Graz University of Technology
Institute for Computer Graphics and Vision

Dissertation

Unconventional 3D User Interfaces for Virtual Environments

Ernst Kruijff
Graz, Austria, October 2006

Thesis Supervisor
Prof. Dr. Dieter Schmalstieg
Graz University of Technology

Advisor
Prof. Dr. Steffi Beckhaus
University of Hamburg

ABSTRACT

In this dissertation, the potential of the human body will be investigated, with the aim to design, develop, and analyze new spatial interaction methods which surpass performance or application possibilities of currently available techniques. In contrast to desktop interfaces, spatial interaction methods potentially make use of all six degrees of freedom and are generally referred to as 3D user interfaces (3DUIs). These interfaces find wide applicability in a multitude of different kinds of Virtual Environments, ranging from those techniques that allow for free movement through a room with large, possibly stereoscopic displays, up to the usage of helmet-like or full-encompassing ("immersive") display systems.

Due to the experimental characteristics, most of the presented techniques can be labeled as being unconventional, even though many of the techniques can find great applicability in the more traditional work environments. Hence, through investigation of human potential, the design space of 3DUIs can be broadened.

More specifically, the basics of 3D User Interfaces and related terminology will be explored (chapter 1), after which an extensive and detailed look will be taken at the possibilities of the different human "input and output channels," relating the psycho-physiological possibilities to technology that is currently existent, or will be developed in the foreseeable future. A reflection on possible applications is included (chapter 2). In chapter 3, issues that are specific to designing and developing unconventional 3DUIs are investigated, ranging from the boundaries of human performance, specific human-computer interface matters, to social and technical issues. Following (chapter 4), a total of seven case studies illuminate multiple sides of designing, developing, and analyzing unconventional techniques, looking at both pure spatial and unconventional setups, and so called hybrid interface techniques. More specifically, *Shockwaves* and *BioHaptics* explore the usage of alternative haptic feedback, either through usage of audio and air-based shockwaves, or neuromuscular stimulation. Also dealing with haptics, *Tactylus* explores multisensory binding factors of a device using coupled visual, auditory, and vibrotactile feedback. The fourth study, *Cubic Mouse*, explores a prop output (control) device, resembling a coordinate system, in order to find specific performance advantages or flaws in comparison to generally used spatial controllers. It, thereby, makes use of a new spatial trajectory analysis method. The final three studies all focus on hybrid interfaces, integrating 2D and 3D I/O methods. *ProViT* deals with integrating a PenPC with a spatial pen device, and the Cubic Mouse to control engineering applications, focusing, foremost, on flow of action factors. *Capsa Arcana* are two consoles used in museum applications that integrate MIDI controllers and desktop devices to allow for more interesting and potentially unconventional control. Finally, with *Eye of Ra*, a new input device form is presented. The Eye of Ra has been specifically designed for closely combining the control of 2D and spatial actions for use in medical scenarios.

The final chapter concludes this dissertation by providing a short summary and reflection, including a road map of open issues and fields of further research.

CONTENTS

ABSTRACT 2
CONTENTS 3
PREFACE 6
CHAPTER 1 14
3D user interaction 14
1.1 What are 3D user interfaces? 14
1.1.1 Basic terminology 14
1.1.2 Which fields are related to 3DUIs? 15
1.1.3 Where and why to use 3DUIs? 16
1.2 Basic 3DUI tasks and techniques 17
1.3 3DUI Research directions and issues 19
1.3.1 Main research directions 19
1.3.2 Main goals 19
1.4 The relation between 3DUIs, human I/O and unconventional interfaces 20
1.5 Summary 21
CHAPTER 2 23
The human I/O system 23
2.1 Introduction to the human I/O channels 23
2.2 Human input 24
2.2.1 Vision 26
2.2.2 Audition 31
2.2.3 Somatic and kinesthetic 35
2.2.4 Chemical 40
2.2.5 Vestibular 44
2.3 Human output 47
2.3.1 Head – mouth and throat 51
2.3.2 Head – face and eyes 54
2.3.3 Hand and arm 56
2.3.4 Feet and legs 62
2.3.5 Body 66
2.3.6 Biopotential 69
2.4 Human Behavior 78
2.5 Summary 81
CHAPTER 3 83
Designing unconventional interfaces 83
3.1 Related fields of research 84
3.2 Human potential analysis 86
3.2.1 Multisensory versus multimodal processing 86
3.2.2 Limits and possibilities of human potential 87
3.2.3 Four stage analysis model 89
3.2.4 Classification 92
3.3 Sensory and control substitution 94
3.3.1 Human input substitution, addition and integration 96

3.3.2 Human output substitution, addition and integration 100
3.4 Interaction flow and feedback.. 103
3.4.1 Flow of action 103
3.4.2 Feedback mechanisms 105
3.5 Application and transfer issues 106
3.6 Developing devices using garage interface design methods.................. 111
3.7 Evaluation 114
3.8 Summary 116
CHAPTER 4 118
Case studies 118
4.1 Introduction... 118
4.2 Shockwaves... 119
4.2.1 Background 119
4.2.2 Related work 122
4.2.3 Interface design and development.. 123
4.2.4 Evaluation and results 127
4.2.5 Reflection.. 127
4.3 BioHaptics 128
4.3.1 Background 129
4.3.2 Related work 130
4.3.3 Experiment setup.. 130
4.3.4 Evaluation and results 132
4.3.5 Discussion and roadmap 133
4.3.6 Reflection.. 135
4.4 Tactylus 135
4.4.1 Background 136
4.4.2 Multisensory feedback binding 137
4.4.3 Interface design and development.. 138
4.4.4 Experiment setup.. 141
4.4.5 Evaluation and results 143
4.4.6 Reflection.. 145
4.5 Cubic Mouse 146
4.5.1 Background 146
4.5.2 Related work 147
4.5.3 Interface design and development.. 148
4.5.4 Experiment setup.. 148
4.5.5 Evaluation and results 149
4.5.6 Reflection.. 152
4.6 ProViT 153
4.6.1 Background and related work 153
4.6.2 Interface design and development.. 154
4.6.3 Evaluation and results 156
4.6.4 Reflection.. 158
4.7 Capsa Arcana 159
4.7.1 Interface design and development.. 159
4.7.2 Adding unconventional controls using MIDI sensors 163
4.7.3 Experiment setup.. 166
4.7.4 Evaluation and results 167
4.7.5 Reflection.. 168

4.8 Eye of Ra 169
  4.8.1 Background and related work 169
  4.8.2 Interface design and development 170
  4.8.3 Experiment setup 173
  4.8.4 Evaluation and results 174
  4.8.5 Reflection 177
4.9 Summary 178
CHAPTER 5 180
Conclusion 180
5.1 Reflection 180
5.2 Contributions 182
5.3 Road map 183
REFERENCES 185

PREFACE

Most of today's computer systems integrate several modalities in order to allow the user to interact with a computer. These systems are mostly focused on the usage of visual and auditory output and control via hand-coupled input devices. However, the human body offers many more input and output possibilities. These possibilities can enrich interaction with more traditional systems, and also give rise to new, more experimental kinds of systems that surpass today's systems on both functional and system technical level. Moreover, the possibilities of the human body can be used for interaction in so-called *Virtual Environments*. Virtual Environments (VE) are computer generated spatial environments that stimulate human senses in such a way that a certain feeling of "immersion" will be created, something like "being in another world".

