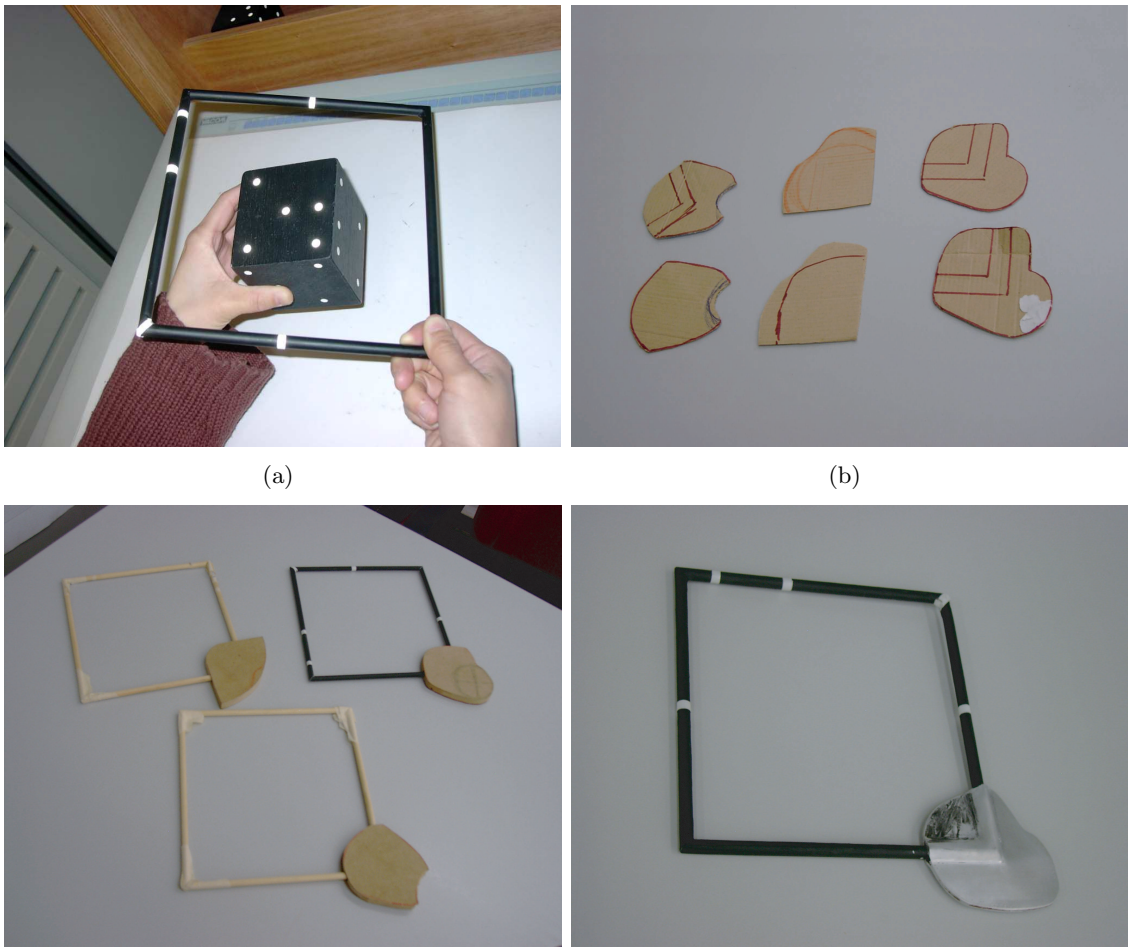
The author designed a prototype that uses a square-shaped metal frame to control the virtual clipping plane. The five infrared-reflecting stripes on three of its sides form a unique planar

pattern as is shown in Figure 4.9a. The six DOFs of the frame (three DOFs for position and three DOFs for rotation) are monitored continuously by the vision-based tracking algorithm. The appearance of this device makes its purpose very obvious. While using this prototype, the user positions the cube with his/her non-dominant hand, and grasps the frame with his/her dominant hand on the side that has no dots on it. The physical cube intersects with the physical frame in a way that agrees one-to-one with the intersection of their virtual counterparts on the screen. As a result, even though all six DOFs are enabled when moving the plane frame, the prototype does not seem difficult to control. This is expected to result in an interface that is easy to learn and use. The cube can again be placed on a small pedestal in case the user prefers to perform clipping operations with a single hand. In order to improve the ease of use, the author pursued a further design by adding different handles to the frame. Three different types of handles were designed in order to find out the best solution. Firstly, the paper mockups of three kinds of handles were made in order to have an impression how they would look (see in Figure 4.9b). Three wooden plane frames with differently designed handles were constructed, as shown in Figure 4.9c, to let users experience how they felt.

### 4.3.3 Informal Evaluation of Prototypes

The author asked several colleagues from to try out the prototypes and give informal feedback. The tangible clipping interfaces received many positive feedbacks. Further survey among the users indicated that there is no difference in preference regarding the handle shapes, which meant the shape of a handle will not be major factor in the future user experiment. The final design of this clipping plane frame is shown in Figure 4.9d.

The next logical step was to undertake a more structured and formalized experiment with some of the prototypes. The author adopted the tasks performed in the previous chapter as the starting point since the systems tested in the previous experiment did not provide satisfying user performance. More specifically, the experimental goal is to ask users to observe volumetric data with the clipping plane while answering the same questions regarding data properties and volume structures as in the previous experiment. By repeating the same experimental tasks, the author wishes to establish the knowledge regarding the effect of those tangible prototypes on the user's performance (response time and accuracy)



(a) (b)

(c) (d)

Figure 4.9: (a) The frame-like tangible interface for virtual clipping plane; (b) paper models of three handles; (c) wooden models of three handles; (d) The final design of the plane-like tangible interface for virtual clipping plane.

in case of a desktop VR environment. The author also wishes to verify whether or not the observed preference from the informal evaluation for the presence of a tangible clipping frame and its 2D intersection image holds in terms of user performance for those visualization tasks. Moreover, the author wants to study the effect of two-handedness and of the form factor of the interaction device (a fixed virtual clipping plane or a tangible planar frame).

## 4.4 Empirical Work

### 4.4.1 Experiment Design

In order to manipulate the 3D rendering results of any scientific data, one needs either 2D input devices, such as mouse, or 3D input devices. Next to the fact that they provide different DOFs for a user during operation, they also have a lot in common: the user manipulates the data by means of the input devices, observes from different viewpoints and analyzes his/her observations in order to plan the next action. As pointed out by Ware, human spatial perception implies that "we should design with a 2 $\frac{1}{2}$ D attitude" for 3D information processing [War01]. More specifically, he suggests that the following principles should be adhered to:

1. Use 3D objects to represent data entities;
2. Emphasize 2D layout and use it to support navigation in 3D spaces.

Following these principles, the author is interested in investigating the performance differences of tangible user interfaces for data analysis tasks compared to a traditional 2D user interface (mouse). The added-value of a 3D clipping plane function with/without visualization of the 2D intersection image is of great interest as well. The relative performances with various set-ups are compared for four generic tasks involved in the analysis of high-density volumetric data.

The author has formulated the following hypotheses:

- *There will be a performance difference between a 3D tangible interface and a mouse-based 2D interface.* 3D object perception theories suggest that users will manipuate

an object more easily when it is controlled by a 3D input device that is similar in shape. Introducing tangible user interfaces in 3D interaction is regarded as an approach towards applying this principle, since tangible interfaces offer opportunities to couple more tightly input and output and to emphasize affordance. The chance to reduce the mental workload for a user while performing 3D manipulation and data analysis tasks is also expected to be higher in case of tangible interfaces. Moreover, 3D tangible interface provides more DOFs for manipulating a volumetric data set. A mouse interface can only provide 2 DOFs at a time.

- *The inclusion of a 3D clipping plane function and a seperate 2D intersection image that corresponds to the clipping plane will make difference in the user performance for the visualization tasks.* In details, this hypothesis means the follows:

  1. *A fixed clipping plane will produce different user performance compared to no clipping plane.*
  2. *Including a clipping plane that is operated by a tangible frame will enable users obtain different performance compared to the case with a fixed clipping plane and the case without any clipping plane.*
  3. *The inclusion of a separate 2D intersection image that corresponds to the clipping plane (fixed or arbitrary) will help users produce different performance compared to the case without an intersection image.*
  4. *An arbitrary clipping plane that is operated by a tangible frame will enable users obtain different performance compared to the case with a fixed clipping plane, while an intersection image is presented.*

According to Ware [War01], "where object structures are important, the design of a 3D user interface should make this structure principally to be laid out in a 2D plane, especially if a static view is required". In the context of this study, a solution that adheres to this principle is a 3D clipping plane. A clipping plane can cut through a 3D data set, which enables a user to observe the inside structure of a volume. Therefore, it may have extra effects on those data analysis tasks. At the same time, an arbitrary clipping plane provides users with more flexibility during the operation and observation than a fixed clipping plane. Hence, it may produce different user

performance as well. Furthermore, adding a 2D intersection image that corresponds to the clipping plane can map the inside structures and details of a 3D volume into a 2D plane, hence making them visually more accessible. Therefore, interpreting these data can become easier, which may contribute to improving user performance.

### 4.4.2 Apparatus

In order to achieve real-time rendering performance in the user study, a different hardware configuration was needed from the one reported in the previous section. The new configuration is organized around two Dell graphics workstations with several specific interface components.

The first workstation is the one discussed in previous section and mainly used for the purpose of tracking the tangible interfaces. It consists of the following components:

- One DELL workstation Precision 530 (Pentium IV, 2.4 GHz, 512 MB RAM) with ATI *FireGL 4* graphics card.
- Two analog Leutron Vision LV-8500 progressive scan CCD cameras (720x576 pixels, 50Hz frame rate) with COSMICAR/PENTAX lenses with a focal length of 12mm and infrared transparent filters (that block visible light); these cameras are connected to two synchronized Leutron Vision PictPort H4D frame grabbers.

The second workstation includes the following components:

- One DELL workstation Precision 670 (Intel Xeon, Dual CPU, 3.2 GHz, 2.0 GB RAM) with NVidia *Quadro FX 4500* graphics card coupled to an infrared emitter from StereoGraphics Inc.
- A 14' CRT monitor from DELL with a vertical refresh rate up to 120Hz, so that stereoscopic images can be created with the help of a pair of active liquid crystal shutter glasses (CrystalEyes 3).
- A second 15' LCD display from DELL that will be used to display the 2D intersection images.

The separation of tracking and 3D rendering functions into two different machines helps achieve better performance during 3D interaction with large data sets.

The rendering engine still uses hardware-supported 3D texture mapping in OpenGL. The volume rendering algorithm used in the prototype platform has been tuned in order to fully use the power of the Nvidia *Quadro FX* graphics card. The algorithm that renders the non-polygonal isosurfaces in the data sets is based on the approach proposed by Westermann and Ertl [Wes90]. In a preprocessing step, the gradient vector is computed for each voxel using the central differences method. The vector components must be normalized to adjust their signed range [-1; 1] to [0; 1]. The original scalar component is similarily normalized to the unsigned range [0; 1]. The three components of the normalized gradient vector together with the normalized scalar value of the data are stored as RGBA quadruplets in a 3D texture. The alpha test allows discarding incoming fragments, which depends on the outcome of a comparison of the incoming alpha value with a user-specified reference value. In our case, the alpha channel contains the scalar intensity value and the alpha test is used to discard all fragments that do not belong to the isosurface specified by the reference alpha value. The specific set-up for the OpenGL alpha test is as follows:

- glDisable(GL_BLEND);
- // Enable Alpha Test for isosurface
- glEnable(GL_ALPHA_TEST);
- glAlphaFunc(GL_EQUAL, IsoValue);

The 3D image on the display is the result of volume rendering for a fixed virtual camera position that is behind the actual observer's viewpoint to avoid perspective distortion. The distance and height of the virtual camera in 3D rendering is calculated based on the *view frustum* [Fer]. A user needs to work with the tangible interfaces in a comfortable position while the objects in the experimental volume are shown with reasonable sizes in the center of the screen.

The stereoscopic shutter glasses and head tracker were not equipped during the experimental study, which was different from the previous VR study. This means that no

stereoscopic viewing will be presented to a subject, and there is no motion parallax effect. The detailed discussion can be found in the later section regarding the reason of such set-up.

### 4.4.3 Data and Task

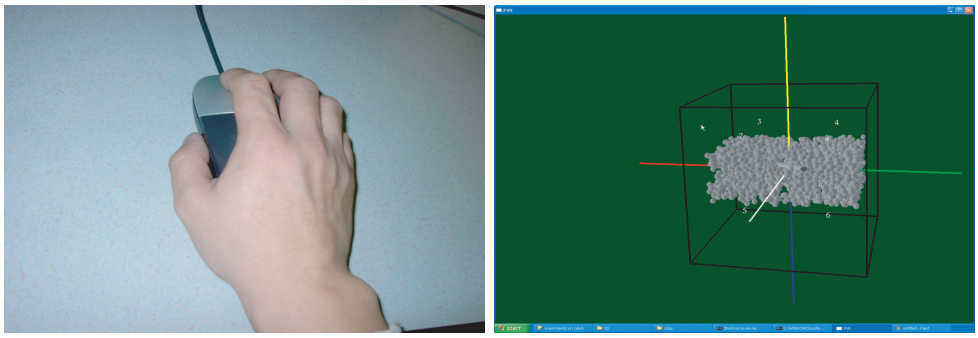
Simulated volumetric data are generated to act as trials in this study. A random number of two to four types of differently-shaped objects (sphere, ellipsoid, cylinder, and curved tube) are inserted with random positions within a rectangle box, whose length, width and height are 200 by 200 by 200 OpenGL units. These objects may overlap with each other to become connected. The objects' properties (size, shape) and the volume density form experimental conditions that vary between trials. The bounding box of the volume is positioned in the center of a cube and subdivided into equally-sized eight regions (a $2 \times 2 \times 2$ array in the $x$, $y$, and $z$ directions) within which object density may differ. Regions are labeled with unique numbers (1 through 8) to enable subjects to identify them.

There are always spheres and at least one curved tube within every volume. Trials may also contain ellipsoids, cylinders, and up to three additional curved tubes. Spheres sizes may vary between four possible radii in the range from 6 to 12 OpenGL units. The density of objects within each region is controlled to be sparse, medium, or dense. A single dense region (the "densest"region) exists within each volume. Sparse regions contain between 10%–60% of the number of objects in the dense region, while medium regions contain between 60%–90% of this number.

Subjects were asked to provide their answers as accurately as possible and to minimize response time as well. The size, density, and counting curve tubes questions were presented in a multiple choice format. Subjects were asked to describe the name of each kind of object for the shape question and to specify the region numbers for the tracking tube question. All the answers from the subjects were recorded by the experimenter on the answer sheets.

### 4.4.4 Experimental Procedure

A between-subject design was used, with interface type as the independent variable (Table 4.1):



(a) (b)

Figure 4.10: Condition 1: (a) mouse input and (b) a 3D view of the data set.

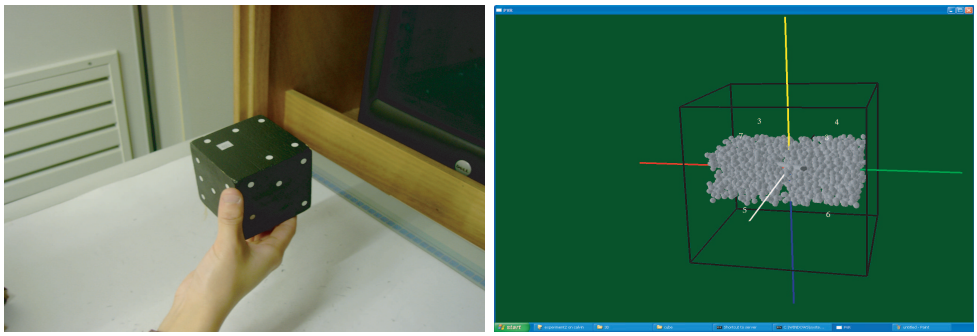
1. Condition 1 (abbreviated as mouse or M): the interaction device is a mouse. The visual feedback consists of a perspective view on the data set. A user can use the left button of the mouse to rotate the data set. The right button can be used to zoom in or zoom out (Figure 4.10). The image of the data set is always centered in the middle of the screen.

2. Condition 2 (abbreviated as cube or C): the 3D system is a non-immersive VR system with a tangible cube to orient and position the data set, which acts as the baseline system. The perspective view on the data set is the only visual feedback available (Figure 4.11);

3. Condition 3 (abbreviated as fixed-plane or CF): baseline system with a fixed virtual clipping plane. A user can manipulate the cube and cut through the data using a fixed clipping plane (Figure 4.12a and b). The fixed virtual plane is oriented orthogonally to the viewing direction;

4. Condition 4 (abbreviated as tangible-frame or CT): baseline system with a tangible cube for manipulating the data set and a square frame for manipulating the clipping plane. The movement of the virtual clipping plane follows that of the physical planar frame. The visual feedback on the screen shows the 3D representation of both the data set and the virtual clipping (Figure 4.13a and b);

5. Condition 5 (abbreviated as fixed-intersection or CFI): the interaction devices are the



(a) (b)

Figure 4.11: Condition 2: the rendered volume follows the position and orientation of the physical cube, as shown in the perspective view.

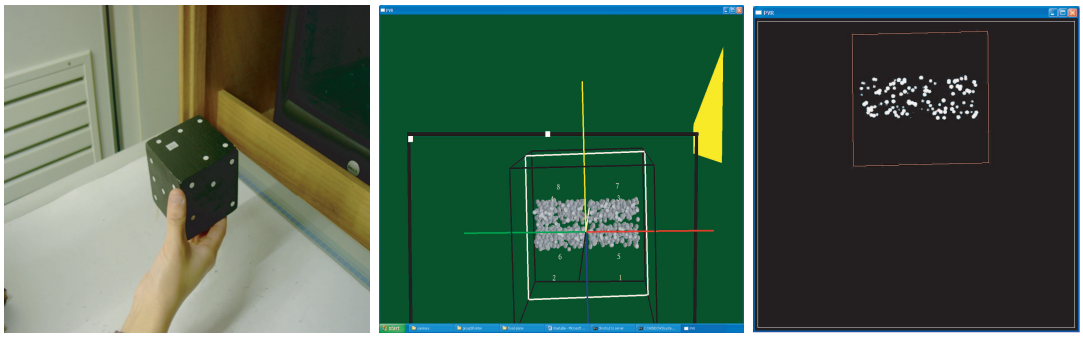
same as in condition 3. Next to the visual feedback of the 3D scene, a second display shows the synchronized 2D intersection image (Figure 4.12a, b and c);

6. Condition 6 (abbreviated as tangible-intersection or CTI): the interaction devices are the same as in condition 4. The visual feedback consists of both the 3D scene (the data set and the clipping plane) and a synchronized 2D intersection image on a separate display (Figure 4.13a, b and c).

Table 4.1: Variables List

| **Independent Variable** | **Dependent Variables** |
|---|---|
| User Interfaces<br>(M, C, CF, CT, CFI, CTI) | time,<br>the accuracy of the tasks (shape, size, density, connectivity) |

Subjects were randomly assigned to one of six groups. They completed several steps during the experiment. As part of an initial interview session, they signed a consent form, answered basic demographic questions (age, gender, and occupation or major field of study), and identified their frequency of computer use and prior experience with any kind of 3D VR visualization system. Then a training session introduced the equipment and described the tasks to be performed. Next, the formal experiment session was conducted. Each experiment session included 11 trials, with each trial containing a single volumetric data

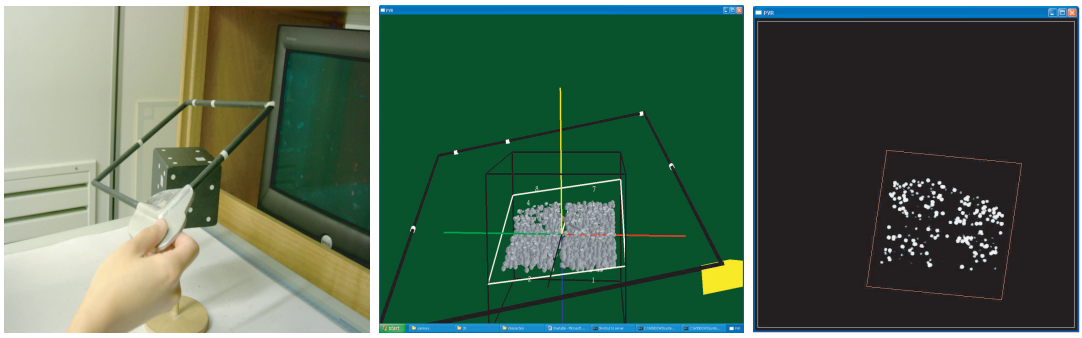


(a) (b) (c)

Figure 4.12: Condition 3 and 5: condition 3 includes a cube for manipulating the data (a) and a fixed virtual clipping plane (b); while condition 5 also includes a separate display of the 2D intersection image (c).

set. The decision on the amount of trials was based on the experience gained in the previous study and on the results of the pilot test. The experience from the previous VR study has indicated that it takes one subject at least 2 hours on average to finish all 16 trials even with the simplest fish tank VR system (without haptic feedback). And most of the subjects became less concentrated after 10 trials in both the previous VR study and the pilot test. Therefore, the author reduced the number of trials to 11 trials in this study, including one training trial. The consequence of this choice is that the power of the study is lower than in the study in the previous chapter, the consequence of which will be discussed in a later section. The 11 data sets were different from one another, and varied by object properties (type, size, position) and volume density. However, the same set of trials (all 11 data sets) were used for all 6 groups (M, C, CF, CT, CFI and CTI) in the same order.

Two dependent variables, the time taken to respond for each trial and the subject's answers for all four tasks were recorded (see Table 4.1). A short break was provided every half hour or whenever a subject asked for one. After completing the last trial in the formal experiment session, subjects filled out a questionnaire describing their preferences, any suggestions they had on how to improve the system, and so on. The study ended with a short debriefing during which the experimenter summarized the study goals.



(a) (b) (c)

Figure 4.13: Condition 4 and 6: condition 4 includes a cube for manipulating the data (a) and an arbitrary virtual clipping plane (b); while condition 6 also includes a separate display of the 2D intersection image (c).

## 4.5 Results

30 subjects volunteered for the experiment, 12 males and 18 females. The subjects were randomly assigned into one of the six groups: 5 subjects (3 males and 2 females) for the C group, 5 subjects (3 males and 3 females) for the CF group, 5 subjects (3 males and 3 females) for the CT group, 5 subjects (1 male and 4 females) for the CFI group, and 5 subjects (3 males and 2 females) for the CTI group. The M group also had five subjects (1 male and 4 females).

The age of each subject and the frequency of computer use were recorded before the experiment began. The measurement of computer experience questions used a standard seven point rating scale. Average ages were 25, 30, 28, 29, 25, and 29 for the M, C, CF, CT, CFI and CTI groups respectively. The frequencies of computer use were 6, 6, 6, 7, 7 and 6 for the six groups. These data suggested subjects' profiles within each group are similar in terms of ages and computer experience.

### 4.5.1 Summary

Two types of measures of performance were analyzed for each trial that a subject completed: response time $rt$ and error rate $\widehat{P_e}$. A single $rt$ value representing the total time in seconds

needed to complete all four tasks was captured for each trial. The author did not record the individual $rt$ for each task since it was difficult to record separately. Separate $\widehat{P_e}$ values for four tasks subjects completed were obtained for the four tasks that subjects completed.

- For the shape, size, density and numerosity questions, the way to code the answers was the same as the one used in the previous chapter (1 for correct and 0 for incorrect). Error rate $\widehat{P_e}$ was still the proportion of wrong answers among all the answers;
- For the spatial region question, subjects' answers were still coded in two parameters: the probabilities of *false negative* $\widehat{P_{FN}}$ and *false positive* $\widehat{P_{FP}}$, as used in a ROC.

Table 4.2: Description of the group pair notation for Fisher's Exact Test

| **Group pair** | **Explanation** |
|---|---|
| ⋆ C vs. M | 3D cube vs. 2D mouse |
| ⋆ CF vs. CT | Fixed vs. Arbitrary clipping plane ( without 2D intersection image) |
| ⋆ CF vs. CFI | Without vs. With 2D intersection image (Fixed clipping plane) |
| ⋆ CT vs. CTI | Without vs. With 2D intersection image (Arbitrary clipping plane) |
| ⋆ CFI vs. CTI | Fixed vs. Arbitrary clipping plane (with 2D intersection image) |
| M vs. CF | 2D mouse vs. Fixed clipping plane (without 2D intersection image) |
| M vs. CT | 2D mouse vs. Arbitrary clipping plane (without 2D intersection image) |
| M vs. CFI | 2D mouse vs. Fixed clipping plane (with 2D intersection image) |
| M vs. CTI | 2D mouse vs. Arbitrary clipping plane (with 2D intersection image) |
| C vs. CF | 3D cube vs. Fixed clipping plane (without 2D intersection image) |
| C vs. CT | 3D cube vs. Arbitrary clipping plane (without 2D intersection image) |
| C vs. CFI | 3D cube vs. Fixed clipping plane (with 2D intersection image) |
| C vs. CTI | 3D cube vs. Arbitrary clipping plane (with 2D intersection image) |
| CF vs. CTI | Fixed (without 2D image) vs. Arbitrary clipping plane (with 2D image) |
| CT vs. CFI | Arbitrary (without 2D image) vs. Fixed clipping plane (with 2D image) |

For $rt$ statistics, trials were analyzed by interface condition (M, C, CF, CT, CFI and CTI). For $\widehat{P_e}$ statistics, trials were analyzed by interface condition (M, C, CF, CT, CFI and CTI) and task (shape, size, density, or connectivity). The logarithm of $rt$ on average and $\widehat{P_e}$ for different conditions were then compared. For each task, the answers for every combination of two conditions were compared in order to locate the significant effects, such as between CT and CTI. All possible combinations (in total 15) are listed in Table 4.2. **The**

**author is particularly interested in those pairwise comparisons that are labelled with $\star$ since they corresponds to the original hypotheses.** In the following sections, for each task, only significant results are listed. In summary, the following important differences in performance were identified:

1. Condition CTI had significantly longer $rt$, than the other five conditions. Condition M had the shortest $rt$.

2. $\widehat{P_e}$ was significantly higher in condition CTI than in condition C for the size task. For the rest of the tasks, condition CTI had significantly lower $\widehat{P_e}$ than condition C. For the tasks of counting the number of tubes and locating the longest tube, the CFI condition was significantly more accurate than the C condition . Similarly, condition CTI was significantly more accurate than condition CT for the same tasks.

3. For the tasks of counting the number of tubes and locating the longest tube, the error rate $\widehat{P_e}$ and only $\widehat{P_{FN}}$ in both condition CFI and condition CTI were significantly lower than in condition M. This also applied to condition CFI compared to condition CT.

4. Additional significant effects include that condition CTI was significantly more accurate than condition M and condition CF for the density task; condition CTI was significantly more accurate than condition CFI for the shape task. For the size task, condition CTI was significantly less accurate than condition CT.

### 4.5.2 Detailed Analysis of Quantitative Results

#### Performance times

The response time $rt$ needed to complete the four tasks in each trial was recorded during the formal experiment session. Average time for the M, C, CF, CT, CFI and CTI conditions were $rt = 231.9\,\mathrm{s}, 268.5\,\mathrm{s}, 267.1\,\mathrm{s}, 244.3\,\mathrm{s}, 247.8\,\mathrm{s}$ and $329.6\,\mathrm{s}$ respectively. Overall, the C condition spent 16% more time compared to the M condition. And the CF condition spent 23% more time compared to the CFI condition. However, the condition CT spent 35% less time compated to the CTI condition. Condition CF spent more time than condition CT

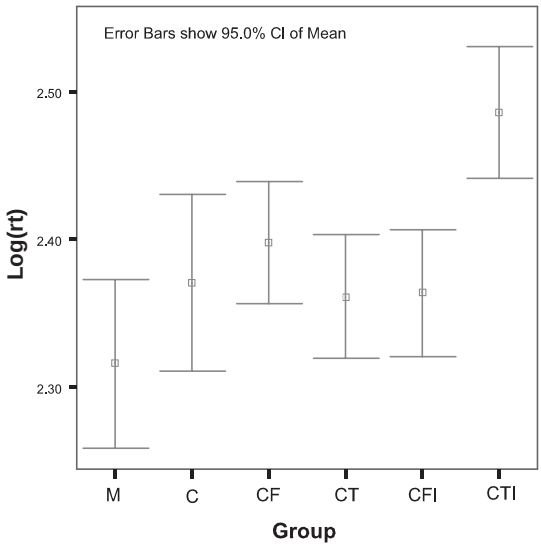



Figure 4.14: ANOVA of lg($rt$) as a function of experimental conditions, together with 95% confidence interval.

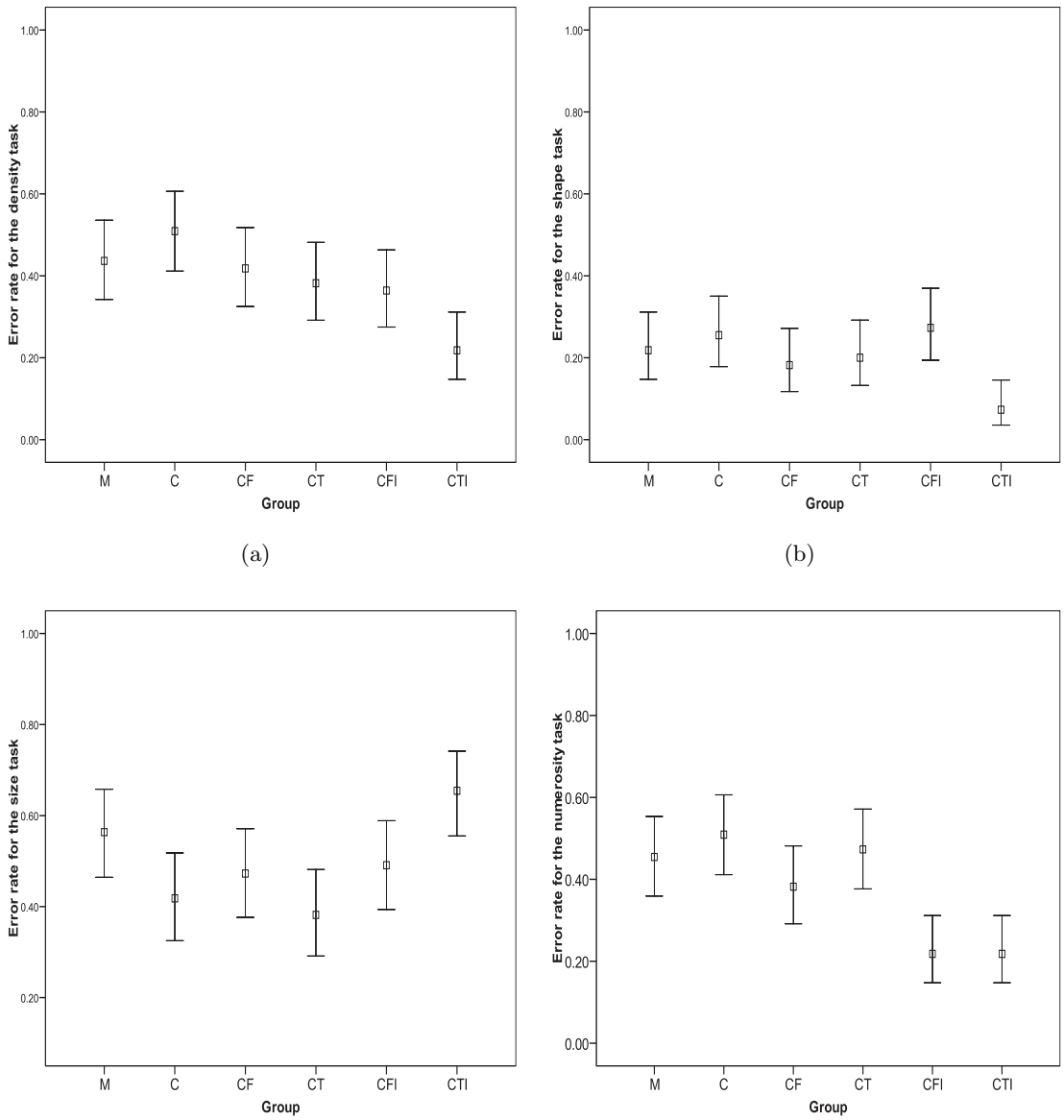
(9%), while condition CFI spent less time than condition CTI (33%). The *ANOVA* for the logarithm of $rt$ indicates that the amount of time spent was influenced by the experimental conditions, $F(5, 324) = 5.527, p = 0.000 < 0.02$ (Figure 4.14). Post-hoc paired comparisons showed that in condition CTI, users spent significantly more time than in the other five condtions. No significant effect is found across other conditions.

Table 4.3: Fisher's Exact Test for the density task

| **Group pair** | **Results** |
|---|---|
| M vs. CTI | **0.025** |
| C vs. CTI | **0.003** |
| CF vs. CTI | **0.040** |

### Accuracy in the density task

For the density task, the answers for every combination of two conditions were compared through a Fisher's exact test to find out whether or not there is an association between the error rate of locating the densest region and the interface used. The results are shown




(a) (b)

(c) (d)

Figure 4.15: $\widehat{P_e}$ for the different experimental conditions. All results are divided by condition (M, C, CF, CT, CFI and CTI), error bars represent 95% confidence interval: (a) $\widehat{P_e}$ for the density task; (b) $\widehat{P_e}$ for the shape task; (c) $\widehat{P_e}$ for the size task; (d) $\widehat{P_e}$ for counting the number of curved tubes.

in Table 4.3 and summarized as follows: none of our hypotheses has been proven to be significant. However, the experiment demonstrates that the subjects in the M , C and CF conditions: produced significantly more errors than the subjects in condition CTI.

In absolute terms, subjects in all six conditions were fairly inaccurate in answering this density question, with error rates $\widehat{P_e} = 0.436, 0.509, 0.418, 0.382, 0.364, 0.218$ for the M, C, CF, CT, CFI and CTI conditions respectively. subjects in the CTI condition are more accurate than their counterparts in the other five conditions. Error rates $\widehat{P_e}$ for the six conditions are shown in Figure 4.15a with 95% confidence interval.

Table 4.4: Fisher's Exact Test for the shape task

| **Group pair** | **Results** |
|---|---|
| $\star$ CFI vs. CTI | **0.010** |
| C vs. CTI | **0.019** |

### Accuracy in the shape task

The results of the Fisher's exact analysis for the shape task are shown in Table 4.4 and the conclusions as to the relative performance of all six conditions are the following: one of our hypotheses is proven to be significant, i.e., subjects in condition CFI performed significantly worse than their counterparts in condition CTI. In addition, subjects in condition C were significantly less accurate than their counterparts in condition CTI. In absolute terms, subjects in all six conditions were fairly accurate in judging the shapes, with an error rate $\widehat{P_e} = 0.218, 0.255, 0.182, 0.200, 0.273, 0.073$ for the M, C, CF, CT, CFI and CTI conditions, respectively. Subjects in the condition CTI again were more accurate than their counterparts in the other five conditions. The error rates $\widehat{P_e}$ for the six conditions are shown in Figure 4.15b respectively with 95% confidence interval.

Further analysis indicates that user performances of the six conditions depend on the experimental condition in terms of the number of shapes (Figure 4.16). In particular, condition CFI had significantly higher error rate than the other conditions when all four shapes were presented.

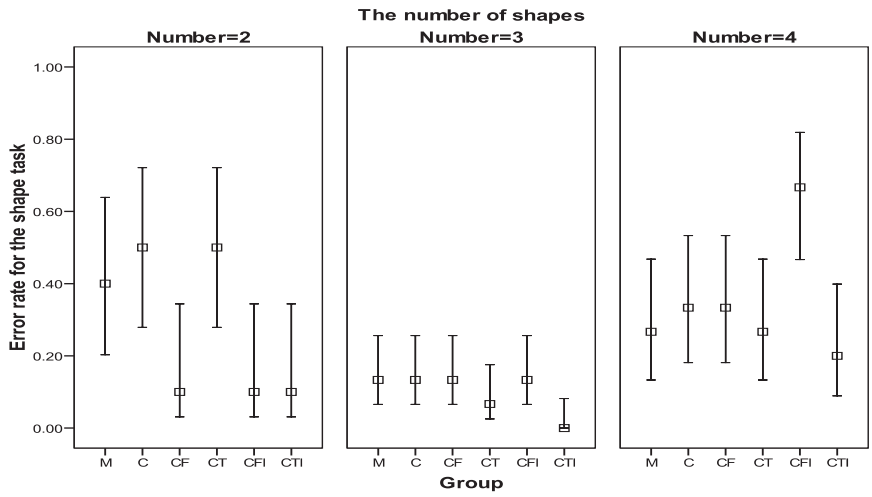



Figure 4.16: $\widehat{P_e}$ for the shape task based on the number of shapes.

Table 4.5: Fisher's Exact Test for the size task

| Group pair | Results |
|---|---|
| ⋆ CT vs. CTI | **0.007** |
| C vs. CTI | **0.021** |

### Accuracy in the size task

The results of the Fisher's exact analysis for the size task are shown in Table 4.5. Significant difference in performance can be observed between condition CTI and condition CT, which matches one of our hypotheses. In addition, there is significant difference between condition C and condition CTI. In general, subjects in condition CTI demonstrated the worst performance. In absolute terms, subjects in all six conditions were not very accurate in counting the number of differently sized spheres, with an error rate $\widehat{P_e} = 0.564, 0.418, 0.473, 0.382, 0.491, 0.655$ for the M, C, CF, CT, CFI and CTI conditions, respectively. The error rates $\widehat{P_e}$ with 95% confidence interval for the six groups are shown in Figure 4.15c.

In the previous analysis, user performance was only compared according to the error rate, that is, equal to the proportion of wrong answers among all the answers. When

Table 4.6: Fisher's Exact Test for the size task in terms of estimation difference

| Group pair | Results |
|---|---|
| ⋆ CF vs. CFI | **0.001** |
| ⋆ CT vs. CTI | **0.017** |
| ⋆ CFI vs. CTI | **0.022** |
| M vs. CF | **0.001** |
| M vs. CT | **0.037** |
| C vs. CF | **0.02** |
| C vs. CTI | **0.022** |
| CF vs. CTI | **0.036** |

performing the size task, three cases can arise. The number of sizes can be estimated correctly, over-estimated or under-estimated. The resulting Fisher's exact analyses, based on these three instead of two (right or wrong) categories, are reported in Table 4.6. This more refined analysis reveals a significant difference between condition CTI and conditions CFI and CT. Significant differences was also found in the comparison between condition CF and condition CFI. All these match our hypotheses regarding to the effects of 2D intersection image and different control modes of a clipping plane (fixed or arbitrary). Moreover, there were significant differences in the error rate between condition M and conditions CF and CT. Significant differences exist between condition C and conditions CF and CTI as well.

In absolute terms, the proportion of underestimation varied between 20% and 45%, with 0.24, 0.22, 0.42, 0.25, 0.20 and 0.45 for the M, C, CF, CT, CFI and CTI conditions, respectively. Although mistakes come from both underestimation and overestimation of the number of sizes for sphere objects (Figure 4.17), obviously the proportion in each condition is different.

### Accuracy in the connectivity task

In the connectivity task, subjects answered two questions: the total number of curved tubes in a volume (numerosity question), and which regions of the volume the longest tube passed through (spatial region question).

The results of the accuracy analysis in case of the numerosity question are reported

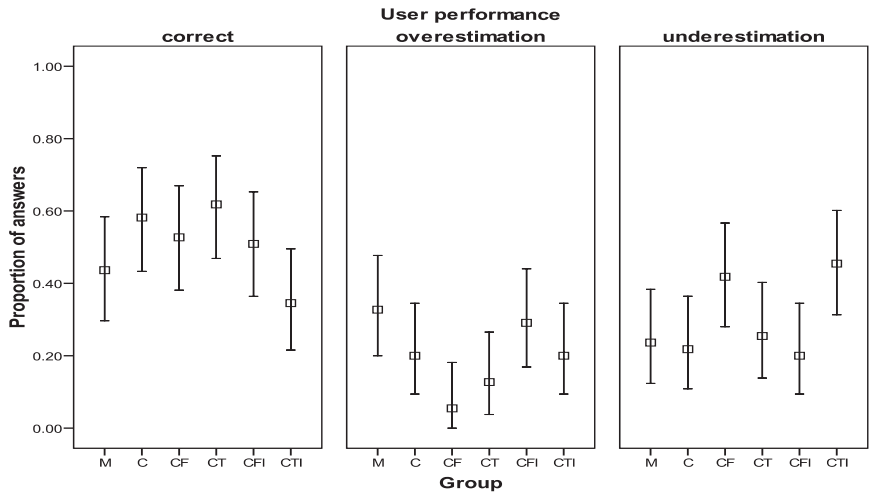



Figure 4.17: $\widehat{P_e}$ for the size task based on estimation difference.

Table 4.7: Fisher's Exact Test for counting the total number of curved tubes

| Group pair | Results |
|---|---|
| ⋆ CT vs. CTI | **0.009** |
| C vs. CFI | **0.003** |
| C vs. CTI | **0.003** |
| M vs. CTI | **0.015** |
| M vs. CFI | **0.015** |
| CT vs. CFI | **0.009** |

in Table 4.7. Significant difference was found between condition CTI and condition CT, which matches one of our hypotheses. In addition, the CTI condition was significantly more accurate than the C and M conditions, while the two latter conditions performed similarly. The analyses also reveal significant differences between condition CFI and conditions M, C and CT. In absolute terms, the error rates $\widehat{P_e}$ in the M, C, CF, CT, CFI and CTI conditions were 0.455, 0.509, 0.382, 0.473, 0.218 and 0.218, respectively. subjects in the CTI and CFI conditions were quite accurate ($\widehat{P_e}$ below 25%) to find out all the curved tubes, while $\widehat{P_e}$ was above 50% in the C condition. The 95% confidence intervals of $\widehat{P_e}$ for six conditions are shown in Figure 4.15d.

Further analysis based on task condition in terms of the number of curved tubes provided

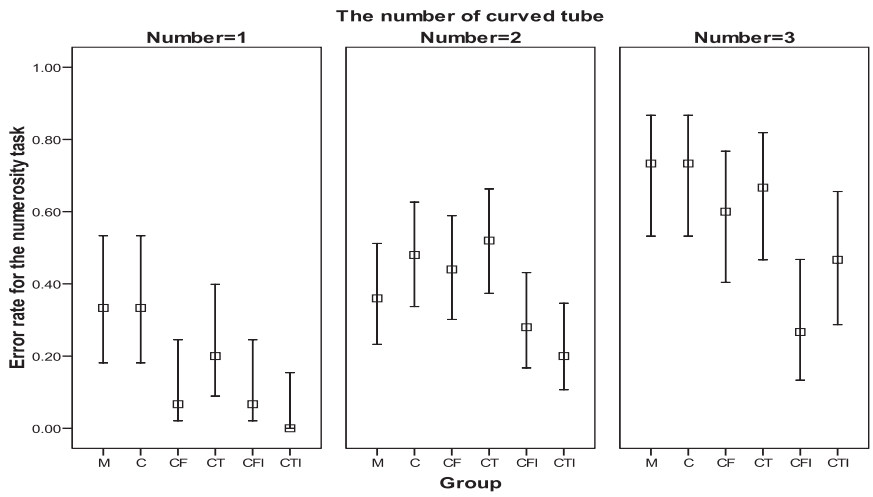



Figure 4.18: $\widehat{P_e}$ for the counting task according to the number of curved tubes.

more detail about how subjects in the six conditions differed in performance (Figure 4.18). Irrespective of the number of curved tubes in a trial, subjects in the C condition always had a higher error rate than their counterparts in the other five conditions. In case of one curved tube, subjects in the CTI condition were significantly more accurate than their counterparts in the M and C conditions. When there were two curved tubes, subjects in the CTI condition were significantly more accurate than their counterparts in the CT condition. When there were three curved tubes, subjects in the CFI condition were significantly more accurate than their counterparts in the M and C conditions.