The kind of interaction afforded by a VE is of a spatial nature, and takes place via a 3D user interface (3DUI). More specifically, this dissertation focuses on how the potential of the human body can be used for the design and development of new kinds of 3DUIs. In doing so, the biological or physiological characteristics of the separate parts of the body, from head to toe and from skin to heart, are explored, showing how their sensor (input) and control (output) capabilities can be applied in experimental, possibly unconventional techniques. The input and output possibilities are seen as "system," leading to the notion of a human I/O system (the "human processing unit"), thereby confronting many system-centered design approaches. Taking this approach does not mean that humans are seen as machines, even when sometimes the borders between both get very close. The majority of approaches described in this work are rather more human-centered. The point is that human potential often reveals incentives to develop a novel interface.

Using the term "unconventional" to describe interfaces is a bit tricky. The term refers to what the user does, how she does it, and by which means. It deals with experimental psycho-physiological ways of performing (general) tasks, performing tasks with experimental devices, or experimental tasks. Unfortunately, it is not always easy to clearly label techniques as unconventional. In order provide some more clarity, the table on the next page provides several possible axes of unconventionalism, the left side showing more conventional directions, the right side more unconventional approaches. All these directions can be found back when reading through the chapters of this work. The table is meant to give some clarity, not to provide a means of categorization – an analyzed technique might show multiple characteristics stated at the unconventional side and still be regarded as conventional. Since unconventionalism is in the eye of the beholder, it is best to view the table from the perspective of a general user that has a reasonable background with computer systems. Experienced researchers in the field of 3DUIs will find multiple known examples in this work. This should be no surprise due to the focus on spatial interfaces. Even so, this work introduces multiple directions that are highly novel and unexplored to many of these researchers too.

← conventional	**unconventional →**
Perception and control	
Hand-based control / audiovisual feedback	Alternative input and output
Stimulating dominant sensory systems (vision, audition)	Stimulating non-dominant sensory systems (haptics, smell, taste)
Provide low perceptual resolution	Provide high perceptual resolution (close to human sensory abilities)
Single or limited input and output modalities	Multiple input and / or output modalities, up to full-body interfaces
Standard methods for control and perception	Control or sensory substitution
Technology / hardware	
Using well-known technical solutions	Using new technology or available technology in new way
Using technology with known usage effects / capabilities	Using technology with unknown usage effects / capabilities
Desktop / 2D interfaces	3D (spatial), free movement interfaces
Separation between computer interface and human environment	Integration of computer interfaces in human environment
Non invasive	Invasive
Application	
Using techniques applied by a large user group	Using techniques applied by a small user group
Techniques for daily tasks	Techniques for special or new tasks
Using socially and ethically accepted techniques	Using socially and ethically questionable techniques
Daily-life solutions	Artistic solutions
Single person, non-distributed usage	Large group, possibly distributed usage
Using natural metaphors	Using “magical” / unnatural metaphors

Possible axes of unconventionalism

As will be illuminated in this dissertation, most unconventional interfaces have experimental human and system sides. The ultimate goal is to find out how the potential of the human body can be used to design, develop and analyze new spatial interaction methods that surpass performance or application possibilities of currently available techniques. The theme surpasses general “innovative” interfaces that mostly enhance current techniques into new dimensions.

In order to investigate such kinds of unconventional interfaces, several questions are addressed in this dissertation:

- What is the potential of the human input and output channels from a human-computer interaction perspective?
- How and why can we use this potential in 3DUIs?
- How does human potential drive the design of new and unconventional (hardware) interfaces?
- Which implications can be derived from the usage of this potential in spatial computer generated or adapted environments?

Why do I write about these topics? Probably the most important reason is sheer interest in the potential of the human body in human-computer interfaces. Even more important, I hope that 3D user interfaces can be further enhanced, in ways that interaction becomes better and “richer” (accurate, exciting) through more advanced control and feedback possibilities. I foresee new ways of interaction that are currently hardly covered, as well as new or changed fields of application.

Organization

In order to advance through the topics involved in unconventional interfaces, this dissertation is separated in two parts: a theoretical and a practical part. Within the theoretical part, I explain the basics of 3D user interaction in a virtual environment (chapter 1): what is a virtual environment, how can we interact with it, what are the basic problems, and finally, what is the connection between 3D user interfaces and human I/O channels? The human I/O system is also the core of chapter 2, in which I explore the potential of the human body. I describe the different human sensory channels at a basic level, and how they can be connected with by devices, and for which purpose (application).

Within the practical part, I explain how one can design unconventional control (human output) and human input methods: which approaches can be taken, where can they be applied, how can they be created, and what do developers need to think about when methods or devices are created (chapter 3).

In chapter 4, I present seven case studies ranging from haptic-like devices, up to mixed (hybrid) interaction techniques, illustrating theoretical and practical aspects. Finally, in chapter 5, I summarize this dissertation, and reflect the work by providing a small road map for further research.

Boundaries

The topic of this dissertation is bound to several fields of research, thereby being highly interdisciplinary. Topics from the “high level” research field human-computer interaction (HCI) will be mixed with more specialized fields like virtual reality technology and 3D user interfaces, itself a mixture between HCI and virtual reality related themes. On the other hand, human factors (including bio-physiological descriptions) come into play when examining the potential of the human body.

The dissertation is *not* a complete and in-depth “medical” overview of the human body, and not every different kind of device that is currently available is described. Also, even though many HCI themes are mixed in, a complete design process of how to create unconventional techniques is not provided, since it largely overlaps with general (3D) interface design rules - only the specific factors are handled.

There are simply too many topics related to the theme of this dissertation making it impossible to explore all topics in full-depth. Hence, wherever possible, sources for further reading are provided to overcome possible chasms or to serve the reader when there is more interest in a specific topic.

Contributions

Several main contributions can be elicited from the body of work presented in this dissertation:

- Expanding the **design space**, identifying new possibilities for creating (unconventional) 3DUI techniques by providing a comprehensive **investigation**

of human I/O potential, analyzing its psycho-physiological background, available technology and possible application areas.

- Based on own experience, user studies and background investigations, **guidelines for designing and developing unconventional 3DUI techniques** are provided, next to providing ways for **porting** these interfaces to general / more traditional work environments through hybrid interface methods.