Table 4.8: Fisher's Exact Test for $\widehat{P_{FN}}$ in locating the longest curved tube

| **Group pair** | **Results** |
|---|---|
| $\star$ CT vs. CTI | **0.001** |
| C vs. CFI | **0.000** |
| C vs. CTI | **0.001** |
| M vs. CFI | **0.009** |
| M vs. CTI | **0.005** |
| CT vs. CFI | **0.002** |

For the spatial region question, the answers are analyzed using the method of ROC. Details regarding how to code the answers can be found in the previous chapter (see Table 3.7).

The performance in terms of *false negative* $\widehat{P_{FN}}$ is summarized in Table 4.8. subjects in condition CTI performed significantly better than their counterparts in condition CT, which verifies one of our hypotheses. Significant differences were also found between condition CFI and conditions M, C and CT. In addition, the experiment reveals the significant differences between condition CTI and conditions M and C. In absolute term, the *false negative* $\widehat{P_{FN}}$ were 0.22, 0.26, 0.17, 0.23, 0.09 and 0.08 for the M, C, CF, CT, CFI and CTI conditions respectively, which were all relatively low. This is illustrated in Figure 4.19.

Table 4.9: Fisher's Exact Test for $\widehat{P_{FP}}$ in locating the longest curved tube

| **Group pair** | **Results** |
|---|---|
| $\star$ CT vs. CTI | **0.021** |
| C vs. CF | **0.028** |
| C vs. CFI | **0.046** |
| C vs. CTI | **0.002** |

The analyses for the *false positive* $\widehat{P_{FP}}$ are summarized in Table 4.9. In this case, the difference between the CT condition and the CTI condition still was shown to be statistically significant, which matches one of our hypotheses. Moreover, there were also significant differences between the C condition and the CF, CFI and CTI conditions. The performance in terms of *false positive* $\widehat{P_{FP}}$ was pretty good , evidenced by the low value, i.e., 0.05, 0.08, 0.04, 0.07, 0.04 and 0.03 for the M, C, CF, CT, CFI and CTI conditions, respectively. Subjects in condition C had a higher probability of making such errors compared to their counterparts in the other five conditions. This is illustrated in Figure 4.19 as well.

The overall performance of each group can be quantified by attributing costs to both the *false negative* ($\widehat{P_{FN}}$) and the *false positive* $\widehat{P_{FP}}$ (false alarm). Subjects in the C condition was the least accurate in both finding the regions the longest curved tube passes through and ignoring the regions that the longest curved tube does not pass through. Subjects in the CTI and CFI conditions were more accurate than their counterparts in the other four conditions, with almost the same percentage of the *false negative* $\widehat{P_{FN}}$. The frequencies of

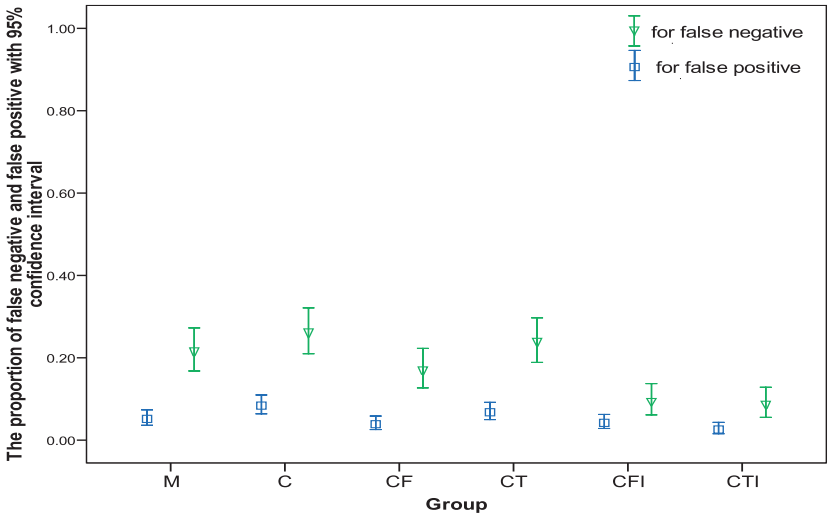


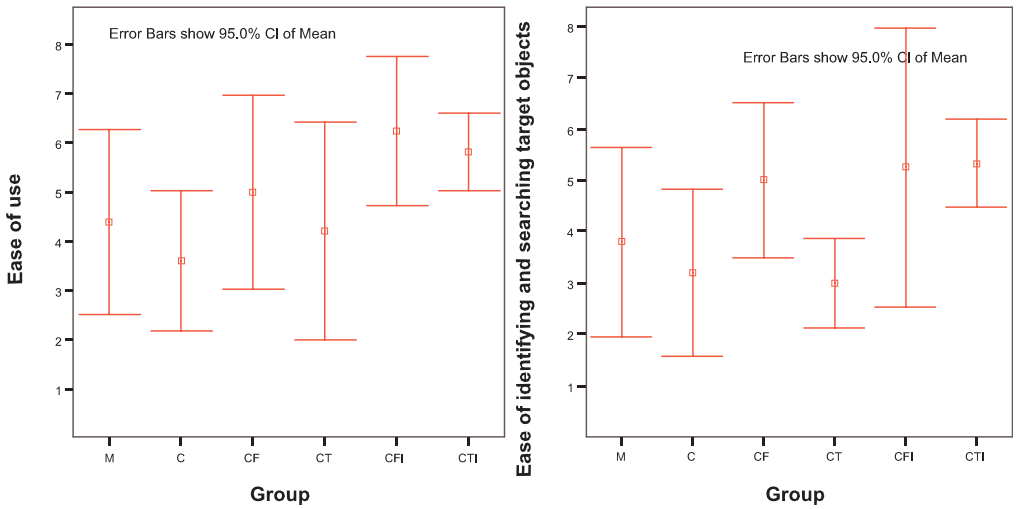

Figure 4.19: $\widehat{P_{FN}}$ and $\widehat{P_{FP}}$ for the different experimental conditions, all results are divided by condition (M, C, CF, CT, CFI and CTI), error bars represent 95% confidence interval.

the *false positive* $\widehat{P_{FP}}$ in all six conditions were approximately equal, although in the CTI condition it was still the lowest in this perspective. Therefore, subjects in the CTI condition was the most accurate. These results indicate that adding a 2D intersection image helps to reduce the *false negative* error ($\widehat{P_{FN}}$) markably.

### 4.5.3 Subjective Results

Subjective results were obtained through analyzing the post-experiment questionnaire. The questionnaire was similar to the one used in the previous chapter with moderate modification by adding additional questions about the tangible user interfaces (see Appendix C). Most of the questions used a standard seven point rating scale (some used a five point rating scale). The answers indicated that overall, subjects preferred the inclusion of a virtual clipping plane controlled by a tangible planar frame and the corresponding 2D intersection image, due to perceived ease of use, usefulness for the tasks, and ease of learning. The findings over the following categories of questions are summarized.

**Overall perception of a VR system with tangible interfaces.** The first category

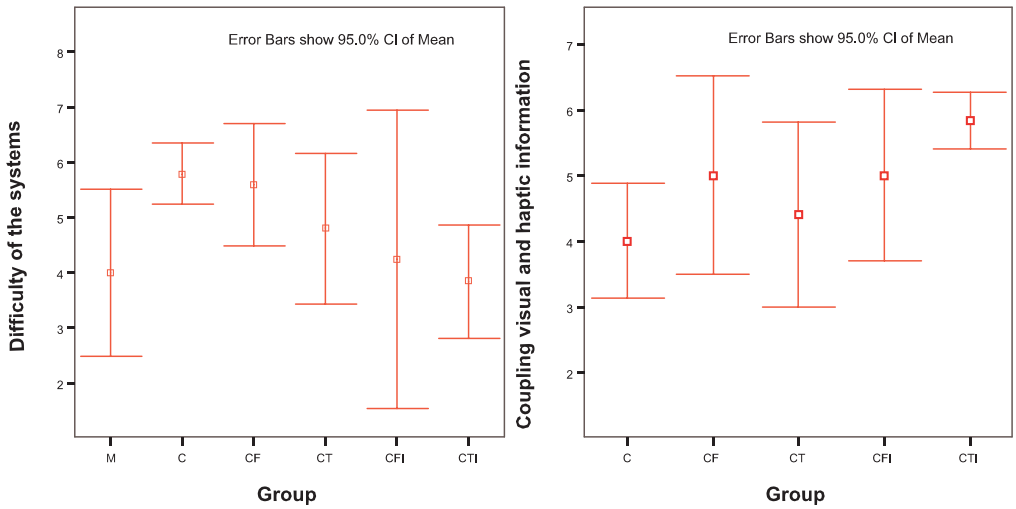



(a) (b)

Figure 4.20: Mean values for the different usability questions of different interfaces, all results are divided by condition (M, C, CF, CT, CFI, CTI), error bars represent 95% confidence interval: (a) mean rank for the ease of use; (b) mean rank for the ease of identifying and locating individual shapes.

of questions in the questionnaire concerns the perceived properties and characteristics of a VR system with tangible interfaces, including the characteristics of immersion, presence, depth cues, and spatial relationships. For most of the questions in this category, there were no significant differences among the six conditions. However, subjects did appreciate the inclusion of a 2D intersection image and the clipping plane function, compared to only working with a cube or a mouse. It was reflected by the answers for the questions "the sense of immersion" and "the spatial relation between objects", in which the CTI and CFI conditions were ranked higher than the other four conditions.

**Usability of a VR system with a tangible interface.** The ease of learning and using an interface is the main focus of this category. Answers to the question: "how well can you examine objects from multiple viewpoints" were similar for subjects from every condition, indicating the act of observing a data set from place to place was judged to be equivalently natural and easy. There were no obvious differences on the question about system delay. No subject from any condition complained about the resolution or the frame rate. For the question: "how easy to use is the interface", there was a significant difference




(a) (b)

Figure 4.21: Mean values for the different usability questions of different interfaces, all results are divided by condition (M, C, CF, CT, CFI and CTI), error bars represent 95% confidence interval: (a) mean rank for the difficulty of using the interfaces; (b) mean rank for the degree of coupling visual and haptic information (without mouse condition).

among these six conditions due to the significant difference between conditions C and CTI, $F(5, 24) = 2.766, p = 0.041 < 0.05$. The CFI condition was the highest, with the absolute rankings of 4.6, 3.6, 5.0, 4.2, 6.2 and 5.8 for the M, C, CF, CT, CFI and CTI conditions respectively (Figure 4.20a). Especially the inclusion of the 2D intersection image seems to improve perceived ease of use. There was also a significant difference on the question: "How easy to identify the individual shape and location of an object", $F(5, 24) = 3.235, p = 0.023 < 0.05$, due to the significant difference between conditions CT and CTI. The CTI condition ranked as the easiest among the six conditions, with scores of 4.4, 3.2, 5.0, 3.0, 5.2 and 5.4 for the M, C, CF, CT, CFI and CTI conditions respectively (Figure 4.20b). As described before and also reflected in the feedback from the subjects, the experimental data sets are regarded as very dense. The subjects were also asked to rank the difficulty they experienced in each condition while carrying out the tasks with the data sets. There was a significant difference with this question, $F(5, 24) = 3.025, p = 0.030 < 0.05$ (post hoc comparison $p = 0.008 < 0.05$ between condition C and condition CTI). In absolute terms, the C condition is regarded as the most difficult to use, with scores of 4.0, 5.8, 5.6,

4.8, 4.2, and 3.8 in difficulty for the M, C, CF, CT, CFI and CTI conditions respectively (Figure 4.21a). With the presence of tangible interfaces in those five conditions, the subjects were asked to evaluate how well the visual and passive haptic feedback worked together. The feedback on this question indicates that there was a significant difference, $F(4, 24) = 3.763, p = 0.019 < 0.05$, due to the significant difference between conditions C and CTI, (post hoc comparison $p = 0.017 < 0.05$), with scores of 4.0, 5.0, 4.4, 5.0, 6.0 for the C, CF, CT, CFI and CTI conditions, respectively (Figure 4.21b).

**Task difficulty and user performance.** Subjects were asked to select the easiest and most difficult task both before the experiment and after they had finished all the trials. Up front, 55% of the subjects thought counting the number of shapes was the easiest task, and 25% of the subjects thought the task of counting the number of curved tubes would be the easiest one. At this stage, finding out the densest region was regarded as the most difficult task by 50% of the subjects, while another 30% of the subjects thought the spatial region question about the longest curved tube was the most difficult one. After the experiment, the easiest task was considered to be the one of judging the number of shapes (35%) or the one of counting the number of curved tubes (30%). At this stage, the most difficult task was considered to be the density task (55%) while the size task was a remote second (20%).

Additional findings can be deduced from these subjective rankings about the difficulties of different tasks, especially when compared to actual user performances. The shape task was regarded as the easiest one by many subjects even though the percentage falls after the experiment (55% to 35%). It is consistent with actual user performances (the accuracies for all conditions are around 70%). It explains the reason that there is no association between the user performance (error rate) and interface choices: the task is easy therefore interfaces in different experimental conditions do not differ much. Although counting the number of curved tubes is another task that was regarded as relatively easy, the actual performance was quite varied between conditions. It indicates that the user performance with this task was influenced by different interface set-ups.

At least 50% of the subjects thought the density task was the most difficult one no matter whether or not it was judged at the beginning or in the end of the experiment. This feedback is consistent with actual user performance as well (see Figure 4.15a), which verifies that the density task is indeed a difficult one and subjects using different interfaces do not differ

markably in their performance. Another noticeable fact is that the subjects' performance on the size task was quite low in all six conditions although it was not regarded as difficult. The fact is that subjects very often under-estimated or over-estimated the number of sizes.

## 4.6 Discussion

The general belief is that a 3D interface should be more suitable than a 2D interface for 3D tasks since it provides a user with simultaneous control over more DOFs. However, several experiments have proven that this is not necessarily true for all 3D tasks, for example the docking task [BKLP04]. Masliah and Milgram [MP00] demonstrated that users manipulated rotational and translational DOF as separate subsets in a 6 DOFs docking task. Zhai [Zha98] also pointed out that with free moving 6 DOFs devices fatigue becomes a significant problem for a 3D docking task. There are a number of possible reasons for these findings. First, the higher numbers of DOF create extra difficulties for operating a 3D interface, compared to a 2D one. It is easier to control a mouse to move from one point to another in 2D than to move a 3D input device to reach a position in 3D space. Second, the accuracy also differs for current 2D and 3D interaction devices. 6 DOFs interfaces are difficult to control precisely because of lack of coordination and no support for a user's arm. To find a proper compromise between such conflicting requirements, for example more DOFs and higher precision, is a big challenge for 3D interface designers [WH99].

### 4.6.1 Response Time

A graph of the time spent on each trial as a function of the trial number indicates response time decreases as the subjects complete more trials (Figure 4.22). The first two trials for all conditions except for the mouse condition show the strong learning tendency. After the first two trials, user performances stabilizes in time.

In the previous analyses, the response time needed to complete trials in the CTI condition was significantly longer, compared to in the other five conditions. One explanation is that in the CTI condition subjects were required to use both the cube and the tangible frame with two hands, and at the same time to observe the intersection image and the 3D result on two separate screens. In this way, it obviously takes more time to explore the

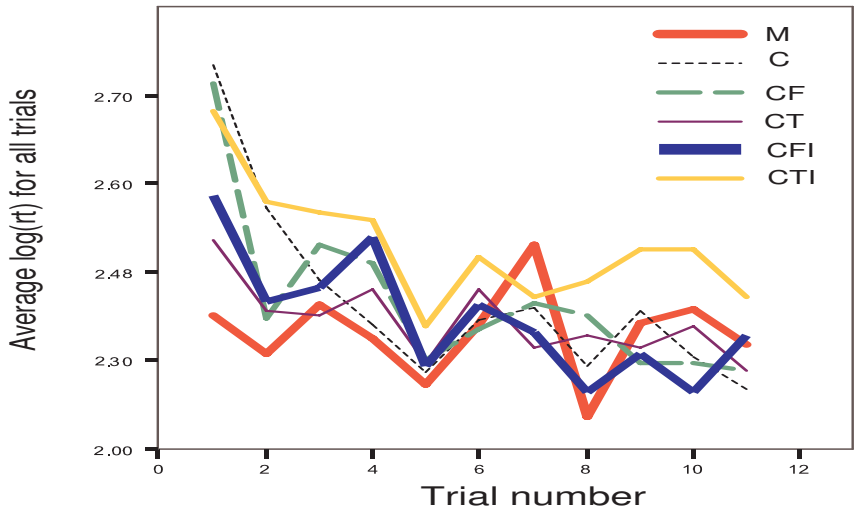



Figure 4.22: Time curves for all conditions, in the order that subjects completed the trials.

experimental data compared to manipulating the cube with one hand (the C, CF and CFI conditions) or working with two hands but only observing the 3D result on a single screen (the CT condition). Another critical issue was the reported difficulty in the CTI condition to reposition the tangible frame precisely where the subjects had previously seen a target object (for example a curved tube) within a volume. Because of the high density of the data sets, subjects would often have to re-search the volume for objects they had located before, but had "lost" as they moved the tangible frame by chance. This can be solved by introducing a clutching mechanism in the future system. Finally, subjects may simply be more familiar with holding a cube object than grasping and manipulating a frame. In the other five conditions, none of them was significantly faster or slower than another. As pointed out by Zhai [Zha98], a user typically takes tens of minutes, a significant duration for learning computer interaction tasks, to gain controllability of 3D input devices. It may take hours of practice to approach the level of constant operational speed. The author believes that the longer response time in condition CTI will decrease without lowering the performance if the subjects spent more time on practice.

### 4.6.2 3D Manipulation with 2D or 3D Interfaces

One basic interaction activity when manipulating the visualization result of 3D data is the modification of the viewpoint, which can be achieved through a either 2D or 3D interface. Together with investigating the effects of different 3D tangible interfaces on data analysis tasks, a comparison of the effects of manipulating 3D viewing on the same data analysis tasks was carried out between 2D and 3D input devices. The 2D input device in this study is a mouse.

Previous Fisher's exact analyses do not reveal any significant differences in user performance (the response time $rt$ and the error rate $\widehat{P_e}$) for all the tasks between condition M and condition C, while $rt$ was shorter for the M condition. The fact that a mouse is more familiar to the subjects can be a reason for that. In order to further investigate the effects of using a 2D or 3D interface to manipulate the viewpoints on all the data analysis tasks, a linear regression model was used to analyze the relationship between the response time $rt$ and the interface choice (a 2D mouse or a 3D cube). A binary logistic regression model [FLP03] is applied to analyze the relationship between the error rate $\widehat{P_e}$ and the interface choice for each task. The regression equations are the follows:

$$\log(rt) = \beta_0 + \beta_1 \times I$$

and

$$\ln(\frac{\widehat{P}}{1-\widehat{P}}) = \alpha_0 + \alpha_1 \times I$$

I represents the interface variable, which can be a 2D mouse or a 3D cube interface.

The results of regression analysis (Table 4.10) confirm the previous conclusion that there is no direct relationship between the user performance (the response time and accuracy) and different interfaces choices (either condition C or condition M). Traditional 2D interfaces require a user to mentally reason the spatial relationships between the objects, which solely relied on the rendered scene on the 2D screen. 3D interfaces (here refers to a cube) provides more DOFs for a user while manipulating the viewpoint. However, those additional DOFs alone did not shorten the response time and help to improve the user performance for all the tasks. The author concludes that controlling the view transformation with different

Table 4.10: Results of regression analyses for $rt$ and $\widehat{P_e}$ (2D versus 3D interface)

| **Dependent Variable** | **Regression Equation and Odds Ratio**[1] |
|---|---|
| $log(rt)$ | $2.381 - 0.11 \times I$, $p = 1.91$ |
| $\widehat{P_e}$ for density task | $-0.095 + 0.058 \times I$, $p = 0.445$, $Exp(B) = 1.06$ |
| $\widehat{P_e}$ for shape task | $1.034 + 0.04 \times I$, $p = 0.654$, $Exp(B) = 1.041$ |
| $\widehat{P_e}$ for size task | $0.447 - 0.117 \times I$, $p = 0.128$, $Exp(B) = 0.889$ |
| $\widehat{P_e}$ for counting tube | $-0.080 + 0.044 \times I$, $p = 0.567$, $Exp(B) = 1.045$ |
| $\widehat{P_{FN}}$ for longest tube | $0.987 + 0.051 \times I$, $p = 0.383$, $Exp(B) = 1.052$ |
| $\widehat{P_{FP}}$ for longest tube | $2.287 + 0.104 \times I$, $p = 0.113$, $Exp(B) = 1.110$ |

DOFs is not the determinant factor for all analysis tasks performed in this study no matter whether or not it is realized by a 2D interface or 3D interface. Therefore, additional DOFs introduced by a 3D input device will not be substantially useful for analyzing volumetric data.

Table 4.11: Each condition and its interface set-up (excluding the M condition)

| **Condition** | **Clipping plane (CL)** | **Fixed or arbitrary (FI)** | **Intersection image (IS)** |
|---|---|---|---|
| C | No (0) | No (0) | No (0) |
| CF | Yes (1) | Fixed (1) | No (0) |
| CT | Yes (1) | Arbitrary (0) | No (0) |
| CFI | Yes (1) | Fixed (1) | Yes (1) |
| CTI | Yes (1) | Arbitrary (0) | Yes (1) |

[1]Exp(B) is the odds ratio, B is the coefficients; p is the significance of the regression model.

### 4.6.3 Clipping Plane Function and 2D Intersection Image

The experimental configuration enables us to separately substantiate the effects of: a) including a clipping plane function, b) different control mechanisms and c) including the corresponding 2D intersection image. Each condition corresponds to one combination of factors, as shown in Table 4.11. A multiple linear regression model was adopted to analyze the relationship between the response time $rt$ and the interface set-ups A multiple binary logistic regression model [FLP03] was applied to analyze the relationship between the error rate $\widehat{P_e}$ and the interface set-ups for each task. The regression equations are the follows:

$$\log(rt) = \beta_0 + \beta_{CL} \times CL + \beta_{FI} \times FI + \beta_{IS} \times IS + \beta_{FI\_IS} \times FI \times IS$$

and

$$\ln(\frac{\widehat{P}}{1-\widehat{P}}) = \alpha_0 + \alpha_{CL} \times CL + \alpha_{FI} \times FI + \alpha_{IS} \times IS + \alpha_{FI\_IS} \times FI \times IS$$

CL represents the clipping plane variable, which can be true or false (with or without a clipping plane). FI represents the variable that describes the control mode of a clipping plane, which can be fixed or arbitrary. IS represents the intersection image variable, which can be true or false (with or without an intersection image). The derived models based on the experimental results are shown in the Table 4.12.

#### With and Without Clipping Function (No Intersection Image)

The original hypothesis assumed that the inclusion of a clipping plane function would help a user perform all the tasks. The previous Fisher exact tests indicate that a clipping plane function alone has no significantly positive effects on the user's performance for most of the tasks, no matter whether or not it is fixed or arbitrary, under the condition that a 2D intersection image was not present. The regression analysis confirms this conclusion. There is a significant effect on the estimation difference only for the size task after introducing a fixed clipping plane function into the baseline system. For identifying the regions that the longest curved tube passes through, adding a virtual clipping plane in a fixed position (condition CF) significantly lowered the rate of the *false negative* $\widehat{P_{FN}}$.

Table 4.12: Results of regression analyses for $rt$ and $\widehat{P_e}$

| **Dependent Variable** | **Regression Equation and Significance of Coefficients** |
|---|---|
| $rt$ | $2.37 - 0.009 \times CL + 0.036 \times FI + 0.124 \times IS - 0.158 \times FI \times IS$<br>$p_{IS}$[2] $= 0.000$, $p_{FI\_IS} = 0.001$ |
| $\widehat{P_e}$ for density task | $-0.036 + 0.518 \times CL - 0.152 \times FI + 0.794 \times IS - 0.565 \times FI \times IS$ |
| $\widehat{P_e}$ for shape task | $1.075 + 0.312 \times CL - 0.118 \times FI + 1.159 \times IS - 1.682 \times FI \times IS$<br>$p_{FI\_IS} = 0.03$, $Exp(B)_{FI\_IS} = 0.186$ |
| $\widehat{P_e}$ for size task | $0.33 + 0.152 \times CL - 0.373 \times FI - 1.121 \times IS + 1.048 \times FI \times IS$<br>$p_{IS} = 0.005$, $p_{FI\_IS} = 0.05$, $Exp(B)_{IS} = 0.326$, $Exp(B)_{FI\_IS} = 2.852$ |
| $\widehat{P_e}$ for counting tube | $-0.036 + 0.146 \times CL + 0.373 \times FI + 1.167 \times IS - 0.373 \times FI \times IS$<br>$p_{IS} = 0.006$, $Exp(B)_{IS} = 3.213$ |
| $\widehat{P_{FN}}$ for longest tube | $1.038 + 0.123 \times CL + 0.43 \times FI + 1.22 \times IS - 0.525 \times FI \times IS$<br>$p_{IS} = 0.001$, $Exp(B)_{IS} = 3.388$ |
| $\widehat{P_{FP}}$ for longest tube | $2.391 + 0.231 \times CL + 0.59 \times FI + 1.01 \times IS - 1.092 \times FI \times IS$<br>$p_{IS} = 0.017$, $Exp(B)_{IS} = 2.743$ |

### Fixed and Arbitrary Clipping Plane (No Intersection Image)

The author hypothesized that an arbitrary clipping plane controlled by a tangible planar frame would be more helpful for a user to perform all the analysis tasks compared to a fixed clipping plane. The results showed indeed that the error rates for all the tasks were different between these two conditions (condition CF and condition CT). However, the differences are not significant, which indicates that whether or not a clipping plane is fixed or arbitrarily controlled is not an important factor. The regression analysis further confirms that there is no direct relationship between the performance (the response time and the error rate) and the control mechanism of a clipping plane (fixed or arbitrary).

[2]only significant $p$ is shown ($p < 0.05$)

### With and Without Intersection Image

In the earlier hypothesis, it was believed that the inclusion of a 2D intersection image would be beneficial for a user to perform all the tasks. The experimental results show that the presence of an intersection image has different effects on different tasks. The Fisher's exact tests indicate that adding an intersection image significantly increased the error rate for the size task when a clipping plane was controlled by a tangible planar frame. However, for counting the number of curved tubes and locating the longest tube, the error rates were significantly lower after adding an intersection image. There is no significant effect on the user performance after introducing an intersection images for all the tasks when the clipping plane is in a fixed position. The regression analysis further reveals that there is a negative relationship between adding an intersection image and the response time, which means that the presence of an intersection image prolongs the response time. It also confirms that there is a direct positive relationship between the error rate and the inclusion of an intersection image for the tasks of counting the number of the curved tubes and locating the longest tube. But, such relationship is negative for the size task.

### Both Clipping Plane and Intersection Image

Prior to the experiment, the author supposed that a subject would perform better while using a clipping plane that is controlled by a tangible planar frame compared to a fixed clipping plane, under the condition that the 2D intersection image is present. The regression analysis confirms that there is a significant interaction between adding a clipping plane function and including a 2D intersection image, which means that the combination of these two features has significant impacts on the user performance. There is a positive relationship between including an intersection image and the response time while a clipping plane function is present, which means that the presence of an intersection image shortens the response time. The regression result also reveals that the interaction between a clipping plane function and a 2D intersection image has significant but opposite effects on the shape and size tasks. This coincides with the results of Fisher's exact test. The Fisher's exact test indicates that the error rates in the case of using a tangible frame was significantly lower than in the case of using a fixed-position clipping plane for the shape task, while an intersection image was included. On the contrary, for the size task such combination has

nearly significant effect on the error rate.

### 4.6.4 Comparison with Other Designs

There are some differences between the design presented here and the previous existing designs of tangible user interfaces. Firstly, the cube-shaped interaction device is a generic (abstract) tactile device that can be used in many diverse applications. This is different from earlier studies where a dedicated device with a data-related shape was used to position and rotate a data set. For example, a head prop was used by Hinckley to resemble his anatomical data. This created some confusion with his subjects, while the position of the nose in the head prop did for instance not necessarily coincide with the position of the nose in the virtual data. In De Guzman's prototypes, the physical models of a human body may exercise similar problems. Secondly, the prototypes presented here can provide more depth cues to users than what was done in the experiment, because they can also work in stereoscopic mode. Ware [WAB93] has experimentally established the importance of a stereoscopic display in (certain) 3D tasks. Stereo was not included in this study because the focus of the experiment was on the effects of tangible devices, not on verifying the stereoscopic effect. The experiment does hence not address possible interaction effects between stereoscopic displays and 3D tangible objects. Thirdly, the optical tracking used in the interfaces allows for wireless interaction devices, hence avoiding possible interaction problems associated with the cables existed in the interfaces of Hinckley and De Guzman. Even though the optical tracking can suffer from tracking loss due to occlusion, the experimental set-up and the design of the tracking patterns on the interfaces in this study helped to avoid this problem as much as possible. Fourthly, the hardware-accelerated texture mapping allows for real-time 3D interaction with volumetric data. Other interfaces can only render geometrical models or slices in real time.

### 4.6.5 3D Clipping Task Analysis

People do not necessarily posses similar skills for 3D viewing and manipulation. These individual differences mean that some people will probably experience more difficulties than others while performing 3D intersection task. Many Intelligent Quotient (IQ) tests actually contain a section that is specifically aimed at measuring an individual's ability to perform

such 2D to 3D mappings. The subjects in the experiment were randomly assigned to the different conditions in the hope of averaging out individual differences in spatial ability. No spatial acuity test was performed with the subjects.

Often the reason to include a clipping plane function (especially in 3D) is that it enables the inside structure of observed data to be visible, which can be helpful for analysis tasks. The analysis of the 3D intersection task reveals how many and what kind of universal interaction tasks are required. Hinckley [HTP+97] recognized that intersection operations require two sub-tasks: viewing and clipping. He suggested that viewing could be subdivided into orienting and zooming. In total, ten separate parameters (yaw, pitch, roll, and zoom for the view on the data; x, y, z, yaw, pitch, and roll for a clipping plane) need to be specified. Of course, only the relative orientation and position of a clipping plane with respect to a data set (the cube) is relevant. Previous experimental results indicate that adding a clipping plane alone can be beneficial for most of the tasks (except for the size task) compared to only manipulating the viewpoint without intersection operation, but not significantly. The reason might be that interaction with a 3D clipping plane requires spatial reasoning. Spatial reasoning is a mental cognitive process that requires thinking about relationships between objects. Therefore, counteracting the benefit from including a 3D clipping plane is the increased complexity of the system.

For 3D information processing, Ware proposed the following principles[War01]:

1. Use 3D objects to represent data entities;
2. Emphasize 2D layout and use it to support navigation in 3D spaces.

In this experiment, the introduction of 3D objects (particularly the tangible planar frame) is motivated by the first principle and for the purpose of reducing complexity introduced by the 3D clipping task. Subjects did give positive feedback: the tactile or passive force feedback helps to maintain spatial awareness and identify spatial relationships. They also reported that 3D objects helped them learn how to perform the intersection task with minimal training effort, because the shape and function of the interaction elements are easily understood. In theory, using a pair of tangible interfaces with two hands can reduce the entire hierarchy of the clipping task into a single transaction that directly corresponds to the task that the user has in mind. In the study, it helped to simplify the apparent

complexity of manipulating a clipping plane relative to an object, which can be reflected by the shorter response time $rt$ in condition CT compared to condition C. At the same time, the user performance did become better as well, although not significantly.

The positive effects of a clipping plane reach a significant level for certain tasks after introducing the corresponding intersection image. This is particularly true while a clipping plane is controlled by a tangible planar frame and for tasks like counting the number of tubes and locating the longest tube. The reason might be that with the help of an intersection image, a user can re-exam or verify his/her judgements about certain objects' properties (shape, density) that are made from the 3D observation. However, the response time $rt$ became longer as well. The explanation is that the procedure of observing the 3D and 2D views together obviously takes more time. Another reason is that the chances to perform the analysis tasks better are getting higher with the presence of a 2D intersection image and subjects are willing to spend more time on the tasks, as lots of subjects reported.

In short, these facts and discussions can lead to several conclusions:

1. Introducing a clipping plane function (in 3D) enables a user to observe the inside structure of the volumetric data set, but brings additional complexity as well because of the task hierarchies of 3D clipping.

2. Using 3D objects to represent volumetric data and a clipping plane in particular can make the 3D clipping task easier to learn and improve the user performance, but not significantly.

3. A 2D intersection image next to the 3D view is a necessary feature in a 3D clipping task. However, this additional feature tends to increase the interaction time.

### 4.6.6 Comparison with Previous Study

The user performance in the experimental study is quite different from the one in the previous chapter. The hardware set-ups and the data sets in these two studies are not exactly the same, although very related. Furthermore, the sample sizes are different as well. In the first user study, each subject completed 16 trials; 11 trials were completed in the second study. Moreover, there were 13 subjects per condition in the previous one,

compared to 5 subjects per condition in this one. As pointed out by Fleiss et al. [FLP03], different sample sizes can lead to different *powers of the studies*, denoted by $\delta$, even studies nevertheless aim to detect the same level of statistical significance. In this thesis, both studies set $\alpha = 0.05$ as the significant level. For example, the error rates in the HMD VR condition and the fish tank condition are $P_{fish} = 0.38$ and $P_{HMD} = 0.60$ for the density task in the first study. The power of the study for this task with 16 trials per subject is $\delta_1 = 0.8$. In the second study, the error rates in the CTI condition and the C condition are $P_{CTI} = 0.218$ and $P_C = 0.436$. The power of the study with 11 trials per subjects for the task is $\delta_2 = 0.7$. Even though these results about *the power of the study* are rough estimate, these differences hamper a direct comparison between both studies. The discussion of relative performance between conditions in each study nevertheless reveals some valuable facts.

Table 4.13: Summary of the error rates in the previous user study of VR

| **Tasks** | **The range of error rate** | **Lowest error rate** | **Highest error rate** |
|---|---|---|---|
| Density | 38%- 62% | Fish tank | HMD |
| Shape | 26%-38% | Fish tank | HMD |
| Size | 76%-80% | HMD | Haptic |
| Number of tube | 26%-52% | Haptic | HMD |
| Longest tube ($\widehat{P_{FN}}$) | 31%-39% | Fish tank | HMD/Haptic |
| Longest tube ($\widehat{P_{FP}}$) | 17%-20% | Fish tank/ Haptic | HMD |

The error rates for all the tasks in both studies are listed in Table 4.13 and Table 4.14. The error rates were comparable in both experiments for the density task. Error rates were as low as 38% or as high as 60% in the first study. In the second study, error rates were between 30% and 50%, except for the CTI condition, where error rate was as low as 20%.

In both studies, error rates were around 20% to 40% for most of the conditions (see Table 4.13 and Table 4.14) for the shape task, even though they were slightly higher in the previous study. Using a virtual clipping plane that is controlled by a tangible frame and

adding its correspondent 2D intersection image into the baseline system (the CTI condition) lowered the error rate from 25% to 7%. The error rate only decreased from 38% to 26% when changing from a fish tank VR condition to a HMD VR condition.

The error rates were around 70% to 80% for the size task in the first study,. The error rates were around 50%, varying between 40% and 70%, in the second study (see Table 4.13 and Table 4.14). The subjects in the HMD VR condition were more accurate than in the other two conditions, but only with 4% difference maximally. Adding a 2D intersection image did not help to lower the error rates when a virtual clipping plane is controlled by a tangible frame, but increased it from 38% to 65%, which is unexpected. Similarly, adding haptic force feedback to a fish tank VR system increases the chances to be incorrect compared to working with a HMD VR system (from 76% to 80%) or a fish tank VR system (from 79% to 80%), even though not significantly. Because both adding a 2D intersection image and adding haptic force feedback to a fish tank VR system amounts to adding additional inside-out information, it is believed that such information is not useful for the size task.

Table 4.14: Summary of the error rates in the user study of this chapter

| **Tasks** | **The range of error rate** | **Lowest error rate** | **Highest error rate** |
|---|---|---|---|
| Density | 22%-51% | CTI | C |
| Shape | 7%-27% | CTI | CFI |
| Size | 38%-66% | CT | CTI |
| Number of tube | 22%-51% | CTI/CFI | C |
| Longest tube ($\widehat{P_{FN}}$) | 8%-26% | CTI | C |
| Longest tube ($\widehat{P_{FP}}$) | 3%-8% | CTI | C |

For counting the number of curved tubes, the range of the error rates in both studies is quite close, with the lowest error rate as low as 20% and the highest rate as high as 50%. In the second study, adding an intersection image significantly increases the chances of being correct (with the error rate as low as 20%) when a virtual clipping plane is controlled by a

tangible frame.

For locating the longest curved tube, two kinds of errors are discussed. The error rates $\widehat{P_{FN}}$ between different conditions are close, from 31% to 39% in the first study. In the second study, the error rates are quite different, with the lowest error rate as low as 8% and the highest rate as high as 26%.

Table 4.15: Conditions with significant effects for the tasks in the first and second user study (VR and tangible interfaces)

| **Tasks** | **First study** | **Second study** |
|---|---|---|
| Density | Fish tank vs. HMD; Haptic vs. HMD | M vs. CTI; C vs. CTI; CF vs. CTI |
| Shape | Fish tank vs. HMD; Haptic vs. HMD | CFI vs. CTI; C vs. CTI |
| Size | | CT vs. CTI; C vs. CTI |
| Number of tube | Fish tank vs. HMD; Haptic vs. HMD | CT vs. CTI; C vs. CFI; C vs. CTI; M vs. CFI; M vs. CTI; CT vs. CFI |
| Longest tube ($\widehat{P_{FN}}$) | Fish tank vs. HMD; Fish tank vs. Haptic | CT vs. CTI; C vs. CFI; C vs. CTI; M vs. CFI; M vs. CTI; CT vs. CFI |
| Longest tube ($\widehat{P_{FP}}$) | Fish tank vs. HMD | CT vs. CTI; C vs. CF; C vs. CFI; C vs. CTI |

All the significant effects in both studies are summarized in Table 4.15. Overall, the major significant effects can be identified as the following. The error rate was always the highest for all the tasks except for the size task in a HMD condition. Subjects in the fish tank condition had significantly better performance than their counterparts in the HMD condition for most of the tasks except for the size task. The haptic feedback significantly improved the user performance compared to the HMD condition for density, shape and counting the number of tubes tasks. Subjects in the CTI condition had significantly lower error rates compared to their counterparts in the C condition for most of the tasks except for the size task in the second study. Subjects in condition CTI was significantly more

accurate than their counterparts in condition CT for counting the number of tubes and locating the longest tube, which indicates a 2D intersection image, is useful for this kind of task. Subjects in condition CFI was significantly more accurate than their counterparts in condition C for the same tasks. In addition, subjects in both the CTI and CFI conditions are significantly more accurate than their counterparts in the condition M for counting the number of tubes and locating the longest tube ($\widehat{P_{FN}}$).

The shape task is regarded as the easiest one for both studies. The absolute error rates were highest for the size task in both studies even though the subjects did not think of the task as difficult. Even though the subjects did not think that counting the number of curved tubes was an easy task, the user performances were good for both studies in absolute term.

### 4.6.7 Guided Search Model

There are related studies in psychology that try to understand the routine visual behavior of a user. Among these studies, the "Guided Search" model from Jeremy Wolfe is a quite relevant one that can be used to analyze the behaviors of the subjects in both studies [Wol94]. In both studies, subjects were asked to search particular objects within the volumetric data in order to collect information and make judgements regarding their sizes, shapes, densities, and etc. It is similar to the general visual behavior that tries to find one item in a visual world that is filled with other distracting items.

As a very important ability of humans, performing visual search has been studied extensively in the past 15 years. The physical and biological limitations of the humans' neutral system lead to the fact that performing all visual functions at all locations in the visual field at the same time is extremely difficult. There are two kinds of solutions for a visual system. Discarding input is the first approach. This is reflected by the anatomy of the visual system. At the fovea, the retinal image is full of details. However, sampling in the periphery is rather coarse. The second approach is to process selected information only. Particular visual processes are restricted to particular portions of the visual field.

Neisser indicated that there are parallel processes and limited-capacity processes in visual search behavior [Nei67]. Parallel processes continuously take place over a large portion of the visual field. However, the portion of the visual field for limited-capacity processes is

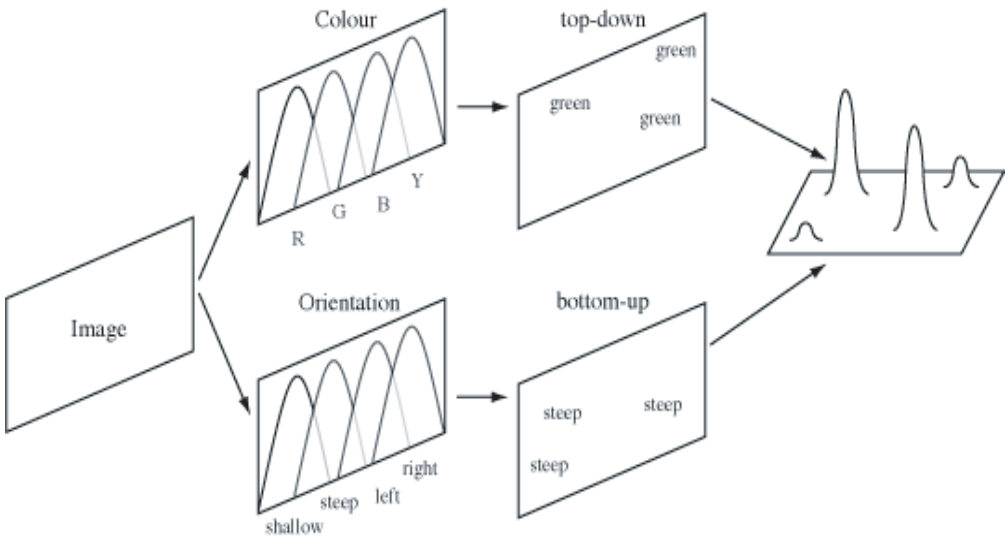



Figure 4.23: The guided search model. (Photograph reprinted from [Wol94]).

relatively small. In order to cover the entire visual field, these limited-capacity processes need to work in a serial mode and need to move from one place to another. It is well accepted that humans deploy the limited resources of their visual system in a controlled fashion. Attention controls the deployment of the resources. Either the properties of the visual stimuli or the demands of the "user" of the visual system decide the way of deployment. A subject's wish to search for a specific visual stimulus will have effect on this deployment.

A brief description of the "Guided Search" procedure proposed by Wolfe is that, at the early stage of searching information, features from all places are processed in parallel (Figure 4.23). Only a limited amount of information is extracted from the visual inputs, however. An active feature map is created from the extracted information. The feature map embodies the estimation by the visual system about locations of target items. The following visual processes deploy attention in a serial mode from item to item in order to find the target item or assert there is no target item. This is a very efficient approach for deploying limited attention resources. The important mechanism underlying this model is pre-attentive processing, because it takes place before conscious attention. Pre-attentive processing determines which visual objects an observer's attention turns to. Many experiments have been carried out to test which kinds of features are processed pre-attentively. The features that are pre-attentively processed can be divided into four categories [WF96]:

form, color, motion and spatial position. Within the spatial position category, the information can be one of these: stereoscopic depth, shape from shading, position on the XY plane and motion. Nakayama and Silverman indicated the combination of stereoscopic depth and color, stereoscopic depth and movement also can be pre-attentively processed [NS86]. He and Nakayama showed that humans could not easily focus attention across isodepth loci unless they were part of a well-formed surface with locally coplanar elements [HHKP96]. Yet, humans can easily spread their attention selectively across well-formed surfaces that span an extreme range of stereoscopic depths.

The "Guided Search" model and the related theory of pre-attentive processing indicate that organizing the visualization results of 3D data should understand and take advantage of the pre-attentive process in guided search. In particular, Wolfe's "Guided search" model has illustrated how basic features, such as color, stereoscopic depth, size, and etc, play a part in a guided search. In terms of 3D interaction technique and interface design, researchers can use this model as a theoretical basis and collect more quantitative evidence through experimental studies on how a basic feature or a combination of several basic features from an interface can influence a user's performance for a 3D analysis task. This evidence can further help to improve the design of a proper interface that can assist in building the feature map during the search process in order to shorten the visual search time.