- **A new input device** (*Tactylus*) is presented which **combines visual, vibrotactile, and auditory feedback** to successfully support collision detection and texture recognition through application of sensory substitution methods.

- Two techniques are presented that make use of the potential of the human body to sense **haptic feedback via alternative methods**. *Shockwaves* makes use of generating pseudo-haptic sensations by using audio or air-based shockwaves. *BioHaptics* triggers muscle contractions by using neuromuscular stimulation.

- A **performance study** compares the *Cubic Mouse* with a Stylus and gloves and shows the strength and preference of users for using the "prop" device for controlling fine-grain actions, but also illuminates its deficiencies for coarse actions. The study also presents **a new trajectory analysis method** using 3D movement paths logged during the evaluation.

- Several studies focus on integrating unconventional interaction methods in more traditional, possibly desktop work environments by using **hybrid interface techniques**. The *ProViT* study predominantly focuses on flow of action factors, including the chunking of actions, device switching influences, and focal attention factors. *Capsa Arcana* investigates the usage of MIDI sensors in a more traditional console form for usage in public space. Finally, the *Eye of Ra* is a device that integrates 2D and 3D functionality in a rather radical physical (device-) form to support a surgical planning application.

All together, the body of work provides new ideas for designing, developing and analyzing unconventional 3D user interface techniques, hopefully being an incentive for further research.

Publication of results

This section provides a commented overview of the main publications in which results of this dissertation have been published, or on which chapters in this work have been based.

- BOWMAN, D., E. KRUIJFF, J. LAVIOLA and I. POUPYREV (2005). 3D user interfaces: theory and practice, Addison-Wesley.

- BOWMAN, D., E. KRUIJFF, J. LAVIOLA and I. POUPYREV (2001). An Introduction to 3D User Interface Design. Presence: Teleoperators and Virtual Environments 10(1).

  These two publications (book, journal publication) and related presentations (courses at IEEE Virtual Reality, ACM Virtual Reality Software and Technology, and SIGGRAPH) can be regarded as the standard reference for designing and developing 3DUIs and is the basis for the general factors concerning spatial interfaces in this dissertation (chapters 1 and 3).

- BECKHAUS, S. and E. KRUIJFF (2004). *Unconventional Human Computer Interfaces*. Course at SIGGRAPH 2004.

  In this course, for the first time the human potential view on designing and creating spatial interfaces was presented. It forms the basis for the overview on unconventional human computer interfaces (chapter 2) and some theoretical and practical reflections in chapter 3.

- KRUIJFF, E., D. SCHMALSTIEG and S. BECKHAUS (2006). *Using Neuromuscular Electrical Stimulation for Pseudo-Haptic Feedback*. Proceedings of the ACM Symposium on Virtual Reality Software & Technology 2006 (VRST 2006).

  Within this article, a study and roadmap was presented regarding the provision of alternative haptic feedback using neuromuscular electrical stimulation. It forms the basis for the BioHaptics study (section 4.3).

- KRUIJFF, E., G. WESCHE, K. RIEGE, G. GOEBBELS, M. KUNSTMAN and D. SCHMALSTIEG (2006). *Tactylus, a Pen-Input Device exploring Audiotactile Sensory Binding*. Proceedings of the ACM Symposium on Virtual Reality Software & Technology 2006 (VRST 2006).

  The results and background of the Tactylus study (section 4.4) were presented in this article. It describes how the Tactylus, a new input device, makes use of vibrotaction and visual feedback to enhance collision detection and texture recognition tasks by analyzing the results of two user studies.

- BORNIK, A., R. BEICHEL, E. KRUIJFF, B. REITINGER and D. SCHMALSTIEG (2006). *A Hybrid User Interface for Manipulation of Volumetric Medical Data*. Proceedings of the 2006 Symposium on 3D user interfaces (3DUI 2006), IEEE Virtual Reality Conference (VR2006).

  This article describes a hybrid user interfaces for a liver planning system, mixing 2D and spatial interaction techniques. It includes a description and evaluation of interfaces techniques connected to the Eye of Ra input device, described in section 4.8.

- KRUIJFF, E. and A. PANDER (2005). *Experiences of using Shockwaves for Haptic Sensations*. Proceedings of 3D user interface workshop, IEEE Virtual Reality Conference (VR2005).

In this article, the design and usage experiences of audio and air-based shockwaves for providing haptic feedback are described. It forms the basis for section 4.2.

- KRUIJFF, E., S. CONRAD, P. PALAMIDESE, P. MAZZOLENI, F. HASENBRINK, M. SUTTROP and Y.-M. KWON (2004). *Remote Virtual Guidance in Immersive Museum Applications*. Proceedings of the 2004 Conference on Virtual Systems and Multimedia (VSMM 2004).
- CONRAD, S., E. KRUIJFF, M. SUTTROP, F. HASENBRINK and A. LECHNER (2003). *A Storytelling Concept for Digital Heritage Exchange in Virtual Environments*. Proceedings of the 2003 International Conference on Virtual Storytelling.

These two publications describe the interaction framework for virtual guidance scenarios, integrating 2D and spatial interaction techniques and virtual storytelling mechanisms. The Capsa Arcana interface consoles presented in section 4.7 have been designed to employ this interaction framework.

- CAO, W., H. GAERTNER, S. CONRAD, E. KRUIJFF, D. LANGENBERG and R. SCHULTZ (2003). *Digital Product Development in a Distributed Virtual Environment*. VRAI, Proceedings of the SPIE.
- KRUIJFF, E., S. CONRAD and A. MUELLER (2003). *Flow of Action in Mixed Interaction Modalities*. Proceedings of HCI International.
- MUELLER, A., S. CONRAD and E. KRUIJFF (2003). *Multifaceted Interaction with a Virtual Engineering Environment using a Scenegraph-oriented Approach*. WSCG, Plzen.

In these three publications, the hybrid interface techniques used in the ProVit project are described. Focusing on the combination of 2D interfaces techniques (a GUI at a tablet PC) and spatial interaction techniques (using a tracked Stylus or a Cubic Mouse). Principles of hybrid interfaces are illuminated, which are used in sections 3.4 (flow of action) and 4.6 (ProViT).

Collaboration statement

Most of the research presented in this dissertation has been performed in cooperation with other research scientists and assistants. This section provides an overview of the main collaborators.

The general factors concerning the design and development of 3DUIs (especially chapter 1) is connected to a body of work performed in cooperation with Doug Bowman, Joe LaViola, Ivan Poupyrev, foremost presented in (Bowman, Kruijff et al. '05). The specific focus on unconventional techniques has found its way in this work through close cooperation with Steffi Beckhaus by ways of a course presented for SIGGRAPH 2004 (Beckhaus and Kruijff '04). Results of this cooperation have flown into chapter 2, and partially in chapter 3.

A large number of people cooperated in the case studies presented in chapter 4. At a general level, this included project members from two projects I directed, namely DHX and ProViT. For detailed design and implementation issues, closer cooperation was sought out with specific people.

For Shockwaves (section 4.2), Joachim Gossmann aided in building the first prototype (SoundMaster 2000) and performing initial tests. The further audio devices were created by myself and Aeldrik Pander, who also programmed the audio interfaces. Sebastian Reinders helped in the design and development of the air cannon.

Whereas the BioHaptics study (section 4.3) was performed solely by myself, a larger number of people aided in the Tactylus study (section 4.4). Gerold Wesche, Gernot Goebbels and Martijn Kunstman worked with me on the ergonomic design and development of the device. Kunstman created the CAD models and made the final prototype. In order to interact, Kai Riege implemented a collision detection module and created the I/O interfaces to connect to the buttons and the vibration element of the device. Riege also set up the technical side of the experiment. Finally, Gerhard Eckel aided in the creation of the audio files for the audio feedback, for which we made use of the audio interface module provided by Juergen Wind.

The Cubic Mouse study (section 4.5) is based on collaboration with the inventors of the device, Bernd Froehlich and John Plate, and also Jakob Beetz and Hartmut Seichter. The latter two greatly aided in the design and development of new interaction techniques, programmed most of the interfaces, and performed the experiments with me.

For the interaction concepts presented in the ProViT study (section 4.6), I was supported by Stefan Conrad and Arnold Mueller, who designed and programmed the GUIs for the Tablet PC interface. Weiqun Cao directed several scenario developments (applications) we made use of during experiments.

The design and development of the consoles and interaction devices in the Capsa Arcana (section 4.7) was done by myself, and Albert Baumhauer programmed the dialog interface that I designed visually, using modules developed by Johannes Strassner. Klaus Troche and Pavel Frolov implemented the audio and video streaming methods used to connect to the communication devices. Berthold Kirsch greatly supported the setup of the DHX evaluation in Milan.

Finally, for the development of the Eye of Ra, I was greatly supported by Alexander Bornik and Thomas Pock, with whom I made the design and the final model of the device. Programming was mainly done by Bornik, whereas, especially during the discussion on the experiment we were advised by Reinhard Beichel and Dieter Schmalstieg.

This work is partially sponsored by the EU (project DHX EU-IST 2001-33476), the BMBF (project ProViT BMBF-01AK939), and the Austrian science fund FWF (under grants P17066 and Y193).

Acknowledgements

Extending the previous section, I'd like to thank a large amount of people who aided in my research efforts. First of all, I would like to thank Dieter Schmalstieg for accepting me to do my dissertation under his direction at TU Graz, for all the great comments and help provided both during the writing of this dissertation, and its related publications. Being at TU Graz has been … and is a great pleasure!

Secondly, I would like to thank Steffi Beckhaus for being the second advisor (and thereby great "commentor") of this work, and the extensive cooperation performed for the unconventional interface course at SIGGRAPH in 2004, which has been a great aid in writing this work.

My praise also go to the “book buddies” Doug Bowman, Joseph LaViola and Ivan Poupyrev, with whom I have spent a huge amount of time discussing, writing, and presenting both the 3DUI courses and book since 1999. You all have been an incredible source of inspiration and help. Thereby I should also not forget Mark Mine, with whom I cooperated during multiple occasions, including the SIGGRAPH course in 2000.

A large part of this work has been performed at GMD/Fraunhofer IMK since 1999, for which I am thankful to whole the group of researchers for their aid and comments. I specifically would like to thank Gerold Wesche, Gernot Goebbels, Aeldrik Pander, Bernd Froehlich, Stefan Conrad, Joachim Gossmann, Arnold Mueller, Sebastian Reinders, Kai Riege, John Plate, Albert Baumhauer, Berthold Kirsch, Weiqun Cao, Juergen Wind, Gerhard Eckel, Pavel Frolov, Johannes Strassner, Andreas Simon and Klaus Troche for their cooperation in projects, and extensive discussions and comments on the work performed. Thereby, I also would like to thank Martin Goebel and Manfred Bogen for being great bosses, and giving me the space to perform the research in the labs. Thanks also go to TU Graz: I’d like to thank the group there, especially Alexander Bornik and Thomas Pock, with whom I have spent numerous hours on designing and creating the Eye of Ra. I also would like to thank all the people who were close to our rooms, because of their patience with both the smell and noise during production. My thanks also go to some people from Bauhaus-University Weimar, especially Holger Regenbrecht, Hartmut Seichter, Jakob Beetz and Stefan Hansen for discussions and cooperation on several studies, and the rest of igroup both for idea discussions, and for plain fun. Furthermore, some more people dispersed around the world have aided in this work, including Rob Lindeman and Martijn Kunstman, for which I am grateful. Also, I would like to thank Jan van Leeuwen and Joske Houtkamp whom indirectly aided in accomplishing this dissertation at TU Graz. In addition, I hereby also would like to thank *anyone* whom I might have forgotten… including those who worked in the projects DHX and ProViT, and the people who aided during the 3DUI and UHCI courses, including the 3DUI mailing list members.

A big “extra” thanks goes to Kris Blom, not only for commenting the SIGGRAPH 2004 course, but also for getting out those nasty grammatical mistakes… greatly improving and, thereby, “polishing” this work.

I’d like to thank my parents, for giving me the chance to study, even though the field of interest has slightly changed. This might also have been caused by my brother, who introduced me to the “world of SGI,” and inspired me to do research on multiple occasions and in different ways, for which I am thankful – you’re a great brother! At that point I’d also like to thank Ivana, for being a great sister in law and her patience in coping with me and my brother.

Finally, I would like to thank Christiane, for being patient and supportive during and outside the time when I was writing this work and for being there as my partner in life.

CHAPTER 1
3D user interaction

In the field of desktop applications, the user interface, the way a user “communicates” with a computer, has received a huge amount of attention. Since the creation of the Xerox STAR interface in the 80’s, a vast amount of research has been put into creating effective user interface techniques and metaphors. Even though Virtual Reality (VR) systems have been around a long time, in comparison to desktop interfaces, this field of interest has received much less attention. Dealing with spatial digital worlds and interfaces, the first efforts in creating VR systems started around 1965 with Sutherland’s head-mounted display construction (Sutherland '65). The real advent of VR systems development has started about a decade ago. Next to improving 3D hardware, developers are now also focusing on the creation of 3D user interfaces (3DUIs) to provide users tools and techniques to interact in a Virtual Environment.

This chapter provides a concise overview of the main issues of 3DUIs, stating the main terminology, tasks and techniques, and research directions. It also includes a basic explanation of how a 3DUI relates to investigating the potential of the human biological / physiological systems. This chapter does not focus on providing detailed descriptions of 3DUI techniques or I/O devices. For a more comprehensive overview of these issues, please refer to either (Bowman, Kruijff et al. '01) or the extensive source (Bowman, Kruijff et al. '05).