The studies of both different VR set-ups and the tangible interfaces for a clipping plane and its related intersection image are such attempts. More specifically, like in this chapter, a clipping plane function is regarded as a mechanism to help reduce the amount of 3D information that a user needs to process during the interaction and then reduces the amount of distractors. The inclusion of an intersection image is another way that helps to form active feature maps based on the tasks during a visual search. For example, more accurate feature maps regarding the location and orientation of different curved tubes can be built after introducing both a clipping plane function and a 2D intersection image, which are demonstrated by the improved user performance for the tasks of counting and locating the curved tubes in the condition CTI compared to the conditions C and M.

## 4.7 Conclusion

In this chapter, the effects of both a clipping plane and different tangible interfaces for controlling a clipping plane on data analysis tasks are investigated. Firstly, three prototypes with 3D interaction devices have been proposed. The motivation of such designs comes from the experimental results in the previous chapter. The capability of having an overview of the data set in a fish tank VR system had been identified as an important feature for the observation of experimental volumetric data. Since subjects with a fish tank VR system achieved better performance than with an immersive HMD based VR system in terms of time and accuracy, the author asked the question whether or not the user performance of a desktop VR system can be further improved by adding tangible elements to the interaction. Therefore, a desktop VR system was selected as the baseline testbed in the experimental study in order to quantitatively measure the added value of tangible interfaces. Because of the dense nature of the data sets used, the author hypothesized that adding a 3D clipping plane function, and a tangible planar frame for controlling a clipping plane or/and a 2D intersection image would all help a user to understand this kind of complex data.

The major results included several important findings.

- Overall, working with a clipping plane that is controlled by a tangible planar frame together with its 2D intersection image, significantly improved a user's performance for the density and shape tasks compared to only working with a cube-based 3D interface. However, such design had a significant negative effect on the size task. Therefore, for the density and shape tasks, an arbitrary clipping plane with a 2D intersection image should be integrated into a 3D interface.

- A clipping plane (either fixed or arbitrary) with its corresponding 2D intersection image significantly improved the accuracy for counting the number of tubes and locating the longest curved tube compared to working with a cube-based 3D interface. For the same tasks, the user's performances with such designs (a fixed or arbitrary clipping plane and an intersection image) were significantly better than a mouse-based 2D interface as well. These results indicate that working with a 3D or 2D interface alone to change the viewpoint is not sufficient for answering property questions about the data precisely. The positive effects of introducing a clipping plane on locating and

searching tubes can be augmented to a significant level after adding an intersection image. Such significant interaction effect between a clipping plane and an intersection image was confirmed by the regression analysis.

- Regarding the response time, subjects in the CTI condition spent significantly more time than in the other five conditions (the response time in the condition CTI is 23% longer than in the condition C). The regression analysis reveals that the inclusion of an intersection image is the major factor that leads to longer response time. One explanation is that it takes more time for a subjects to use the cube and the tangible frame with both hands, and at the same time observe the intersection image and the 3D result on two separate screens during the experiment. Because subjects did not have any experience with using a tangible frame to control a clipping plane, it is believed that response time can be shorter with practice.

- Feedback from the subjects indicate that the tasks present different difficulties. Finding the densest region was always regarded as the most difficult one. Counting the number of shapes was ranked as the easiest. Although locating the longest curved tube was regarded as a relatively easy task in the beginning of the experiment, some subjects realized that it was indeed a difficult one in the end. User performance in terms of error rates mostly coincides with subjective evaluations about the difficulty of the tasks.

Most of the data sets in both studies are high density volumes. The tasks performed are similar to a visual search problem in psychology. A visual search is the routine visual behavior by which one finds specific items in a visual world that is filled with other distracting objects. The "Guided Search" model proposed by Jeremy Wolfe is a way to explain how the visual system utilizes limited resources to perform the search in an economic way. This model indicates that in the early stage of a search, there are several parallel feature processes that collect information from observed images. The processed information then forms activation maps. Based on these maps, attention can be guided from location to location in the visual field to refine the search. According to this model, a successful 3D interface and interaction technique should be able to help the parallel feature processes and assist in the formation of accurate activation maps, which will contribute to improve the user performance. In this study, the tangible interfaces were designed to help maintain the

spatial and temporal correspondence between the interaction objects and the virtual objects being manipulated. Furthermore, with the help of "Guided Search" model, the benefits of certain interfaces can be explained. For example, forming correct active feature maps is identified as a very important step by the "Guided Search" model during a visual search. Accurate feature maps about the location and orientation of different curved tubes are built after introducing both an arbitrary clipping plane function and a 2D intersection image. Therefore, it is reasonable that the inclusion of a tangible frame and its corresponding 2D intersection image significantly improved user performance compared to only using a simple tangible interface to manipulate 3D viewpoint for counting the number of the curved tubes.

# Chapter 5

# Epilogue

3D interaction with scientific data through visualization technology presents multiple challenges to researchers. The research topics investigated in this thesis might look quite diverse from a reader's point of view. However, they reflect the fact that this 3D interaction question is rather complex. Often when a researcher carries out such research, the research question is highly domain-related. Providing a solution for 3D interaction with scientific data analysis not only requires visualization algorithms as the basis to effectively represent data visually or through other modalities, but also needs to take into account the usability of interaction devices or techniques. The perceptual and cognitive capabilities of a human user play an important role in determining and explaining the success or failure of such interaction processes.

## 5.1 Conclusion and Contribution

### 5.1.1 Conclusion

In this thesis, 3D interaction has been studied from three different perspectives. Three individual research questions were formulated :

1. **What are the usability issues with current graphical user interfaces for TF specification ( in particular with the most frequently used method of**

**trial-and-error)?**

2. **What are the performance differences between available VR systems when analyzing object properties within a volumetric data set, such as size, shape, density and connectivity?**

3. **What are the potential effects of using tangible user interfaces when analyzing object properties within a volumetric data set? In particular, can a clipping plane operation, performed using tangible objects, assist in 3D data analysis tasks?**

The presented results lead the author to emphasize that 3D interaction research is not only about the technology development, the computer system, but also about studying the tasks and the users. More specifically, the following conclusions are drawn from the experimental results.

**Graphical User Interface for TF Specification** More specifically, how a GUI for TF specification in direct volume rendering affects the effectiveness of the visual mapping in a 3D visualization application was the focus in this perspective. A user experiment was designed to study usability issues with the trial-and-error method for interactive TF specification.

- The experimental results indicated that adding a histogram did not significantly improve user performance. Although theoretically the histogram of a data set is a useful property and potentially relevant to the task of TF specification, it was not regarded as valuable additional information by users.
- Limiting the shape and hence also the numbers of DOFs of a TF did not improve the interaction process either. Such functionality is useful for a user in the initial exploration phase, but is too limiting to achieve good performance.
- The inclusion of additional TF information also did not assist a user in speeding up his/her exploration and refinement process. However, several users who had less background knowledge in direct volume rendering or TF specification did appreciate this information, and they preferred it to be present, because it helped them to get started. Since the information about possible TF configurations does

not present any active mechanism to shorten the search time, it is not unexpected that it did not improve user performance in terms of execution time.

**Virtual Reality Systems** How much data analysis tasks can benefit from a virtual environment (VE) was the question of interest in this part of the research. Three VR systems (immersive head-mounted display VR, fish tank VR, fish tank VR with haptics displays) were developed to visualize 3D scalar data. The user performance of these VR systems was compared in four specific data analysis tasks. The results were quite surprising.

- The subjects with immersive HMD VR system obtained the worst performance for most of the tasks.
- The subjects with the fish tank VR system performed better than the subjects in the other two systems in terms of time and accuracy.
- The inclusion of touch in the haptic fish tank VR system was welcomed by the users and seemed to help the users with some tasks, although this was not proved to be statistically significant.

The dense nature of the data sets in the trials and the involvement of analytical reasoning processes were the major reasons that the experimental tasks were considered to be difficult and user performances were low. Some of the tasks were just too difficult and exceeded the limits of subjects' cognitive and perceptual abilities. So, extra system/interface support seems to be necessary for them. For example, most of the subjects regarded judging the densest region and identifying the number of different sizes of spherical objects as very difficult tasks and their performances at times were even below chance.

**Tangible User Interfaces** Since subjects using the fish tank VR system in the previous study performed best in terms of time and accuracy, the question arose as to whether or not user performance could be further improved by offering alternative means for inside-out information (next to the available outside-in view). More specifically, the author considered the option of adding a clipping plane function controlled by tangible devices, combined with visual feedback of the corresponding intersection image. A follow-up user study with the same analysis tasks as in the previous one was carried

out to quantify the user performance. The actual study included six different set-ups: a mouse-based non-immersive VR system, a VR baseline system with a cube for controlling the view transformation, with/without either a virtual clipping plane or a tangible planar frame, and with/without additional feedback through a 2D intersection image. The experimental results indicated several findings.

- Overall, adding a virtual clipping plane that is controlled by a tangible planar frame, together with its 2D intersection (condition CTI) image did significantly improve user performance, in comparison to the baseline system, for most of the tasks, except for counting the number of differently sized spherical objects.
- Adding a virtual clipping plane in a fixed position and providing the corresponding 2D intersection image (condition CFI) significantly improved the accuracy of locating the longest curved tube, in comparison to the baseline system.
- Adding an intersection image had negative effects on judging the sizes of spherical objects, irrespective of whether the clipping plane was in a fixed position or controlled by a tangible planar frame.
- Regarding response time, subjects in the tangible-intersection condition (condition CTI) spent significantly more time on finishing the tasks than subjects in the baseline condition (condition C). This is similar to what was observed when haptic feedback was added in the fish tank set-up in the VR study.

The subjective evaluations of these interface prototypes showed that most of the users perceived the tangible interfaces to be easy to use. The feedback from the subjects indicated that the experimental tasks presented varying difficulties. Finding the densest region was always regarded as the most difficult task. Counting the number of shapes was ranked as the easiest. Although locating the longest curved tube was regarded as a relative easy task upfront, several subjects recognized that it was indeed rather difficult after the experiment.

### 5.1.2 Design Guidelines

An interface designer should aim to utilize the strengths of both the interface and the human users to improve the cognitive process of data analysis. The role of an interface is to help

finding patterns in the data and to assist the organization of the information in ways that are meaningful to the observer. An observer expects that the interface allows him or her to apply his or her skills at hand to refine and organize data in a way that is appropriate for the analysis. This interaction process can be regarded as a technology-mediated dialogue. Although there are many successful 3D interaction devices and techniques available, misunderstandings and difficulties in current 3D interaction still exist:

1. Technology still poses many limitations to the wide use of 3D interfaces and interaction techniques. For instance, current heavy HMDs with relatively low spatial resolutions are obviously inadequate for creating user-friendly immersive VR systems.

2. The differences between universal and inclusive design in 3D interaction research are ignored. Universal design focuses on the usability of the technology and tries to make it accessible for all users. Inclusive design focuses on the capabilities of the intended users and tries to build technologies that are accessible to these specific groups. Users or designers often expect that there will be one or several universal interaction techniques or interfaces that can be applied to all kinds of tasks in 3D interaction. For example, some people expect to create a universal device with multiple DOFs for 3D manipulation that is similar to the mouse in 2D interaction. However, the results of many experiments up till now indicate that applying a single solution to all application domains can be quite disappointing and difficult. Selecting or designing a 3D interface often depends on the task being performed and on the application contexts.

3. Visual perception, for example spatial perception, has played an important part in helping or hindering the data analysis process that takes place in 3D interaction with scientific data. Although there are several perceptual or cognitive models available that can be used as a theoretical basis for interaction researchers in order to understand the role of visual perception and cognition during the data analysis process, the mechanisms behind the visual analytical reasoning are still not understood well enough. This leads to the situation that 3D interaction researchers cannot benefit much from the models that are established in other disciplines.

4. In HCI research itself, there is no convincing interaction theory that describes the

characteristics of 3D interaction and that can offer guidelines for design. Although several evaluation approaches for generic user interface design exist, there is still a lack of systematic methods especially for measuring 3D interfaces and interaction techniques.

Despite of these difficulties and misunderstandings, the results of the three user studies reported in the thesis do provide several useful design guidelines:

**Graphical User Interface for TF Specification** TF specification is actually a parameter optimization process, which requires the user to explore the high-dimensional parameter space in order to search for an appropriate mapping between opacity and data value. The experiment in this thesis indicates that the trial-and-error method is very simple to understand for a user. However, only manipulating the transfer function curve does not simplify the search process. Other available methods have not proven successful to make this specification process completely automatic. Therefore, removing a user from the specification process is difficult and probably also inappropriate. **The core of designing a good user interface for TF specification is that the user interface provides mechanisms (enough and useful information) to speed up the iterative loop and in doing so helps a user to identify an adequate TF without having go through too many iterations.** Otherwise, a user can get frustrated easily. In addition, the TF specification is a highly memory-demanding interaction process. Therefore, any mechanism that can be incorporated in the user interface to reduce the short-term memory workload can be expected to be very beneficial. Additional guidelines that relate to particular information are the following:

- Histograms are potentially useful, provided that the representation is straightforward and the user knows how to derive useful data properties. Therefore, histograms should be included in a TF interface as an optional feature.
- Providing TFs curves with limited DOFs in the trial-and-error method can help a user to get started and can speed up the learning process. But, full control should remain available in order to tune initial results with low-DOF TFs. An interface should hence offer both options.

- Providing suggested TF curves can be useful, particularly for novices. The positive effect can be enhanced if the corresponding rendering results of such TFs are presented to the user as well.

**Virtual Reality Systems** There are many different kinds of VR systems available: fish tank VR, CAVE, etc. VR systems can be differentiated in terms of the degree of immersion, navigation style and the way of observing and manipulating the objects (selection and manipulation). For example, the travel technique used in a HMD based VR system is physical movement; on the contrary, users in a fish tank VR system use manual viewpoint manipulation for travel interaction. Hence, selecting a certain kind of VR system depends on what kind of task a user intends to perform. Specific guidelines that the experimental results provide are the following:

- Immersion and presence are important aspects for improving the subjective experiences in a VR system. However, they are not decisive factors that affect user performance in scientific data analysis.
- Overview is an important feature especially when a data set is dense.
- It is necessary to introduce a functionality that can add landmarks and labels into an immersive HMD based VR system because the visual representation of data alone does not provide enough clues for a user to orient himself/herself during navigation.
- Providing haptic force feedback has added-value for a tracking path task, although the effects are not significant. It can be an optional feature since it affects user preference.
- Immersive HMD-based VR systems have proven their added value for applications such as gaming (or more generally, entertainment). However, they are hard to use for data analysis purposes at this stage because of simulation sickness, and bad ergonomic design of the display hardware involved.

**Tangible User Interfaces** TUIs aim at providing improved interaction experiences. Some useful design guidelines concerning TUIs can be derived from the study reported in this thesis:

- Only working with a 3D input device for manipulating the view transformation of a data set and observing the data set from the outside are not sufficient to answer complicated data properties questions precisely. Using tangible user interfaces for 3D manipulation improves the usability of such interaction. However, it does not help users to better understand the observed data, especially when objects inside a 3D volume are obstructed by the outside objects.
- The physical shape of an interaction device is not a decisive factor when trying to understand object properties within a 3D volume. TUIs can help maintain the spatial and temporal correspondence between the interaction devices and the virtual objects being manipulated. However, this correspondence is not very useful for data analysis tasks.
- The inclusion of a 3D clipping plane function can improve user performance for data analysis tasks, but the effects depend on how it is controlled. A clipping plane that is controlled by a similar shape of interaction device is particularly helpful for a tracking task. When a data set is rather dense, a clipping plane function proves to be useful and should be included.
- The overview capability of a desktop VR environment is a very useful feature.
- Adding a 2D intersection image introduces an inside-out property to a desktop VR system, which is an important feature particularly when the related clipping plane can be rotated in arbitrary directions. It can assist a user in certain tasks, for example locating an object inside a volume. This matches the suggestion from Ware about using 2D layout to support navigation in 3D space [War01].
- Some analysis tasks, such as measuring the size of an object, are difficult to perform without the help of a measurement reference. Adding a measurement tool, such as a virtual ruler, can enable a user to perform absolute measurements more accurately.

## 5.2 Future Perspective

User interfaces change gradually as technology advances. Apple first introduced the GUI into desktop PC, with mouse and keyboard as input devices and a bit-mapped graphic

display as output device, which has now become the standard settings for desktop computer architecture. The so-called WIMP (Windows-Icon-Mouse-Pointing) interface continues to be the state-of-the-art paradigm for personal computers.

### 5.2.1 Future Work Related to This Thesis

Previous studies indicate several useful guidelines on designing 3D interaction techniques and devices for data analysis. There are many possibilities to investigate this topic further, from improving the design of certain interfaces to performing additional experimental studies.

**TF Specification** The trial-and-error method is the basic scheme while manipulating TFs. However, this approach is very primitive. The experimental results indicate that additions, such as histograms, pre-defined TFs and limited-DOF TFs cannot significantly facilitate the specification process. Some even hinder the process. The user responses indicate that a TF user interface with limited DOF is a proper starting point for exploring a data set. Therefore, a user interface that can switch between free style control and limited DOF control may create better user performance than an interface with limited DOF control alone. This needs to be verified. Second, the image-based methods discussed in chapter two seem to be worthy to be investigated experimentally. However, those proposed methods do not provide any interface mechanism for dynamically modifying a TF curve. A user interface that combines the features of image-based methods and of the trial-and-error method seems potentially useful. Furthermore, there is still no experimental evidence regarding the usability of those image-based methods. This should also be investigated in the future. Because image-based methods require creating a lot of rendered images, other related studies regarding how to arrange multiple rendering results can act as good references. For example, Robertson et al. proposed a user interface called Data Mountain that enables users to arrange document thumbnails freely on an inclined plane textured with passive landmarks [RCL+98]. His further study showed that such interfaces can take advantage of spatial memory. However, another study indicates that there was no significant difference between 2D and 3D interfaces in terms of arranging multiple thumbnail images efficiently [Coc04]. Therefore, investigating image-based methods should consider investigating the means of better arranging the TF rendering results.

Data-based methods also need to be investigated further. The semi-automatic method proposed by Kindlmann uses data properties to detect object boundaries and applies this knowledge in the rendering [KD98]. However, different data sets from different materials or domains may have very different properties. It is not guaranteed that the proposed algorithm is sufficiently robust to be useful in diverse settings. Therefore, designing a user interface that combines semi-automatic method with extra user-intervention mechanism will probably be a better solution. This of course needs to be verified in future experimental studies.

**Virtual Reality Systems** Immersive HMD-based VR systems usually do not have an overview feature. The experiment in this thesis showed that when users operate in such system to observe a 3D data set, they often moved away from the center of the volume so as to view the data set from the outside. Therefore, adding a functionality that can shrink or restore the data set, for instance by clicking a button, will be helpful. Second, improving an immersive HMD based VR system can be achieved by adding more artificial labels within the VE so as to increase users' spatial awareness, because usually a data set itself does not have any spatial landmarks. The user study only investigated three major kinds of VR systems. CAVE is another kind of immersive VR set-up, and might be worthy to be incorporated in a similar study. To compare the user performance between an immersive HMD based VR system and a CAVE system using the same data sets and tasks can help us identify distinct aspects of both systems. The study of haptic feedback did show positive, but not significant effects on counting the number of curved tubes and tracking spatial topology of the tube objects. One possible reason is that the fish tank haptic system does not have a trace map that can indicate the real-time location of the haptic tool during exploration, which makes the inside-out property less effective. Therefore, future work can be to add an additional path map that shows the location of the haptic tool within the observed volume dynamically. The study in this thesis asked subjects to perform tasks as fast as possible. Most subjects did not have extensive experience with the haptic user interface and this might have contributed to that fact that some subjects wanted to finish the tasks qickly without fully profiting from the haptic feedback function. Therefore, future work should probably include investigating the long-term effects of haptic feedback, amongst others without posing time restrictions.

**Tangible User Interfaces** Current design of TUIs is still primitive in terms of shape and function. The cube and the clipping frame provide a very intuitive way to manipulate and cut through a 3D data set. However, there is no clutch mechanism to freeze the movement of both tangible interfaces when a user finds a region of interest. A user must keep his/her pose steady for a continuous period of time and analyze the data set at the same time, during which his/her attention is very easily distracted. It leads to an intermittent interaction process. Therefore, adding a clutch mechanism that can fix and capture the rendering results whenever a user wants will help the user concentrate on the task. It will also enable the interaction flow to be smoother.

Furthermore, the clipping plane function in this thesis exposes slice details within a 3D spatial context, but removes the part of data between the clipping plane and the observer. However, it may be useful if the data elements and the local neighborhood remain visible. So an alternative design is to open the volume up along a clipping plane, using a book or cutting metaphor (for example [Cow00], [KY97]) so that the context information is pushed aside but not removed or made invisible. For example, in the "Corner Cube" [Reh98], three orthogonal slices of a medical data set are projected to the sides of a cube, allowing them to be viewed in their relative orientations. Volumes of interest (VOI) are drawn in the center of the cube, with outlines projected to the cube walls; thus, the walls provide anatomical context for the VOIs. The drawback of this design is that the context information in the corner cube is limited to three slices in fixed orientations and few details of the VOIs are displayed. Therefore, further work could include the development of clipping plane functions that can open up a volume along several clipping planes with user-defined arbitrary orientation. TUIs may play a role in this.

Moreover, the study in the second chapter indicates that in direct volume rendering, TF can reduce occlusion by making voxels semitransparent so that voxels behind them become visible. Therefore, possible future work can develop a method to highlight spatial areas of interest by manipulating the TF. This method can make certain area more visible, but does not increase the screen space allotted to areas of interest. It may help users to perform some data analysis tasks, which are currently difficult because of occlusion.

### 5.2.2 Future Work for Generic 3D Interaction

3D interaction through AR, VR or TUIs continues to be a challenging topic for interface researchers. Currently, 3D interaction itself is still an experience, instead of a routine. Practically, there is not enough evidence to support that 3D interfaces can increase the speed of interaction, or enable a better understanding of the observed data for a user.

In order to make 3D interfaces and interaction indeed work in real world applications, future investigation should, in the author's view, concentrate on the following topics:

- Advancing technology: technology advances can solve part of the existing problems in 3D interaction. A new generation of 3D LCD displays already can be found in the market in terms of hardware development. They avoid asking a user to wear uncomfortable 3D classes during the interaction. The newly released Wii by Nintendo is another true example of an engaging and responsive interaction in 3D space through a handheld pointing device that detects a user's motion in real time (Figure 5.1). They demonstrate in which way technology can be improved and applied to 3D applications.

- Improving the ergonomics of 3D interfaces: one factor that can lead to frustration with 3D interfaces is the poor ergonomic design. For example, it takes time until a user gets used to heavy goggles, sometimes too long for acceptance and practical usage.

- Further understanding of the perceptual and cognitive issues behind 3D interfaces: Navigating through a 3D space can be natural and attractive in the beginning, but soon after such space and way of interaction may become an obstacle for a user. In some worst cases, 3D interfaces make the user feel frustrated and lose interest. The interaction process slows down. For instance, simulation sickness often takes place with unclear reasons in VR interfaces. Although inappropriate ergonomic design is one of the reasons, further study on perceptual and cognitive issues can uncover additional facts.

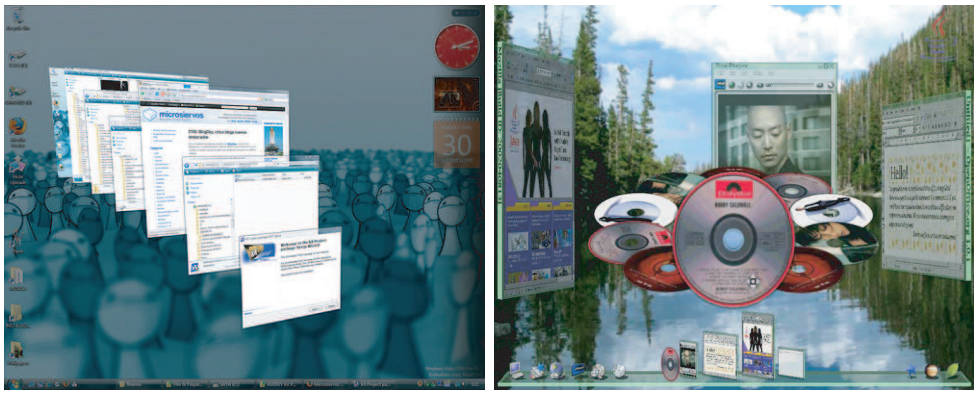
- Finding the "killer" application for 3D interfaces: Despite of the existing difficulties, entertainment industry, auto manufacturers and some educational interfaces have benefited from the fun and engaging nature of 3D interaction, as evidenced by the success


(a) (b)

Figure 5.1: The diagram of Wii: (a) console platform; (b) tangible remote controller. (Photograph reprinted from [Inc]

of "shoot-em-up" video games. There are other successful examples of 3D user interfaces, which can be found in architecture and medical applications. The 3D interface metaphors appeared in the recently released windows operating system Vista from Microsoft and the Project "Looking Glass" by Sun Microsystems are attempts to explore the possibilities of offering a richer user experience through innovative 3D desktop user interfaces (Figure 5.2). These 3D interfaces are not only about looks, they are about creating an engaging user experience, which makes interaction and communication easier. For example, in the prototype of "Looking Glass", desktop applications are not represented by 2D icons and buttons that are stacked upon one another. They are represented as 3D widgets in a 3D environment, and manipulated as 3D objects (Figure 5.2b). This kind of design moves beyond the old configurations and brings revolutions to the use of the desktop, which exposes the potential of 3D interfaces for the near future.

- Establishing proper models: There are already several models available in HCI that describe a user's behavior during interaction or the relationship between a user and physical devices, for example, Fitts' model of the information processing capability of the human motor system and Guiard's model of bimanual control. However, no



(a) (b)

Figure 5.2: 3D user interfaces on Desktop PC: (a) the operating system Vista from *Microsoft*; (b) 3D user interface - the "Looking Glass" Project from *Sun* (Photograph reprinted from [MI])

proper model addresses the issues involved in 3D interaction. Establishing proper models and performance metrics can contribute to a better understanding and design of 3D interaction techniques/devices.

- Proposing better evaluation methods and collecting enough experimental evidence: The novelty and the limitless possibilities of 3D user interfaces and interaction research have resulted in a practice where researchers mostly focus on developing new devices, interaction techniques, and user interface metaphors. Researchers put most of their efforts in exploring the design space without paying much attention to assess how valid the new designs are. Another cause is that most researchers have a technical background so that they will be naturally inclined to emphasize technology, rather than usability. Evaluations also have not been used to actively influence the design process. Therefore, there exists a wealth of interaction techniques and interfaces and some general thoughts about their advantages and disadvantages, but very little experimental evidence about their concrete performance, usability, and usefulness. At the same time, although there are several evaluation methods available in HCI, customized evaluation methods for 3D interaction are still in their infancy. Proposing customized evaluation methods and carrying out more systematic evaluations should

be important future work so that enough experimental evidence can be collected to guide the design of 3D interaction.

In short, with the future development of technology and better understanding of the principles behind 3D interaction, 3D interface will become more useful for data analysis by scientists and professionals.

# Bibliography


[ABW93] Kevin W. Arthur, Kellogg S. Booth, and C. Ware. Evaluating 3D task performance for fish tank virtual worlds. *ACM Trans. Inf. Syst*, 11(3):239–265, 1993.

[Ahu93] A.J. Ahumada. Computational image quality metrics: A review. *Society for Information Display International Symposium, Digest of Technical Papers*, 24:305–308, 1993.

[Bad86] A.D. Baddeley. *Working Memory.* Clarendon, Oxford, 1986.

[Bad90] A.D. Baddeley. *Human Memory.* Allyn and Bacon, Boston, 1990.

[BAR+02] Doug A. Bowman, Datey Ameya, Young Sam Ryu, Umer Farooq, and Omar Vasnaik. Empirical comparison of human behavior and performance with different display devices for virtual environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pages 2134–2138, 2002.

[BB92] J. Beddow and C. Beshers. Designing a visualization interface for multi-dimensional, multivariate data. Tutorial Notes at Visualization, 1992.

[BCE+92] K.W. Brodlie, L.A. Carpenter, R.A. Earnshaw, J.R. Gallop, and R.J. Hubbard. *Scientific Visualization, Techniques and Applications.* Springer-Verlag, 1992.

[BCWea06] Doug A. Bowman, Jian Chen, Chadwick A. Wingrave, and et al. New directions in 3D user interfaces. *The International Journal of Virtual Reality*, 5(2):3–14, 2006.

[Bie87] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94(2):115–147, 1987.

[BIJH00] J. Brederson, M. Ikits, C. Johnson, and C. Hansen. The visual haptic workbench. In *Proceedings of PHANToM Users Group Workshop*, pages 46–49. Aspen, CO, 2000.

[BKLP04] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley, New Jersey, 2004.

[BKPP92] R. Bakalash, A. Kaufman, R. Pacheco, and H. Pfister. An extended volume visualization system for arbitrary parallel projection, 1992.

[BOYBK90] F. P. Brooks, M. Ouh-Young, J.J. Batter, and P.J. Kilpatrick. Project GROPE - haptic displays for scientific visualization. *Computer Graphics*, 24(4):177–185, 1990.

[BP02] C.P. Botha and F.H. Post. New technique for transfer function specification in direct volume rendering using real-time visual feedback. In S. K. Mun, editor, *Proceedings of the SPIE International Symposium on Medical Imaging- Visualization, Image-Guided Procedures, and Display*, pages 227–234. SPIE, SPIE, 2002.

[BPW03] N.H. Bakker, P.O Passenier, and P.J. Werkhoven. The effects of head-slaved navigation and the use of teleports on spatial orientation in virtual environments. *Human Factors*, 45(1):160–169, 2003.

[Bro59] J. Brown. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 10:12–21, 1959.

[Bro99] F.P. Brooks. What's real about virtual reality? *IEEE Computer Graphics and Applications.*, 19:16–27, 1999.

[BS01] C. Basdogan and M. Srinivasan. Haptic rendering in virtual environments, 2001.

[Cas91] Stephen M. Casner. A task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, 10(2):111–151, 1991.

[CM83] S. Card and T. Morgan. *The Psychology of Human Computer Interaction.* Lawrence Erlbaum Associates, Hillsdale, 1983.

[CMS88] M. Chen, S.J. Mountford, and A. Sellen. A study in interactive 3D rotation using 2D control devices. *Computer Graphics*, 22(4):121–129, August 1988.

[CNSD93] C. Cruz-Neira, D.J. Sandin, and T.A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the CAVE. *ACM Computer Graphics*, 27(2):135–142, July 1993.

[COAC97] Wickens C, Olmos O, Chudy A, and Davenport C. Aviation display support for situation awareness. Technical report: ARL-97-10/LOGICON-97-2. Aviation Research Lab, Institute of Aviation, University of Illinois at Urbana-Champaign. http://www.aviation.uiuc.edu/new/html/ARL/TechPdf/97-10.pdf, 1997.

[Coc04] Andy Cockburn. Revisiting 2D vs 3D implications on spatial memory. In *AUIC '04: Proceedings of the fifth conference on Australasian user interface*, pages 25–31, Darlinghurst, Australia, Australia, 2004. Australian Computer Society, Inc.

[Cow00] David J. Cowperthwaite. *Occlusion resolution operators for three-dimensional detail-in-context.* PhD thesis, 2000. Adviser-F. David Fracchia.

[DCH88] R. A. Drebin, L. Carpenter, and P. Hanrahan. Volume rendering. In *(SIGGRAPH '88 Proceedings*, pages 65–74, USA, 1988. ACM Press.

[DGWlHCM+03] E. De Guzman, F. Wai-ling Ho-Ching, T. Matthews, T. Rattenbury, M. Back, and S. Harrison. Eewww!: Tangible instruments for navigating into the human body. In *Extended Abstracts of CHI 2003*, pages 806–807. ACM, 2003.

[DLJ+03] Cagatay Demiralp, David H. Laidlaw, Cullen Jackson, Daniel Keefe, and Song Zhang. Subjective usefulness of CAVE and Fish Tank VR display systems for a scientific visualization application. In *Proceedings of IEEE Visualization*, Seattle, WA, USA, 2003. IEEE Press.

[EB02] M. Ernst and M. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, 2002.

[EF95] A.M. Eskicioglu and P.S. Fisher. Image quality measures and their performance. *IEEE Computer Graphics and Applications*, 43(12):2959–2965, 1995.

[Fer] Antonio Ramires Fernandes. View frustum culling tutorial. http://www.lighthouse3d.com/opengl/viewfrustum/.

[FHSH06] Bernd Froehlich, Jan Hochstrate, Verena Skuk, and Anke Huckauf. The GlobeFish and the GlobeMouse: two new six degree of freedom input devices for graphics applications. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 191–199, New York, NY, USA, 2006. ACM Press.

[FLP03] Joseph L. Fleiss, Bruce Levin, and Myunghee Cho Paik. *Statistical Methods for Rates and Proportions.* Number 0-471-52629-0. John Willey & Sons, 2003.

[FP00] B. Froehlich and J. Plate. The cubic mouse: A new device for three-dimensional input. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 526–531. ACM Press, 2000.

[GH98] Maarten W. Gribnau and James M. Y Hennesse. Comparing single- and two-handed 3D input for a 3D object assembly task. In *CHI '98: CHI*

*98 conference summary on Human factors in computing systems*, pages 233–234, New York, NY, USA, 1998. ACM Press.

[GHD99] G. Gediga, Kai-christoph Hamborg, and I. Duntsch. The isometrics usability inventory: An operationalisation of ISO 9241-10. *Behaviour and Information Technology*, 18:151–164, 1999.

[Gib] W. K. Gibson. *Neuromancer*. Ace Books, New York.

[Gri99] M.W. Gribnau. *Two-handed Interaction in Computer Supported 3D Conceptual Modeling*. PhD thesis, Delft University of Technology, Delft, Netherlands, 1999.

[Gui87] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior*, 19, 1987.

[Hau89] A. G. Hauptmann. Speech and gestures for graphic image manipulation. In *Proc. ACM CHI'89 Conference on Human Factors in Computing Systems*, pages 241–245, 1989.

[HD91] Robert Haimes and D. Darmofal. Visualization in computational fluid dynamics: A case study. In *IEEE Visualization*, pages 392–397, 1991.

[HHKP96] T. He, L. Hong, A. Kaufman, and H. Pfister. Generation of transfer functions with stochastic search techniques. In *Proceedings of IEEE Visualization 1996 (VIS '96)*, pages 227–234. IEEE, IEEE CS, 1996.

[Hin96] K. Hinckley. *Haptic issues for virtual manipulation*. PhD thesis, University of Virginia, Virginia, 1996.

[HJ04] Charles D. Hansen and Chris R. Johnson. *The Visualization Handbook*. Elsevier Butterworth Heinemann, Massachusetts, 2004.

[HPG94] K. Hinckley, R. Pausch, and Kassell N. F. Goble, J.C. and. A three-dimensional user interface for neurosurgical visualization. In *the SPIE Conf. on Medical Imaging*, pages 126–136. SPIE, 1994.

[HTP+97] K. Hinckley, J. Tullio, R. Pausch, D. Proffitt, and N. Kassell. Usability analysis of 3D rotation techniques. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, pages 1–10. ACM, 1997.

[IK95] T. Itoh and K. Koyamada. Automatic isosurface propagation using an extrema graph and sorted boundart cell lists. *IEEE TVCG*, 1(4):319–327, 1995.

[Inc] Nintendo Inc. Wii. http://wii.com/.

[IU97] H. Ishii and B. Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of CHI 1997*, pages 234–241. ACM, 1997.

[JKM00] T.J. Jankun-Kelly and Kwan-Liu. Ma. A spreadsheet interface for visualization exploration. In *Proceedings IEEE Visualization 2000*, pages 69–76. IEEE, 2000.

[Joh04] C.R. Johnson. Top scientific visualization research problems. *IEEE Computer Graphics and Applications*, 20(6):2–6, July/August 2004.

[KB88] Arie Kaufman and Reuven Bakalash. Memory and processing architecture for 3D voxel-based imagery. *IEEE Comput. Graph. Appl.*, 8(6):10–23, 1988.

[KBF+95] W. Kreuger, C. Bohn, B. Froehlich, H. Schueth, W. Strauss, and G. Wesche. The Responsive Workbench: A virtual work environment. *IEEE Computer.*, 28(7):42–48, July 1995.

[KBS94] Paul Kabbash, William Buxton, and Abigail Sellen. Two-handed input in a compound task. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 417–423, New York, NY, USA, 1994. ACM Press.

[KD98] G. Kindlmann and J. W. Durkin. Semi-automatic generation of transfer functions for direct volume rendering. In *IEEE Symposium On Volume Visualization*, pages 79–86. IEEE, IEEE CS, 1998.

[Kea90] R. L. Klatzky and et al. *Procedures for Haptic Object Exploration Vs. Manipulation*, pages 110–127. Ablex, New Jersey, 1990.

[KG01] A. Konig and E. Gruller. Mastering transfer function specification by using volumepro technology. In Tosiyasu L. Kunii, editor, *Spring Conference on Computer Graphics 2001*, pages 279–286. IEEE, 2001.

[KHI+03] Robert Kosara, Christopher G. Healey, Victoria Interrante, David H. Laidlaw, and Colin Ware. User studies: Why, how, and when? *IEEE Computer Graphics and Applications*, 23(4):20–25, 2003.

[KIK01] Y. Kitamura, Y. Itoh, and F. Kishino. Real-time 3D interaction with activecube. In *Extended Abstracts of CHI 2001*, pages 355–356. ACM, March 31 - April 5 2001.

[Kre00] Oliver Kreylos. Interactive volume rendering using 3D texture-mapping hardware. lecture notes, 2000.

[KY97] Yair Kurzion and Roni Yagel. Interactive space deformation with hardware-assisted rendering. *IEEE Comput. Graph. Appl.*, 17(5):66–77, 1997.

[LC87] W. E. Lorensen and H.E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *Proceedings of SIGGRAPH 87*, pages 163–169. ACM, ACM, 1987.

[LCN98] B. Lichtenbelt, R. Crane, and S. Naqvi. *Introduction to Volume Rendering*. Prentice-Hall, New Jersey, 1998.

[Lev88] M. Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(5):29–37, 1988.

[Lew95] J. R. Lewis. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 21(3):57–78, 1995.

[LM03] Liere, R. van and J.D. Mulder. Optical tracking using projective invariant marker pattern properties. In *Proceedings of the IEEE Virtual Reality Conference 2003*, pages 191–198, 2003.

[LR03] R. van Liere and A. van Rhijn. Search space reduction in optical tracking. In J. Deisinger and A. Kunz, editors, *Proceedings of the Immersive Projection Technology and Virtual Environments Workshop 2003 (IPT/EGVE 2003)*, pages 207–214, 2003.

[LZB98] Andrea Leganchuk, Shumin Zhai, and William Buxton. Manual and cognitive benefits of two-handed input: An experimental study. *ACM Trans. Comput.-Human. Interact.*, 5(4):326–359, 1998.

[Ma99] Kwan-Liu. Ma. Image graphs - a novel approach to visual data exploration. In *Proceedings IEEE Visualization 1999*, pages 81–88. IEEE, 1999.

[MABea97] J. Marks, B. Andalman, P.A. Beardsley, and et al. Design galleries: A general approach to setting parameters for computer graphics and animation. In *ACM Computer Graphics (SIGGRAPH '97 Proceedings)*, Computer Graphics Proceedings, Annual Conference Series, pages 389–400. ACM, ACM Press / ACM SIGGRAPH, 1997.

[MB97] Tom McReynolds and David Blythe. Programming with OpenGL: Advanced rendering. SIGGRAPH '97 Course, 1997.

[Mel63] A.W. Melton. Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and verbal Behavior*, 2:1–21, 1963.

[MI] Sun Microsystems Inc. Project Looking Glass. $http : //www.sun.com/software/looking_glass/$.

[Mil56] G.A. Miller. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63:81–97, 1956.

[MIWB02] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks. Physiological measures of presence in stressful virtual environments. In *Proceedings of ACM SIGGRAPH 2002*, volume 21, pages 645–652, 2002.

[ML02] J.D. Mulder and R. van Liere. The personal space station: Bringing interaction within reach. In *VRIC 2002 Conference Proceedings*, pages 73–81, 2002.

[MLK07] J. B. Martens, R. van Liere, and A. Kok. Widget manipulation revisited: a case study in modeling interactions between experimental conditions. In *Proceedings of IPT-EGVE 2007*, 2007.

[MM95] D. P. Mapes and J. M. Moshell. A two-handed interface for object manipulation in virtual environments. *Presence*, 4(4):403–416, 1995.

[Mor86] N. Moray. *Handbook of Perception and Human Performance*. Wiley, New York, 1986.

[MP00] M. Masliah and Milgram P. Measuring the allocation of control in a 6 Degree-Of-Freedom docking experiment. In *Proceedings ACM CHI 2000*, pages 25–32, 2000.

[MQA+04] J.B. Martens, W. Qi, A. Aliakseyeu, A.J.F. Kok, and R. van Liere. Experiencing 3D interactions in virtual reality and augmented reality. In *Proceedings of the 2nd European Union Symposium on Ambient Intelligence (EUSAI 2004)*, pages 25–28, 2004.

[Nei67] U. Neissen. *Cognitive Psychology*. Prentice-Hall, New York, 1967.

[Nor93] Donald A. Norman. *Things that Make Us Smart*. Addison-Wesley Publishing Co., Reading, MA, 1993.

[NS86] K. Nakayama and G.H. Silverman. Serial and parallel processing of visual feature conjunctions. *Nature*, 320(6059):264–265, Mar 1986.

[OPL+97] Randy Osborne, Hanspeter Pfister, Hugh Lauer, Neil McKenzie, Sarah Gibson, and Wally Hiatt. EM-Cube: An architecture for low-cost real-time volume rendering. In Steven Molnar and Bengt-Olaf Schneider, editors, *1997 SIGGRAPH / Eurographics Workshop on Graphics Hardware*, pages 131–138, New York City, NY, 1997. ACM Press.

[PBWI96] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 79–80, New York, NY, USA, 1996. ACM Press.

[PLB+01] H. Pfister, H. Lorensen, C. Bajaj, G. Kindlmann, W. Schroeder, and et al. The transfer function bake-off. *IEEE Computer Graphics and Applications*, 21(3):16–22, 2001.

[PPW97] Randy Pausch, Dennis Proffitt, and George Williams. Quantifying immersion in virtual reality. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 13–18, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[PT90] B. A. Payne and A. W. Toga. Surface mapping brain function on 3D models. *IEEE Computer Graphics and Applications*, 10(5):3341, September 1990.

[PVG95] P. Peruch, J.-L. Vercher, and G. M. Gauthier. Acquisition of spatial knowledge through visual exploration of simulated environments. *Ecological Psychology*, 7(1):1–20, 1995.

[RCL+98] George Robertson, Mary Czerwinski, Kevin Larson, Daniel C. Robbins, David Thiel, and Maarten van Dantzich. Data mountain: using spatial memory for document management. In *UIST '98: Proceedings of the 11th annual ACM symposium on User interface software and technology*, pages 153–162, New York, NY, USA, 1998. ACM Press.

[Reh98] K. Rehm. A symbolic environment for visualizing activated foci in functional neuroimaging datasets. *Medical Image Analysis*, 2(3):215–226, 1998.

[Rob90] Philip K. Robertson. A methodology for scientific data visualization: Choosing representations based on a natrual scene paradigm. In *Proceedings of Visualization'90*, pages 114–123. IEEE Computer Society, 1990.

[RS90] J.W. Regian and W.L. Shebiske. Virtual reality: A instructional medium for visual-spatial tasks. *Journal of Communications*, 42(2):136–149, June 1990.

[RS01] Christof Rezk-Salama. *Volume Rendering Techniques for General Purpose Graphics Hardware*. PhD thesis, Universit at Erlangen-Nurnberg, Germany, 2001.

[RSKK06] Christof Rezk-Salama, Maik Keller, and Peter Kohlmann. High-level user interfaces for transfer function design with semantics. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1021–1028, 2006.

[RW01] S. Z. Kohn Razzaque and M. Whitton. Redirected walking. In *Proceedings of Eurographics 2001*, pages 289–294, Manchester, 2001. ACM, ACM Press / ACM SIGGRAPH.

[SAH00] William J. Schroeder, Lisa S. Avila, and William Hoffman. Visualizing with VTK: A tutorial. *IEEE Comput. Graph. Appl.*, 20(5):20–27, 2000.

[Sar06] N.B. Sarter. Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics*, 36(5):439–445, 2006.