1.1 What are 3D user interfaces?

Nowadays, most people are familiar with using a desktop computer system, controlling monitor-based applications with a mouse and keyboard through a basic set of interaction techniques and metaphors. These techniques, metaphors, and devices are not always usable when working in spatial computer generated environments. In order to understand why, this section gives a brief introduction of the what-and-how’s of these kinds of environments and their ways of interaction.

1.1.1 Basic terminology

With the media hype of the nineties, the term “virtual reality” has been used for everything from a 3D game up to video-like installations. Hence, to avoid misunderstandings this term will not be used in this dissertation.

Rather, the key term is *Virtual Environment* (VE). A VE is a computer generated spatial environment that stimulates human senses in such a way that a certain feeling of “*immersion*” will (can) be created, something like “being in another world”. In order to create such a spatial environment, one makes use of a *Virtual Reality System*, a set of hardware components that allow users to provide input to a computer, and the computer to generate output towards the user. The medium through which the user communicates with a computer is called a *user interface.* Through the user interface, a user performs tasks using *interaction techniques* that couple hardware and software components, applying a specific metaphor. As a result, a 3DUI is a user interface that involves interaction within a spatial context.

There are different kinds of VEs. Mostly, these VEs are characterized by the balance between what can be perceived from the real world and computer generated world. There are VEs that strongly involve the real world, like *Augmented Reality s*ystems that enhance the real world by overlaying synthetic objects. In contrast, fully immersive systems exist in which the real world is more or less cut out. The whole range of systems is generally described as the *Virtuality Continuum*, a term coined by Milgram and Kishino (Milgram and Kishino '94). A graph of this continuum, often referred to as the Mixed Reality Continuum, can be seen below.

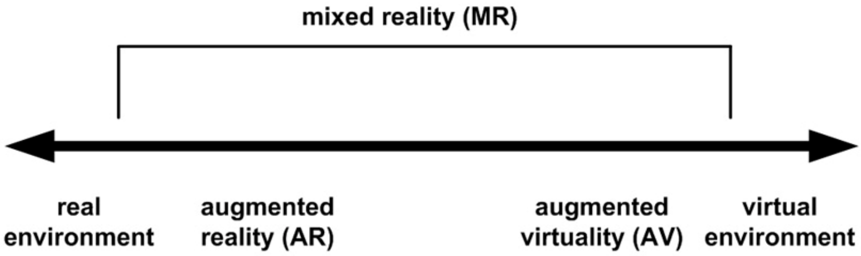


Figure 1.1: *The virtuality continuum.*

The described continuum should be regarded as a rather general view on VEs. For example, an art installation that makes use of a large projection wall and some kind of interaction that involves the full body may not seem to be a VE, but has many aspects that strongly overlap with the basic principles of a 3DUI or VE. The scope of VEs is not strictly limited to “traditional” systems that involve a Head Mounted Display or a large stereoscopic display system. Therefore, wherever unconventional systems are described, a reflection on how they relate to a 3DUI will be given.

For readers interested in reading more on what kind of different components VR systems are built from, it is recommended to read chapters 3 and 4 in (Bowman, Kruijff et al. '05) or (Sherman and Craig '03). Many components will be handled throughout the following chapters, but no complete overview will be given in this introduction. Additionally, the annotated 3DUI Bibliography (Poupyrev and Kruijff '00) provides a good start into the different thematic areas of 3DUIs.

1.1.2 Which fields are related to 3DUIs?

The field of 3DUI is highly interdisciplinary and draws upon many both human-oriented and technical areas.

First of all, basic principles from the field of HCI apply, ranging from general guidelines and interaction theories (Norman '90; Shneiderman '98; Preece '02), up to design and evaluation methods (Nielsen '93).

The design of 3DUI techniques and devices is highly affected by human factors, including spatial perception, cognition, and motor systems. Basic human factor sources include (Salvendy '97; Stanney '02).

At a technological level, the fields of Interactive 3D graphics come into play, including different rendering and visualization techniques (Tufte '90; Watt and Watt '92; Foley '96) and the design of 3D input and output devices. The field of 3D I/O techniques is highly scattered among multiple publications, of which a good amount is used throughout this dissertation when explaining device-level issues. It should be stated that in this dissertation an important role is given to devices that are not generally regarded as computer input or output devices, like brain scanning or temperature measurement hardware.

Finally, popular media has always played a role of inspiration, ranging from movies and TV series like Star Trek or Minority Report, books from Neal Stephenson or William Gibson, up to role playing games like Shadowrun.

1.1.3 Where and why to use 3DUIs?

VEs, and therefore also 3DUIs, have been applied in multiple application areas, each with their particular interest in working with spatial data. Probably the most widely spread is the usage of VEs within design, simulation and visualization, and training scenarios. Particular areas in which VR systems are used are the engineering area (including car, aerospace and architecture) and oil industry.

Directly related to the popularity of 3D games at desktop computers, the entertainment field quickly caught up with technological developments from the area of VEs. There are only a couple of really large scale examples that are well known in the application of VR systems in the entertainment field. Examples include the Spiderman the Ride at Universal Studios, and DisneyQuest at Disneyworld (Pausch, Snoddy et al. '96).

In some forms also entertaining, many artists have been using VE as a medium to create artistic impressions. Some of these installations have been shown at public events like SIGGRAPH or Ars Electronica.

Both as training and simulation tool or as ways of remote diagnosis, VR systems have driven multiple developments in the medical and psychiatric areas, including tele-medicine and the treatment of phobias.

Finally, VEs can be a great means for educations purposes, from explaining “how things work” up to designing 3D artifacts like architectural designs.

Desktop systems have become effective and well tested work environments, used for an incredible number of tasks in our daily life. So, why do we actually want to use a 3DUI when this desktop environment works so well?

The usage of VEs “promises” several improvements in comparison to desktop environments (Bowman, Kruijff et al. '05), some of them well documented and tested, some of them still based on hypotheses.

- *Task improvement*: for true 3D tasks, the usage of a 3DUI can mean a great advantage, since the task characteristics can match the characteristics of a spatial environment well. Especially in training and simulation scenarios, the usage of VEs has proven to be an advantage in comparison to desktop systems.
- *Better understanding:* there are complex 3D data sets that can be explored well within a VE, since the interactive examination often leads to new insights that are impossible to see or understand at a desktop system.
- *More fun:* VEs can truly “pull” a user into a non-existing world. Due to the possibilities of immersing a user into a digital world through stimulation of

multiple senses, a VE can be very expressive and allow users to participate actively within stories told.
- *New or extended possibilities:* VEs truly can enable tasks that are hardly possible at desktop systems. These tasks largely involve the "usage" of immersion, for training purposes or psychiatric experiments.