[SB97] Mandayam A. Srinivasan and Cagatay Basdogan. Haptics in virtual environments: Taxonomy, research status, and challenges. *Computers & Graphics*, 21(4):393–404, 1997.

[SCB04] Kenneth Salisbury, François Conti, and Federico Barbagli. Haptic rendering: Introductory concepts. *IEEE Computer Graphics and Applications*, 24(2):24–32, 2004.

[SFKL05] Jrgen Schulze, Andrew Forsberg, Robert Zeleznik Kleppe, and David H. Laidlaw. Characterizing the effect of level of immersion on a 3D marking task. In *Proceedings of HCI International*, pages 447–452. HCI International, 2005.

[SK04] Hartmut Seichter and Thomas Kvan. Tangible interfaces in design computing. In *Education and Research in Computer Aided Architectural Design (eCAADe 2004)*, pages 159–166, Copenhagen, Denmark, 2004.

[Sne92] B. Sneiderman. *Designing the User Interface: Strategies for Effective Human-computer Interaction.* Addison-Wesley Publishing Company, Reading, Massachusetts, 1992.

[SS97] J. Kenneth Salisbury and Mandayam A. Srinivasan. Phantom-based haptic interaction with virtual objects. *IEEE Comput. Graph. Appl.*, 17(5):6–10, 1997.

[Sut68] I. Sutherland. A head-mounted three-dimensional display. In *Proceeding of the Fall Joint Computer Conference*, pages 757– 764. AFIPS, 1968.

[SWWL01] J.P. Schulze, U. Wussner, S.P. Walz, and U. Lang. Volume rendering in a virtual environment. In *Proceedings of 5th IPTW and Eurographics Virtual Environments*, pages 187–198. Springer Verlag, 2001.

[TC05] J.J. Thomas and K.A. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics.* IEEE CS Press, 2005.

[TCO+97] R. M. Taylor, J. Chen, S. Okimoto, N. Llopis-Artime, V. L. Chi, F. P. Brooks, and et al;. Pearls found on the way to the ideal interface for scanned-probe microscopes. In *Proceedings of Proceedings of IEEE Visualization*, pages 467–470. IEEE Computer Society Press., 1997.

[TL03] F.Y. Tzeng and E. Lum. A novel interface for higher-dimensional classification of volume data. In *Proceedings of the IEEE Visualization 2003*, pages 20–24. IEEE, 2003.

[UI01] B. Ullmer and H. Ishii. *Emerging Frameworks for Tangible User Interfaces.* Addison-WesleyNew Jersey for Human-Computer Interaction in the New Millenium, John M. Carroll, ed., 2001.

[vDLS02] A. van Dam, D. H. Laidlaw, and R. M. Simpson. Experiments in immersive virtual reality for scientific visualization. *Computer & Graphics*, 26(4):535–555, August 2002.

[WAB93] C. Ware, K. Arthur, and K.S. Booth. Fish tank virtual reality. In *Proceedings of CHI 93*, pages 37–42. ACM and IFIP, 1993.

[War01] C. Ware. Designing with a 2 $\frac{1}{2}$D attitude. *Information Design Journal*, 10(3):255–262, 2001.

[WBV+01] Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, and D'nardo Colucci. High-performance wide-area optical tracking -the hiball tracking system. *Presence: Teleoperators and Virtual Environments*, 10(1), 2001.

[Wes90] L. Westover. Footprint evaluation for volume rendering. In *Proc. SIGGRAPH*, pages 2134–2138, 1990.

[WF96] Colin Ware and Glenn Franck. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Trans. Graph.*, 15(2):121–140, 1996.

[WG98] P. Werkhoven and J. Groen. Manipulation performance in interactive virtual environments. *Human Factors*, 40(3):432–442, 1998.

[WH99] D.C. Wickens and G.J. Hollands. *Enginerring Psychology and Human Performance*. Prentice-Hall, New Jersey, 1999.

[WL90] Stephen Wehrend and Clayton Lewis. A problem-oriented classification of visualization techniques. In *Proceedings of Visualization 90*, pages 139–143. IEEE, IEEE, 1990.

[WLPO95] C. Wickens, C. Liang, T. Prevett, and O. Olmos. Egocentric and exocentric dipslays for terminal area navigation. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*, pages 16–20, 1995.

[Wol94] J.M. Wolfe. Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2):202–238, 1994.

[WP82] R.N. Wilton and B. Pidcock. Knowledge of spatial relations: Varying the precision with which locations must be specified. *Quarterly Journal of Experimental Psychology*, 34:515–528, 1982.

[WVG92] J. Wilhelms and A. Van Gelder. Octrees for faster isosurface generation. *CM Transactions on Graphics*, 11(3):201–227, 1992.

[Zha98] Shumin Zhai. User performance in relation to 3D input device design. *Computer Graphics*, 32(4):50–54, 1998.

[ZM93] Shumin Zhai and Paul Milgram. Human performance evaluation of manipulation schemes in virtual environments. In *VR*, pages 155–161, 1993.

# Summary

Users that want to engage in 3D interaction with scientific data through visualization technology experience a range of different problems. The starting point in this thesis is that scientific visualization has technologically progressed to a stage where practically relevant analysis work can in principle be undertaken with scientific data. However, there still exist important barriers to a widespread use of visualization technology, primarily in the interaction techniques and interface design. The author hence does not focus on improving current visualization algorithms or developing new ones, but instead investigates specific usability issues that play a role in the interaction process with 3D data. This thesis presents more detailed insights into existing problems and provides experimental evidence from several user studies that can help to resolve some of the issues. The results also contribute to a better understanding of remaining challenges in the design of 3D user interfaces.

The author identifies several misunderstandings and difficulties within current 3D interaction research:

1. Often, 3D interaction designers expect and take for granted that there will be generic interaction techniques or interfaces that can be applied to a wide range of tasks in 3D interaction. Therefore, researchers try to apply a very limited number of solutions to many different application domains.

2. Technology still poses many limitations to the implementation of 3D interface and interaction techniques. Especially the real-time requirement of interaction is often difficult to meet and necessitates compromises.

3. Although there are several perceptual or cognitive models available that can be used as

a theoretical basis for interaction researchers when trying to understand the role of visual perception and cognition in the data analysis process, the mechanisms influencing the visual perception and the analytical reasoning are still insufficiently understood to derive clear guidelines.

4. The field of Human-Computer Interaction (HCI) does not offer a widely accepted interaction model or performance metrics that can be used to formulate guidelines for or analyze the characteristics of 3D interaction interfaces. A systematic approach for evaluating 3D interfaces and interaction techniques is clearly lacking.

The research methodology that the author adheres to is to create/design and evaluate interaction techniques and interfaces for performing diverse but generic analysis tasks with volumetric data. These tasks are inspired by demands from specific disciplines. Three different perspectives have inspired the research and have helped to shape the concrete empirical studies. More specifically, these studies have addressed: 1) the graphical user interface (GUI) for transfer function (TF) specification, 2) the role of desktop versus immersive Virtual Reality (VR) as the environment for creating the human-system interface, and 3) the role of tangible devices in 3D interaction:

**Graphical User Interface for TF Specification** TF is a critical component in a direct volume rendering system. The TF is a critical component in direct volume rendering. It specifies the relation between physically observed data (such as densities measured in CT or MRI imaging), possibly also including their first- or second-order derivative values, and optical characteristics relevant for rendering, such as color and opacity. The most widely used method today is trial-and-error, in which the TF specification is accomplished by either manually editing a graphical curve that represents the TF, or by adjusting numerical parameters. By visually inspecting the resulting image, the user can assess the impact of his actions. A user study was designed to gather more quantitative data on how different aspects of a GUI for TF specification, using a trial-and-error method, affect the effectiveness of a prescribed visualization task. More specifically, the study established the usefulness of histogram information, possible/suggested TFs and limited DOFs for interactive TF specification. The obtained results can be summarized as follows:

- There is no significant evidence that histograms can help to improve the user performance in terms of time or accuracy.
- Additional information about possible/suggested TFs can prove useful to novel users, but also does not contribute to increased performance in the long run.
- Interfaces that restrict the number of DOF of a TF also do not improve the chance to find a TF that is suitable for the task, and are moreover not better appreciated by users.

Working memory theory can possibly help to explain why subjects failed to finish the specification tasks effectively and efficiently with the trial-and-error interfaces offered to them. Indeed, these interfaces do not provide any mechanism to relieve the workload of working memory. This suggests that a designer of TF user interfaces should take measures to relieve the workload of working memory, for instance by allowing users to view and return to earlier stages in the interaction process.

**Virtual Reality Systems** A VR interface is a combination of 3D input and output devices that aims at creating an experience for humans that helps them to better exploit well-established skills, developed by interacting with the physical world. VR is hence widely proposed as an alternative for helping to improve the interactivity of 3D computer systems. How much scientific visualization can benefit from a virtual environment was the core question in this second empirical study. The author was particularly interest in how different VR systems influence performance for a range of scientific visualization tasks (such as the identification or estimation of the size, shape, density, and connectivity of objects present in a volume). The tasks were derived from data analysis questions that are raised by domain specialists who are studying Cystic Fybrosis (CF). The conducted study compared the user performance of three different stereoscopic VR systems: (1) a head-mounted display (HMD); (2) a fish tank VR (fish tank); and (3) a fish tank VR augmented with a haptic device (haptic). HMD subjects were placed "inside" the volumetric data set and walked around in it to explore its structure. Fish tank and haptic subjects saw the entire volume on-screen and rotated it to observe it from different viewpoints. Response time and accuracy were used to measure performance.

The results show that subjects in the fish tank and haptic groups were significantly

more accurate at judging the shape, density, and connectivity of the objects and completed the tasks significantly faster than the HMD group. Although the fish tank group itself was significantly faster than the haptic group, there were no statistical differences in accuracy between these two groups. Subjects classified the HMD system as an "inside-out" display (looking outwards from inside the volume), and the fish tank and haptic systems as "outside-in" displays (looking inwards from outside the volume). Including haptics added a kind of inside-out capability to the fish tank system through the use of touch. An outside-in system is recommended, since it offers both overview and context, two visual properties that are important for the volume analysis tasks that the author studied. In addition, based on the opinion of the subjects in the haptic group (80% positive) that haptic feedback aided comprehension, supplementing the outside-in visual display with inside-out haptics is recommended whenever possible.

**Tangible User Interfaces** Since subjects using the fish tank VR system in the previous study performed best in terms of time and accuracy, the question arose as to whether or not the user performance could be further improved by offering alternative means for inside-out information. More specifically, the author considered the option of adding a clipping plane function controlled by tangible devices combined with visual feedback of the corresponding intersection image. An extensive user study with the same analysis tasks as in the previous one was carried out to quantify the user performance. The actual study included six different set-ups. A non-immersive VR set-up was equipped with a mouse, a virtual clipping plane, or tangible clipping frame, and additional feedback through a 2D intersection image was provided in two of the six conditions.

The experimental results indicate several findings. Overall, adding a virtual clipping plane that is controlled by a tangible frame, and providing a separate 2D intersection image, did significantly improve user performance, compared to the baseline system, for most tasks, except for counting the number of differently sized spherical objects. An intersection image had negative effects on the size task, irrespective of whether or not the clipping plane was in a fixed position or controlled by a tangible frame. Adding a clipping plane in a fixed position together with the correspondent intersection image significantly improved the accuracy for locating the tube object, in comparison to the baseline system. Regarding the response time, users who worked with both the tangible frame and the intersection image spent more time on finishing the tasks than

with the baseline system.

The subjective evaluations have shown that the introduction of tangible objects bring several benefits to a user. Firstly, by using touch, one of our natural abilities, it is possible to reduce the learning time spent on a 3D user interface. Secondly, tangible objects seem to improve the understanding of 3D space and help to maintain the spatial and temporal correspondence between the interaction devices and the virtual objects being manipulated, both of which are important factors while performing 3D data analysis.

The research questions being investigated in this thesis are quite broad and diverse. They reflect the fact that understanding scientific data through 3D interaction is rather complex. Providing solutions for the observed problems not only requires visualization algorithms to represent data visually or in other forms (for instance, haptically), but also forces us to take into account usability aspects of the interface provided for interaction. The perceptual and cognitive abilities of a human user play important roles in determining the success or failure in the diverse data analysis tasks that are required in the course of a (3D) interaction process. Designing a 3D interface or interaction technique in such a way that pitfalls of visual (spatial) perception are avoided and cognitive processes (such as memory) are supported is crucial. This thesis intends to contribute to the recognition of and the emphasis on a multidisciplinary and experimental approach that the author thinks is essential when trying to establish guidelines in this field. Advancing technology, introducing proper HCI models and performance metrics, proposing better evaluation methods and collecting more experimental evidence remain important objectives for future 3D interaction research.

# Appendix A

# Scientific Visualization

As a result of all these research activities, different algorithms (for example volume rendering), have become available to better assist scientists and professionals in their exploration of scientific data. Generally there are two kinds of algorithms (surface rendering and volume rendering) to represent 3D data visually. Additional algorithms are also available to make best use of the hardware accelerations that are provided by modern graphics boards.

## A.1 Surface Rendering

Surface rendering is a sub-area within geometric graphics. Geometric graphics is the major approach in contemporary 3D computer graphics. It employs a pipeline structure in which a display list of geometric primitives is created together with required transformations. The frame buffer that is used to provide an image to the display is updated from the information in this list after every change in the scene or in the viewing parameters. This structure is powered by geometry graphics hardware, as it has advanced in the past decade, making geometric graphics the state of-the-art in computer graphics, particularly in 3D graphics. Geometric graphics represents a scene as a set of geometric primitives (such as, lines, triangles or polygons). These primitives are transformed, mapped to screen coordinates, and converted by scan-conversion algorithms into a discrete set of pixels. Any change to the scene, viewing parameters, or shading parameters requires the scene generation system to

repeat this process. As vector graphics that did not support rendering the interior of 2D objects, 3D geometric graphics generates merely the surfaces of 3D objects and does not support the rendering of their interior structures.

Surface rendering includes a class of algorithms to extract geometry primitives from volumetric data, and evolved from geometric graphics. The main idea of surface rendering is to design algorithms that extract isosurfaces from 3D scalar field data efficiently and unambiguously. The position of an isosurface, as well as its relation to other neighboring isosurfaces, can provide clues to the underlying structure of the scalar field [WVG92], [IK95]. Marching Cubes is the most widely used algorithm [LC87]. The advantage of this method is that it decreases the amount of data that needs to be processed and it is easy to render such data in contemporary graphics cards. There are also other algorithms that derived from this method, for instance Marching Tetrahedrons [PT90].

## A.2 Direct Volume Rendering

Direct volume rendering is another method of displaying 3D volumetric data as a 2D image. It is more straightforward compared to the surface rendering method. The individual values in a data set are made visible by an assignment of optical properties, like color and opacity, which are then projected and composited to form an image [DCH88], [Lev88]. As a tool for scientific visualization, the appeal of direct volume rendering (in contrast to other rendering techniques such as isosurfacing) is that no intermediate geometric information needs to be calculated, so the process maps from the data set directly to an image. Therefore, direct volume rendering has the advantage of rendering all of the data without throwing away any information. There are two popular rendering algorithms: ray casting and splatting.

### A.2.1 Hardware-accelerated Volume Rendering

Due to the huge amount of data processing required for 3D rendering, creating visual representation of such data often relies on hardware to improve rendering speed. There are two approaches to using hardware acceleration: using customized hardware or using general-purpose hardware. The first approach is done through designing special graphics hardware; the second approach uses the texture mapping available in current graphics cards. Detailed

reviews of these techniques can be found in Chapter 2.

## A.3 Haptic rendering

Haptics is the study of how to couple the human sense of touch with a computer-generated world or how to represent and interact with scientific data through touch. The theoretical and practical basis of such development is that human brain uses multiple sources of sensory information derived from several different modalities, including vision, touch and audition to perceive the external world.

Force feedback is one area of haptics that deals with devices or objects that interact with the muscles and tendons of a user. It gives a user a sensation of a force being applied. These devices mainly consist of robotic manipulators, which push back against a user with the forces. These forces correspond to the environment that the virtual effector is in.

Tactile feedback deals with the devices or objects that interact with the nerve endings in the skin, which sense heat, pressure, texture and etc. These devices typically have been used to indicate whether or not the user is in contact with a virtual object. Some other tactile feedback devices have been used to simulate the texture of a virtual object.

Haptic rendering is the software-driven process that computes and generates forces in response to interactions with virtual objects, based on the position of the force feedback device. Haptic rendering of an object can be seen as pushing the device out of the object whenever it tries to move the inside of the object. The human sense of touch is sensitive enough to require a processing speed of at least 1K Hz in terms of haptic rendering. The haptic rendering needs to provide forces that push the user out of the object. The further inside of the object a user moves, the greater the force pushes him/her out. This makes the surface feel solid.

Several haptic rendering techniques have been developed recently to render virtual objects. Just as in 3D computer graphics, the representation of 3D objects can be either surface-based or voxel-based for the purposes of computer haptics. While the surface models are based on parametric or polygonal representations, volumetric models are made of

voxels. The existing techniques for haptic rendering with force display can also be categorized based on how the probing object is modeled [SS97]: (1) point-based, where the probe is modeled as a point. Exploring and manipulating real objects with only the tip of a stick is analogous to this kind of model. (2) ray-based, where the probe is modeled as a line segment. With this model, it is analogous to exploring and manipulating real objects with the entire length of a stick in addition to its tip. (3) a 3D-object, where the probe consists of a group of points, line segments and polygon. The type of modeling method used in simulations depends on the needs and complexity of the application. Currently, many algorithms exist to render different properties of an virtual object, for example the shape, surface texture, softness, or dynamics [BS01], [SB97], [SCB04].

## A.4 Auditory rendering

Technically speaking, auditory representation in VR systems is implemented by means of a sound system. However, in contrast to conventional sound systems, the auditory representation is non-stationary and interactive, i.e., among other things, dependent on listeners' actions. This implies, for the auditory representation, that very complex, physiologically-adequate sound signals have to be delivered to the auditory systems of the listeners, namely to their eardrums.

One possible technical way to accomplish this is via transducers positioned at the entrances to the ear canals (headphones). Headphones are fixed to the head and thus move simultaneously with it. Consequently, head and body movements do not modify the coupling between transducers and ear canals (so-called head-related approach to auditory representation) - in contrast to the case where the transducers, e.g. loudspeakers, are positioned away from the head and where the head and body can move in proportion to the sound sources (room-related approach). In any real acoustical situation the transmission paths from the sources to the ear-drums will vary as a result of the listeners' movements in relation to the sound sources- the actual variation being dependent on the directional characteristics of both the sound sources and the external ears (skull, pinna, torso) and on the reflections and reverberation present.

VR systems must take account of all these specific variations. Only if this task is

performed with sufficient sophistication will the listeners accept their auditory percepts as real - and develop the required sense of presence and immersion.

# Appendix B

# Questionnaire for User Study of Transfer Function Specification

## B.1 Experiment Instructions

### B.1.1 Introduction

In visualizing volumetric data, typically a transfer function is used to reconstruct the required structure. Transfer functions make different structure of a 3D data set visible by assigning optical properties, for instance transparency, to the numerical values that comprise the dataset. Figure B.1 emphasizes the different structures by changing the transfer function.

Finding a good transfer function is critical to producing an informative rendering. But setting only one variable is already a difficult task. Users with little experience in computers or the mathematical background of volume rendering are usually not able to handle complex paradigms for specifying the transfer functions. Therefore the relationships between the users interface and transfer function should be investigated.

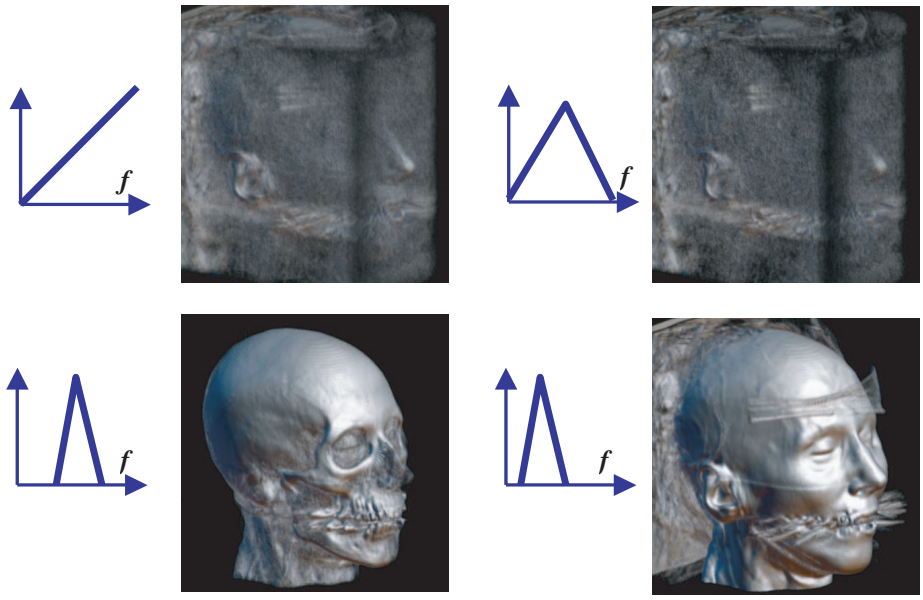



Figure B.1: The rendering results with four different transfer functions for CT scan data of a head.

### B.1.2 Overview of The Experiment

This experiment has been created to study the effects of different interfaces on the transfer function specification. The experiment will last about 25 minutes. It is very important that you remember that we are testing these techniques - we are not testing you. The procedure is the following:

1. Please finish reading this instruction sheet and if you have any questions do not hesitate to ask me.
2. Start to get familiar with the interfaces and the experiment set-up.
3. If you are still not clear how to use the system and how to perform the task, then ask me.
4. Start to do experiment. If you feel like being unable to finish the experiment, you are free to give up.