A main factor that has driven the popularity of VEs is the availability of both affordable and mature technology. Especially through developments in the games industry, real-time graphics can be provided by lower-cost workstations, and low-cost office projection technology, cutting the costs of VR systems.

1.2 Basic 3DUI tasks and techniques

In order to perform a task within a VE, one needs an interaction technique that translates the user's intentions captured by an input device into system actions.
Accordingly, the action should result in an output from the computer. Even though a user can accomplish a huge number of tasks within a VE, some main categories can be defined that overlap with most tasks. These tasks are selection and manipulation, navigation, system control and symbolic input. In the next sections, a closer look will be taken on these tasks. For detailed descriptions, please refer to (Bowman, Kruijff et al. '05).

Manipulation
Manipulation is probably the most fundamental task in a VE. It allows a user to adapt the content of an environment, thereby clearly getting away from just being a "passive observer." When manipulation can not be performed, many application-specific tasks cannot be executed. Most of the times, manipulation is understood as *spatial rigid body manipulation*, which means that object shapes are preserved during execution. Hence, manipulation should not be mistaken with "modeling" tasks, in which spatial bodies can be deformed. As a result, three canonical tasks can be defined. These tasks are selection, positioning, and rotation. The performance of these tasks is highly related to the used input device. A device always has a number of *control dimensions*, generally ranging between 1 (a button) up 6 degrees of freedom (DOF). In addition, using a device for manipulation tasks is bound to how many DOFs a device can perform at once, which is known as *control integration* (Jacob, Sibert et al. '94).
Ergonomically, multiple factors play a role: to which extent can we rotate or position before we need to re-grasp the device (known as clutching), what kind of *grip* do we have on the device, how rough or precise a task can be performed?
A handful of basic techniques exist that can roughly be subdivided in how we select the objects in the VE (with its respective feedback channel), and from which perspective we can manipulate objects (first or third-person view). The perspective issue is probably the most dealt with problem – when objects lay outside the reach of a user, it causes a whole lot of difficulties. Generally, objects out of reach are manipulated via some kind of artificial extension of the arm. A more detailed discussion on the classifying manipulation techniques can be found in (Poupyrev, Weghorst et al. '98b).

Navigation

Navigation refers to the task of moving through a virtual environment and consists of two parts: travel and wayfinding.

Travel is the motor component of navigation, built up from low-level actions to control the position and orientation of a viewpoint. In the real world, this mostly involves physical movement like walking, but within the VE, this is not necessarily the case. Travel takes three forms: exploration, search or maneuver. Search and exploration refer to movement with or without a specific goal, whereas maneuver is the performance of more delicate viewpoint modifications. These tasks consist of three components: direction, velocity and conditions of input (Bowman, Koller et al. '97). In order to perform travel tasks, one can choose either a physical way of locomotion, or a virtual one, which may include only little physical movement. Physical locomotion techniques include real walking, using a treadmill, and vehicles like a bike, whereas virtual techniques make use of pointing, gazing, or specialized steering devices. Traveling using one of these (or other) techniques may include additional aids like zooming or advanced rotation methods to make movement easier.

Wayfinding is the cognitive counterpart of travel. In order to move through a VE, one needs to “plan” where to go: wayfinding is the cognitive process of defining a path through an environment, using or acquiring spatial knowledge, aided by both natural and artificial cues. This activity is mostly an unconscious activity, but come into the forefront when we get lost. Wayfinding can be seen as main “purpose” for an application, either when it is used to transfer spatial knowledge to the real world, or when complex environments are being explored. Different kinds of spatial knowledge exist: landmark, procedural and survey knowledge. When a user moves through a VE, all these kinds of knowledge are possibly built up. A discussion on the importance of these kinds of knowledge can be found in (Thorndyke and Hayes-Roth '82). In order to support a user performing wayfinding actions in a VE both user-centered and environment-centered aids exist. User-centered aids include the provision of a wide field-of-view in a display device, the supply of real motion cues, or the support for search strategies. On the other hand, environment-centered aids include natural cues like atmospheric colors or fog for depth estimation. Other possible techniques are the usage of architectural design methods or *legibility* techniques (Lynch '60) to structure an environment or making available artificial cues like a map or compass.

System control

System control methods are techniques to send commands to an application, to change a mode, or to modify a parameter. As such, they are inherently equivalent to using a widget in desktop environment. Nevertheless, it is not always possible to transfer WIMP (Windows, Icons Menus, and Pointers, (Preece '02)) metaphors from a 2D desktop environment into a 3DUI. The task characteristics in a VE are often too different; hence, using a simple menu technique may lead to extremely low performance. Nonetheless, the usage of such menu techniques is the main method of performing system control actions, simply because up till now, not much attention is given to system control techniques. A limited number of new system control techniques exist that need to be carefully tested. These techniques include graphical menus, voice commands, gestural commands, and tools. Obviously, a strong dependency exists between technique and used input device.

Symbolic input
Symbolic input is the task of communicating symbolic information like text or numbers to the system. This task can be seen as the “traditional” desktop task, normally bound to keyboard input. Within the rising complexity of VE applications, often there is a need for providing symbolic input that simply cannot be performed by using a keyboard. Hence, new techniques have been developed that can also be used within a VE. Mostly, these techniques or devices are derived from other application areas. Techniques include devices that are closely related to a normal keyboard, like miniature or low key-count keyboards, and soft keyboards that mimic the usage of a real keyboard in “thin air.” On the other hand, pen-based, gestural, or voice techniques exist that overcome the need for a keyboard, but are usually less effective or more complex to use.

1.3 3DUI Research directions and issues

1.3.1 Main research directions

Within the field of 3DUIs, several research directions can be identified. The biggest focus is probably still on VR technology – much attention is given to designing better and more advanced *I/O devices*. Developments include the creation of high definition visual displays, spatial audio setups, the construction of vibrotactile arrays, and the design of specialized input devices. Examples of these I/O devices can be found in the next chapter.

A second main issue is the design of *better or specialized 3D interaction techniques*. Investigations focus both on more advanced techniques for universal tasks, as identified in section 1.2, or at techniques for complex or composite tasks. Another field of interest is the development of 3DUIs for 2D devices, either to be used at a desktop environment or for *mixing 2D and 3D techniques* in a VE. Closely related to the development of interaction techniques are efforts on design approaches, including hybrid interaction techniques, multimodal interaction, and two-handed interaction.

Given less attention, but not less important, is the interest in 3DUI software tools or even complete 3DUI toolkits, and evaluation methods.

Reflecting the laid out main research directions, this dissertation mostly focuses on the development of new I/O methods for 3DUIs, as well as the creation of new interaction techniques. It should be stated that a large overlap exists with the main directions general HCI research. Fields like multimodal interfaces and ubiquitous computing are not specific fields of research in the 3DUI domain, but affect multiple topics in this dissertation.