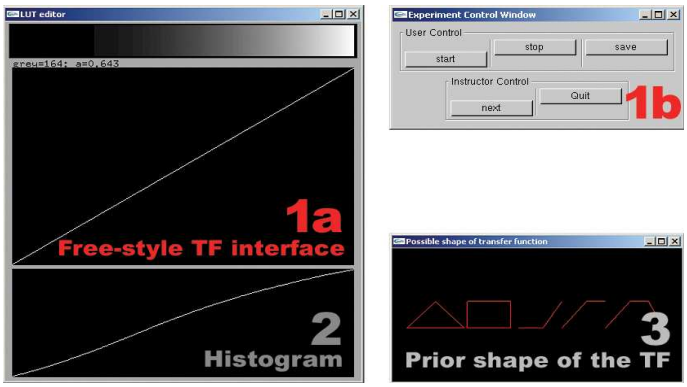



Figure B.2: The user interfaces for experimental conditions 1 (part 1), 2 (part 1+2), 3 (part 1+3) and 4 (part 1+2+3).

5. Upon completion there will be a questionnaires to fill in. We will ask you some questions about the techniques we are testing, the experiment itself, or your performance of the task.

# Thank you for your participation

### B.1.3 Interface

During this experiment, you will be given one user interface to find the ideal transfer function to highlight required structure. To adjust the transfer function, you have a graphical user interface. Figure B.2 shows the graphical interface for this condition. You could add control point by double clicking the left button of the mouse when the mouse cursor is on the curve. You can move the control point by holding the left button of the mouse down. The start and end point of the curve will already be there from the beginning and could only be moved up and down. You can use the "delete" key to delete the control point you add. For each task, when you are ready, please pressing "start" button. After you think you get the structure we ask for and are satisfied, press "stop" button. You also need press the "save" button to save your result.

### B.1.4 Task

With this interface, you will be given to four tasks. At the beginning of each task, you can see the initial visualization result. You would use the graphic interface to search transfer function for the required results. The requirements for these four tasks are the following:

- For the human head, your aim is to emphasize the bone structures.
- For the engine, your aim is to find the complete two circular parts inside this engine.
- For the lobster, your aim is to render the complete lobster structure.
- For the foot, your aim is to highlight the bone structures under the skin.

After experiment, you need compare and rank you results (saved as images) with original requirement from us in order to evaluate your own performance through this interface.

## B.2 General Questionnaire

**(Date: time: )**

1. Specify your job title, if any. If you are a student, indicate your class and major.

---

2. What is your age _____

3. Are you:

a) male

b) female

4. Are you:

a) right-handed

b) left-handed

c) ambidextrous

5. How often do you use a computer? (Circle the best answer)

a) Daily

b) A few times a week

c) A few times a month

d) Rarely or never

6. How often do you play a computer game? (Circle the best answer)

a) Daily

b) A few times a week

c) A few times a month

d) Rarely or never

7. What computer platform(s) are you familiar with? (Circle all that apply)

a) PC

b) Macintosh

c) UNIX workstations

d) Other

8. Which, if any, of these input devices are you familiar with? (Circle all that apply)

a) keyboard

b) mouse

c) joystick

d) touch screen

e) pen/stylus (e.g. Apple Newton, PalmPilot)

f) drawing tablet

g) 3D input devices (e.g. trackers, 3D mice)

h) Other _____

9. Did you have any knowledge about volume rendering and it's transfer function?

a) yes

b) no

c) some

10. Do you think it is hard to understand the transfer function?

a) yes

b) no

c) some

11. Have you ever used volume rendering or visualization application with transfer function?

a) yes

b) no

c) some

If so, Could you please describe the system and the input devices used below (use back if necessary): _____

## B.3 Subjective Evaluation Questionnaire

You will be asked to give the evaluation in different categories:

1. Question for Overall System Usability

Table B.1: Evaluation of the whole system (easy of use)

| **Rating** | Very easy (1) | 2 | 3 | 4 | Very difficult (5) |
|---|---|---|---|---|---|
| easy of use | | | | | |

Table B.2: Evaluation of the delay of the visual feedback

| **Rating** | Not noticeable (1) | 2 | 3 | 4 | Very slow (5) |
|---|---|---|---|---|---|
| delay | | | | | |

2. Questions for the interface

Table B.3: Evaluation of the interfaces for effectiveness (the amount of control over the TF)

| **Rating** | Least Effective (1) | 2 | Effective (3) | 4 | Most effective (5) |
|---|---|---|---|---|---|
| Interface 1 | | | | | |
| Interface 2 | | | | | |
| Interface 3 | | | | | |
| Interface 4 | | | | | |
| Interface 5 | | | | | |

3. Questions for Task Performance

Table B.4: Evaluation of the interfaces for TF control (efficiency to set a TF)

| **Rating** | Least Efficient (1) | 2 | Efficient | 4 | Most Efficient (5) |
|---|---|---|---|---|---|
| Interface 1 | | | | | |
| Interface 2 | | | | | |
| Interface 3 | | | | | |
| Interface 4 | | | | | |
| Interface 5 | | | | | |

Table B.5: Evaluation of the interfaces' look and feel for TF control

| **Rating** | Least Satisfied (1) | 2 | Satisfied | 4 | Most Satisfied (5) |
|---|---|---|---|---|---|
| Interface 1 | | | | | |
| Interface 2 | | | | | |
| Interface 3 | | | | | |
| Interface 4 | | | | | |
| Interface 5 | | | | | |

Table B.6: Overall rating of the interfaces

| **Rating** | very bad (1) | 2 | 3 | 4 | very good (5) |
|---|---|---|---|---|---|
| Interface 1 | | | | | |
| Interface 2 | | | | | |
| Interface 3 | | | | | |
| Interface 4 | | | | | |
| Interface 5 | | | | | |

Table B.7: Evaluation of the tasks (easiness)

| **Data** | Easiest(1) | Easier | Normal | Difficult | Very difficult (5) |
|---|---|---|---|---|---|
| Head | | | | | |
| Engine | | | | | |
| Lobster | | | | | |
| Foot | | | | | |

Table B.8: Evaluation of understanding the TF concept

| **Rating** | Not understand at all (1) | ... | Completely understand (5) |
|---|---|---|---|
| curve information | | | |

Table B.9: Evaluation for the usage of the cumulative histogram

| **Rating** | Not helpful at all (1) | ... | Very helpful (5) |
|---|---|---|---|
| curve information | | | |

Table B.10: Evaluation for the usefulness of the additional curve information

| **Rating** | Not helpful at all (1) | ... | Very helpful (5) |
|---|---|---|---|
| curve information | | | |

Table B.11: Evaluation for the usage of the additional curve information

| **Rating** | Don't use at all (1) | ... | Use all the time (5) |
|---|---|---|---|
| curve information | | | |

# Appendix C

# Questionnaire for User Study of Virtual Reality

## C.1 Participant Information Sheet

Effectiveness of Virtual Reality System on Volumetric Data Observation

**Participant Instructions**

As a participant in the study "Effectiveness of Virtual Reality System on Volumetric Data Observation", you will do several things. This document describes what you will be doing and gives you instructions. The investigators will elaborate on the instructions and answer any questions you have.

**Part I: Preliminaries (Conference Room)**

- We will ask you again if you meet all the qualifications to be a participant in this study.
- We will explain the entire experiment to you by going through these instructions and answering any questions that you have.

- We will show you the equipment the experiment will use:
  - If you are assigned to the immersive virtual reality (VR) system, you will use the head-mounted display (HMD);
  - If you are assigned to the "fish tank" VR system, there will be a monitor, shutter glasses and haptic device only for manipulation;
  - If you are assigned to the "fish tank" VR system with additional haptic devices there will be a monitor, shutter glasses and haptic device for manipulation and force feedback;

  Please ask any questions you have about them.

- We will give you a consent form that describes aspects of the study not included in these instructions. You'll read this form and be asked if you have any more questions about the study.

- After all your questions are answered; we will ask you if you are willing to sign the form agreeing to be a participant in this experiment. Once you have signed the consent form, you are officially a participant in this experiment and you are entitled to payment if you finish 80

**Part II: First Questionnaire Session**

The first questionnaire you'll fill out asks basic information about your age, gender, etc. You also need describe the general state of your health in this short questionnaire. Please answer the questions carefully and thoughtfully; your answers are a key element in making our study produce meaningful and useful results. When you've finished the questionnaires, we will again review the instructions for the VE part of the experiment, and we'll move into the laboratory.

**Part III: The VE Session**

**A. Training**

**If you have been assigned to use immersive VR, please read this paragraph.**

1. We'll put you in the HMD and adjust it so that it fits snugly and comfortably.
2. When you put the HMD on, you'll find yourself in a virtual environment with some spheres, ellipsoids, and other objects.
3. We'll help you adjust the HMD so that you can see the images properly and in stereo.
4. The entire virtual environment you will visit fits inside the open space of the laboratory we will be in. You won't be able to see the real laboratory once you put the HMD on, so one of the investigators will stay near you all the times to be sure that you don't trip on or bump into anything.
5. Next we will train you in how to move in the virtual environment. Your walking in the virtual environment exactly corresponds to your walking within the real environment. When you take a step in the real world, you move the length of that step in the virtual environment. When you stop walking, your viewpoint in the virtual environment also stops.
6. You can stay in the training room until you're comfortable moving around in the virtual environment and you are comfortable with observing the objects. Take as long as you need, up to the maximum session length of 5 minutes.
7. After we start the experiment we will try not to interact with you at all; we will tug on your cables, speak to you, or touch you gently on the shoulder only if we must help you stay in the proper part of our laboratory.

**If you have been assigned to "fish tank" VR without haptics, please read this paragraph.**

1. You will wear shutter glasses and hold a stylus during your VE session. We will show you how to use the stylus from the haptic device and its buttons.
2. We'll put you in the chair and adjust it so that it fits you comfortably.
3. When you put the shutter glasses on, you'll find yourself in a virtual environment with one cubic volume with spheres and other objects inside.

4. You'll need to adjust the position of your head in the front of the monitor so that the head tracker can detect the movement of the shutter glasses you are wearing and properly show the objects in stereo in the right position.

5. The entire virtual environment you will see fits inside the monitor. You will still be able to see the real laboratory even you put the glasses on.

6. Next we will train you in how to move in the virtual environment. Movement of your head within the real environment will cause the relative movement of the objects in the virtual environment exactly corresponding to your movement. When you move forward or backward in the real world, the virtual objects will appear to stay stationary in the virtual environment. When your head is not in the range of tracker, you may lose stereo and tracking. Try several times so that you will feel what it looks like when the stereo and tracker fail.

7. Now we will train you how to manipulate the volume. We will show you the button before you put the glasses on. As long as you hold the button down and rotate the stylus, you will continue "manipulating" (rotating) the volume in your hand.

8. You can stay in the training session until you're comfortable moving your head around in the virtual environment and rotating the objects. Take as long as you need, up to the maximum session length of 5 minutes.

9. When we start the experiment we will try not to interact with you at all; we will speak to you, or touch you gently on the shoulder only if we must help you perform the experiment.

**If you have been assigned to "fish tank" VR with haptic device, please read this paragraph.**

1. You will wear shutter glasses and hold a stylus from a haptic device during your VE session. We will show you how to use the stylus and its buttons.

2. We'll put you in the chair and adjust it so that it fits you comfortably.

3. When you put the shutter glasses on, you'll find yourself in a virtual environment with one cubic volume with spheres and other objects inside.

4. You'll need to adjust the position of your head in the front of the monitor so that the head tracker can detect the movement of the shutter glasses you are wearing and properly show the objects in stereo in the right position.

5. The entire virtual environment you will see fits inside the monitor. You will be able to see the real laboratory even you put the glasses on.

6. Next we will train you in how to move in the virtual environment. Movement of your head within the real environment will cause the relative movement of the objects in the virtual environment exactly corresponding to your movement. When you move forward or backward in the real world, the virtual objects will appear to stay stationary in the virtual environment. When your head is not in the range of tracker, you may lose stereo and tracking. Try several times so that you will feel what it looks like when the stereo and tracker fail.

7. Now we will train you how to touch and manipulate the volume. Firstly, we will let you try out the dice demo from the device manufacturer. You will know how to manipulate the haptic device essentially. During the experiment, you can move the cursor by moving the stylus. When an object intersects the cursor, you will feel force feedback from the stylus. You will not be able to punch through the objects like sphere, ellipse. You will fall through the inside when the cursor touches the tube shape object, for example the curved tube or cylinder. We will have shown you the button before you put the glasses on. As long as you push the button down and move the stylus, the cursor will escape from the tube and you could continue "manipulating"(rotating) the volume in your hand.

8. You can stay in the training session until you're comfortable moving around your body in the virtual environment and feeling and rotating the objects. Take as long as you need, up to the maximum session length of 5 minutes.

9. When we start the experiment we will try not to interact with you at all; we will speak to you, or touch you gently on the shoulder only if we must help you perform the experiment.

**B.Task**

1. Your task in the virtual environment is to observe the objects inside the volume, describe the topology of the objects, identifying the distribution of specific shape of the object (for example, the densest region of the spheres), judge whether two objects are connecting with each other (for example, a question like "please indicates the area where the longest curved tube crosses" will be asked).
2. You will be given a screen print-out to tell you what kind of objects you need to identify.
3. You will have 21 trials to do, each with one volume data set, in which the first five for training purpose. During each trial, we'll start a timer when you put the shutter glasses/HMD on. Even if you have not finished the experiment, we will stop this trial session after 5 minutes.
4. When you've finished all trials, you're done. We will help you take the HMD/shutter glass off and return to the conference room. You can also ask to discontinue the experiment at any time if you do not feel comfortable.
5. You will answer the post-experience questionnaires and debrief at the conference room.

**Part IV: Second Questionnaire Session**

You'll fill out second questionnaire after the whole task session. Please answer the questions carefully and thoughtfully; your answers are a key element in making our study produce meaningful and useful results. The questionnaire will ask you about how you are feeling about the task and the VR system.

**Part V: Debrief Session**

When you've finished the questionnaires, the investigator will ask you if you have any other comments about the experience or questions that you'd like to ask.

## C.2 Pre-experiment Questionnaire

**Group No: Participant ID:**

**A. General questions:**

1. Gender, Age, and Race/Ethnicity:

☐ Male ☐ Female

Age _____

Race/ Ethnicity:

☐ American Indian or Alaskan Native

☐ Asian or Pacific Islander

☐ Black, not of Hispanic Origin

☐ Hispanic

☐ White, not of Hispanic Origin

☐ Other _____

2. What is your University status?

My status is as follows:

☐ Undergraduate student

☐ Graduate student

☐ Research Associate

☐ Staff member - systems/technical staff

☐ Faculty

☐ Administrative staff

☐ Other (please write in) _____

3. To what extent do you use a computer in your daily activities?

I use a computer...

(Not at all) 1 2 3 4 5 6 7 (Very much)

4. Which input devices for a computer do you use regularly?

☐ None

☐ Mouse

☐ Joystick

☐ Touch pad (of laptop)

☐ Stylus (handheld like PalmPilot)

☐ Drawing tablet

☐ Console controller

☐ 3D input devices (e.g. trackers, 3D-mouse), namely

☐ Other: _____

5. To what extent do you play computer games?

I play computer games...

(Not at all) 1 2 3 4 5 6 7 (Very much)

6. Which kind of Virtual Reality (VR) have you used, if any (check all that apply)?

☐ Head-Mounted Display

☐ Desktop system with stereoscopic viewing (possibly with 3D input device)

☐ Large stereoscopic projection screen (standard, curved or multiwall)

☐ Responsive Workbench (or similar table-based stereoscopic display system)

☐ Augmented Reality (see-through glasses)

☐ None

7 If you have experienced immersive virtual reality (with HMD):

I have experienced HMD virtual reality...

(Never before) 1 2 3 4 5 6 7 (A great deal)

8. If you have experienced Fish tank virtual reality (with shutter glass) before:

I have experienced "Fish Tank" virtual reality...

(Never before) 1 2 3 4 5 6 7 (A great deal)

9. If you have experienced Phantom or other force feedback device before?

I have experienced Force-feedback virtual reality...

(Never before) 1 2 3 4 5 6 7 (A great deal)

**B. Participant Health Questionnaire:**

Please circle your answers to the following questions:

1. How many hours per week do you exercise?

During an average week, I exercise...

☐ Less than 0.5 hours

☐ 0.5 hours

☐ 1 hour

☐ 1.5 hours

☐ 2 hours

☐ 2.5 hours

☐ 3 or more hours

2. Do you or your family members have a history of epilepsy?

☐ Yes ☐ No

3. Are you colorblind or have an unusual sensitivity to slight differences of color?

☐ Yes ☐ No

If you answered yes please explain briefly in the space below: _____

4. Do you use glasses or contact lenses?

☐ Yes ☐ No

- 4.1 Are you nearsighted? ☐ Yes ☐ No
- 4.2 Are you farsighted? ☐ Yes ☐ No

5. Are you in your usual state of good fitness (health)? ☐ Yes ☐ No

If not, please explain: _____

6. In the past 24 hours, which, if any, of the following substances (including alcohol) have you used? Please select all that apply.

☐ Sedatives or tranquilizers

☐ Decongestants

☐ Anti-histamines

☐ Other

**C. Further Comments**

Please write down any further questions that you wish to ask about the experiment.

_____

_____

**D. Reminder - all answers will be treated entirely confidentially.**

Thank you once again for participating in this study and helping with our research. Please do not discuss this with anyone until the end of the semester. This is because the study is continuing, and you may happen to speak to someone who may be taking part.

## C.3 Post-experiment Questionnaire

**Group No: Participant ID:**

The following questions relate to the Virtual Reality (VR) system your have experienced during the experiment. Please select the correct one:
☐ Immersive HMD VR
☐ Fish Tank VR
☐ Fish Tank with Force-feedback (haptic)

**A. Experiment experience**

1. Please rate your sense of being in the virtual environment that has the simulation data on the following scale from 1 to 7, where 7 represents your normal experience of being in a place. I had a sense of being in the virtual environment containing the simulation data:
(Not at all) 1 2 3 4 5 6 7 (Very much)

2. Please rate any sense of immersion you experienced when looking into the data set. The sense of immersion I experienced was...
(Not at all) 1 2 3 4 5 6 7 (Very much)

3. How difficult or straightforward was it for you to understand spatial relationships between objects in the virtual environment while working with the system? The spatial relation was...
(Very difficult) 1 2 3 4 5 6 7 (Very straightforward)

4. Did you find it relatively simple or relatively complicated to move through the virtual environment and the simulation data? To move through the virtual environment was...
(Very complicated) 1 2 3 4 5 6 7 8 9 (Very simple)

5. The act of moving from place to place in the virtual environment can seem to be relatively natural or relatively unnatural. Please rate your experience of this. The act of moving from place to place seemed to be...
(Very unnatural) 1 2 3 4 5 6 7 (Very natural)

6. How often did you to feel you were in virtual environment when observing the simulation data and searching for the required structure? I felt I was in virtual environment...
(Very few) 1 2 3 4 5 6 7 (Very often)

7 To what extent were there times during the experience when you felt dissatisfied with the interface? There were times during the experience I felt dissatisfied...
(At no time) 1 2 3 4 5 6 7 (Almost all of the time)

8. To what extent did you have a sense of acting in the virtual space, rather than operating something from outside?
(Not at all) 1 2 3 4 5 6 7 (Very much)

9. To what extent did you feel that the virtual environment surrounded you?
(Not at all) 1 2 3 4 5 6 7 (Very much)

10. To what extent did you feel like you just perceived pictures?
(Not at all) 1 2 3 4 5 6 7 (Very much)

11. How much did your experience in the virtual environment seem consistent with your real-world experience?
(Not at all) 1 2 3 4 5 6 7 (Very much)

12. To what extent do you think the virtual reality system you experienced helped you identify the structure within the volumetric simulation data? I thought the virtual environment helps me...
(Not at all) 1 2 3 4 5 6 7 (Very much)

13. How easy to use was this virtual reality system?
(Hard to use) 1 2 3 4 5 6 7 (Very easy)

14. How effective do you feel you were when working with the virtual reality system compared with a traditional desktop system?
(No difference) 1 2 3 4 5 6 7 (Very effective)

15. How helpful were the depth cues in this virtual reality system compared to the traditional desktop system?
(No difference) 1 2 3 4 5 6 7 (Very much)

16. Rate the degree of difficulty in carrying out the task, for the virtual reality system you experienced:
(Not difficult at all) 1 2 3 4 5 6 7 (Very difficult indeed)

17. How well could you examine objects from multiple viewpoints?
(Very difficult) 1 2 3 4 5 6 7 (Very easy)

18. How much delay did you experience between your actions and expected outcomes?
(None at all) 1 2 3 4 5 6 7 (Very much)

19. When exploring the virtual space, did the objects appear too compressed or too magnified?
(Not at all) 1 2 3 4 5 6 7 (Very compressed)
(Not at all) 1 2 3 4 5 6 7 (Very magnified)

20. During the experiment, your general level of confidence in the accuracy of your answers was:
(Just guessing) 1 2 3 4 5 6 7 (Very sure)

21. During the task, identifying the individual shape and location of an object within the environment was:
(Difficult) 1 2 3 4 5 6 7 (Easy)

22. During the task, identifying the global topology of the simulation data was through the VR system:
(Difficult) 1 2 3 4 5 6 7 (Easy)

23. During the task, what was the level of demand on your memory?
(Small) 1 2 3 4 5 6 7 (Large)

24. How effective was the sense of perspective (further objects appeared the correct size compared to nearer objects)?
(Ineffective) 1 2 3 4 5 6 7 (Effective)

25. Did you think that the VR system you experienced change the way you observe and analyse the data - in comparison to traditional media?
☐ Didn't change at all
☐ Changed just a little bit
☐ Changed slightly
☐ Changed quite some
☐ Changed radically

## Answer these questions if experienced Fish Tank VR with haptic device

1. How consistent or inconsistent was the information coming from your various senses (visual and haptic feedback)?
(Inconsistent) 1 2 3 4 5 6 7 (Very Consistent)

2. How well could you actively survey or search within the virtual environment using touch? Searching within the virtual environment through touch was
(Very difficult) 1 2 3 4 5 6 7 (Very easy)

3. Do you think the addition of multimodal sensory stimulation (haptics) would help you?
A) Create a better understanding of space:
☐ Not at all

☐ Small chance

☐ Possibly

☐ Likely

☐ Very likely

B) Create better understanding the structure and topology of the data:

☐ Not at all

☐ Small chance

☐ Possibly

☐ Likely

☐ Very likely

**B. Further Comments**

Please write down any further comments that you wish to make about the experiment. In particular, what things helped you finish the task, and what things were needed but missing? Also describe the features of the system you do not like.

# Index