1.3.2 Main goals

Several main research goals in the field of 3DUIs can be identified that are strongly coupled to the research directions identified in the previous section. Related to the creation of new I/O devices, much attention is given to make *technology better*. This includes creating a higher resolution or lighter display, and making audio available for difficult display setups. Better technology is often a way of making *VEs more realistic*. For the purpose of accuracy (engineering applications), impact (training or games), or esthetics (art), both hardware and software technology are used to create more life-like and detailed computer input and output.

However, better technology or realistic VEs do not always result in effective applications. Better technology can result in better usability, for example when an input device with better ergonomics is developed. Nonetheless, most of the time, much effort also needs to be spent on *making 3D user interaction "easier and better,"* by focusing on the actual interaction techniques being used. Necessarily, this also includes work on the dependencies between task and technique. With the rise of more complex applications, more specialized tasks cannot always be matched well by just using the techniques developed for the universal tasks identified in section 1.2. Thus, specialized techniques need to be developed.

Finally, a good chunk of work is performed on *understanding the impact of VEs*. The most dominant field of work is probably *presence,* related to understanding why people "feel" like being in another world (Slater, Usoh et al. '94).

1.4 The relation between 3DUIs, human I/O and unconventional interfaces

3DUI developments and VEs have partially been driven with the aim to make virtual worlds "more realistic." Realism often directly relates to the simulation of multiple human senses, in order to represent digital content as close to "the real world" as possible. As such, multimodal (or multi-sensory) output and to a lesser extent also input has been driven forward. Visual and auditory output has received most interest, followed by haptic output, whereas latest developments include the stimulation of the smell and taste senses.

Even with the interest in human-centered or user-centered design approaches (Raskin '00), these developments have been driven by technology and mainly observed from the computer-human communication viewpoint. Interaction is predominantly regarded as "computer input and output".

With (Beckhaus and Kruijff '04), the viewpoint has been turned around in order to truly focus on the human itself. As a result, not the computer but the human is regarded as an input and output structure. Even though a technological influence is always observable, this approach is highly focused on the human possibilities of providing output to a computer or to receive input. As a result, observing the communication flow between human and machine has been reversed:

machine input = human output
machine output = human input

But why do so? At a first look, this approach may look rather peculiar, but when taking a closer look, some particular advantages can be identified … even when it needs a little twist in the mind to start with. Basically, we want to regard the user as "human processing unit," to see what is actually possible to perform with the human body, without directly restricting developments (or ideas) by technology. We want to come up with new tasks and application areas, basing the exploration on a human-oriented (requirement) analysis. We perform these analyses in order to develop technological solutions, without being bound to current technological solutions in the first way. At the same time, it can be seen how current technology (or technological advancements) can be connected to the human body to fully unlock its potential, for example by re-focusing its originally intended functionality. For example, think about using a brain activity scanning device for interaction purposes. Don't be fooled if this may sound like a system-oriented approach: The potential of the human body will still be the main focus, to which technology will be adapted and not vice-versa. Technology oriented developments are currently still dominating, being used by the general user in daily life and can greatly restrict or confuse a user in certain task domains.

In regard to human potential, this dissertation should be understood as having the aim to present interfaces that solely make use of the potential of the human body to its furthest extents under all circumstances. The unlocking of human potential should be kept in line with the effort to be spent on a task. Physically or cognitively overloading a user would mostly lead to a bad interface: the human is not a machine.

Exploring the possibilities of human output and input possibilities fits perfectly well in the development of 3DUIs. As stated before, the development of 3DUIs highly involves the "usage" of multiple human senses and control channels (from now on called the *human I/O channels*) in order to observe and interact with VEs. Thereby, the evaluation of actual potential of the human I/O channels can greatly stimulate new developments in the field of 3DUIs or bring current developments into new light.

Driven by the human I/O framework, it can be concluded that unconventional interfaces are viewed upon from two directions:

- *Human oriented*: What potential do the human I/O channels have and how can they be applied in a 3DUI, or can we even drive interfaces beyond the potential of the human body?
- *Device oriented:* Are there devices that can enable the full potential of the human body, for example by using devices from non-computing application areas, adapting the functionality of a device for other kinds of usage, or using emerging technology?

As a result, it will be questioned how the possibilities of the human I/O system be *applied*, and how new technologies may drive new 3DUI developments or increase the possibilities of current efforts. Also, some new application areas are illuminated that are not directly driven by human or technological potential.

1.5 Summary

In this chapter, the basic background of 3DUIs has been discussed, presenting the terminology and basic issues.

Interfaces for spatial, computer generated environments. 3DUIs are being used in a range of different spatial environments that are characterized by the amount of the real world being replaced (blocked out) by digital content. This range is known under the name of the Mixed Reality continuum.

The field of 3DUIs is highly interdisciplinary. 3DUI research integrates different disciplines, including computer graphics, human-computer interfaces, psychology, and design.

Interfaces for different needs. Interfaces can aim at different purposes, including task improvement, better understanding of data, fun, and new or extended application possibilities.

Several top-level 3DUI techniques can be identified. Most of the currently available techniques can be categorized in a small group of tasks, which include selection, manipulation, navigation (travel and wayfinding), system control, and symbolic input.

Most research in the field of Mixed Reality is still focused on better / more advanced I/O techniques. Research, predominantly, still focuses on how to make technology better, but other themes, such as making spatial environments more realistic, the design and development of 3DUIs, or understanding the impact of VEs, are gaining more interest.

Exploring the human I/O to develop (multisensory) 3DUI techniques. This dissertation takes a predominantly human-centered approach, identifying the potential of the human body to develop new or more advanced 3DUIs. For analytical purposes, the human is seen as a "human processing unit," leading to the notion of the human I/O system.

Summarizing and reflecting, this dissertation addresses several main 3DUI research goals and directions identified in section 1.3.

- *More advanced I/O devices*: in this dissertation, many new and/or technologically advanced devices are presented, both from other researchers (chapter 2), and as a result of my own work (chapter 4). Own techniques (devices) specifically focus on increased ergonomics, the usage of sensory substitution methods, or the combination of 2D and spatial techniques (see next point).
- *Mixing 2D and 3D techniques*: three case studies focus on the combination of 2D and spatial techniques in so called hybrid interaction techniques (chapter 3 and 4) to advance interaction in tasks that are not solely of a spatial nature.
- *Making VEs more realistic*: several techniques are presented that focus on the combination of multiple sensory or control systems in order to create more "vivid" interactive environments (chapter 4).
- *Making 3D user interaction "easier and better"*: this dissertation specifically focuses on making interaction better, by looking at specific factors, such as the advantages, disadvantages and problems of multisensory processing, flow of action in complex applications, and the development of advanced feedback mechanisms. This includes the work on *specialized interaction techniques*, such as those focusing on the exploration of textures (chapter 4). Guidelines and /or explanations are included that make it easier for other researcher to replicate results (chapter 3).

As a next step, in the following chapter a closer look will be taken at "human processing unit," by taking a detailed look at the different human I/O channels.

CHAPTER 2
The human I/O system

In this chapter, an overview of the human I/O system will be given. The different input and output channels of a human being will be described at a basic psycho-physiological level, focusing on the specific potential that could be applied in human computer interfaces and, more specifically, in 3DUIs. Therefore, a specific focus will be put on how the potential of the human can be accessed or enabled via hardware interfaces, and how they can be applied at a practical level. As a result, the human input and output channels are described at three levels: perception, hardware interfaces and application. Furthermore, some top-level factors that deal with behavior are handled that affect both human input and output. This chapter does not provide a complete overview of all possible hardware devices for a 3DUI, since this would be outside the scope of this dissertation. To a large extent, the content of this chapter is based on (Beckhaus and Kruijff '04), adapted and specifically focusing on the field of 3DUIs.

2.1 Introduction to the human I/O channels

In order to create a basic understanding of what is meant by the human I/O system, a principal description of the main issues will be given in this section. These issues will be the focus of the following sections, when describing the different I/O channels, and in the next chapter, when the focus will be on the more practical aspects of designing unconventional 3DUIs.

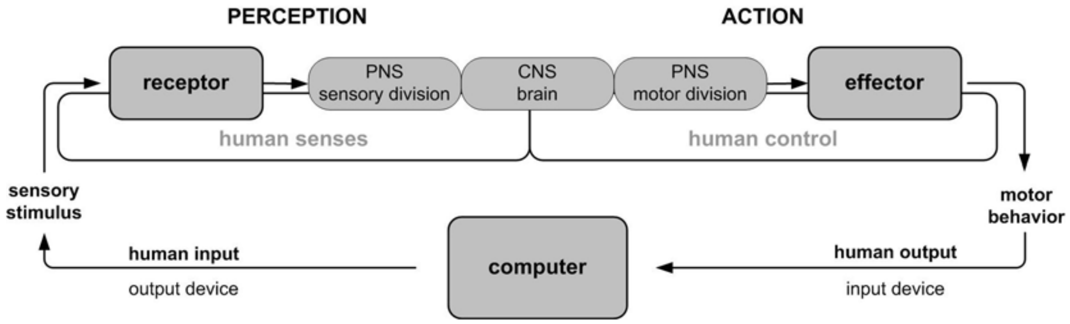


Figure 2.1: *Information processing in the human I/O system (the human as processing unit).*

The human I/O system and its relation to human computer interfaces can be best described by providing an information processing diagram as shown in Figure 2.1. The information processing loop consists of a perception and an action sector. In a normal ("traditional") human computer interface, a user will receive a stimulus from an output device that stimulates a receptor. The human sensory system, as is well known, generally is regarded as having five modalities: vision, audition, olfaction (smell), gestation (taste), and somasthesis (haptics/touch). The vestibular system, which contributes to the sense of balance, is sometimes also seen as modality.

When the body receives a stimulus, it extracts four kinds of information: the sensory modality or submodalities, the intensity of the stimulus, the duration of the stimulus, and its location. Each sensory modality has its specific receptors, which converts physical stimulus energy into electrochemical energy. As such, a receptor “fires” impulses that are carried to the central nervous system (CNS) via the sensory division of the peripheral nervous system (PNS). The central nervous system (including the brain) handles these signals with its specific areas (sensory receiving areas) and sends impulses out via the motor division of the PNS. These impulses can result in so called *voluntary actions* that involve the stimulation of the muscles, for example to control an input device. The PNS also controls *involuntary actions,* like the control of the heart beat. As such, they do not fall under motor output as described in Figure 2.1 (Kandel, Schwartz et al. '00).

One of the particular focuses of the human I/O system is the close coupling of human and hardware interface. Often nicknamed “bioware,“ some medical directions provide new perspectives on how closely human potential, or human information processing, can be influenced by digital devices. Throughout this chapter, the importance of these devices for 3DUIs will be illuminated to certain extent. As with the general descriptions of the specific sensory and motor (or control) channels of the human body, the information on psycho-physiological issues will be kept within boundaries. For a complete overview and detailed discussion, please refer to (Shepherd '94; Kandel, Schwartz et al. '00; Goldstein '02) or to the specific references given within the text.

The sections describing the different human systems are split up in three subsections. The first part introduces the psycho-physiological background of the system, followed by a (focused) overview of different hardware interfaces that can be connected to the human system. Finally, in the application part it is explained how the coupling of the human system and hardware interface may lead to new or adapted interfaces or sensations or how human potential could be used in different ways, thereby demanding new hardware interfaces and techniques.

2.2 Human input

Human input refers to the processing of a stimulus that can be provided by an external source like an output device. A stimulus is handled through a sensory pathway, as can be seen in Figure 2.2. When a stimulus arrives at a receptor, it is encoded by the receptor neurons into sensory information that travels to the brain’s specific receiving area. For most modalities two or more parallel pathways convey sensory information. Each neuron of a sensory pathway is composed of an input side (the cell body and dendritic trees) and an output side (the axon with its branches). Transmission between neurons happens between the synapses, and the axonal endings and dendrites of another neuron (Kandel, Schwartz et al. '00).

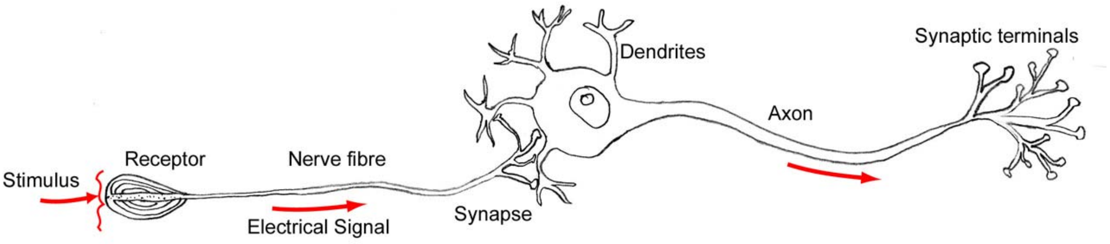


Figure 2.2: *Sensory pathway, adapted from (Goldstein '02). Courtesy of S. Beckhaus*

Receptor cells are specialized and are sensitive to different stimuli. Hence, different modalities are associated with different receptors (Figure 2.3).

Modality	Stimulus	Type of receptor	Receptor
Vision	Light	Photoreceptor	Rods, cones
Audition	Air-pressure waves	Mechanoreceptor	Hair cells (cochlear)
Olfaction	Chemical	Chemoreceptor	Olfactory sensory neurons
Taste	Chemical	Chemoreceptor	Taste buds
Somatic	Mechanical, noxious (chemical), thermal	Mechanoreceptor, Nocireceptor, Chemoreceptor, Thermoreceptor	Dorsal root ganglion neurons
Vestibular	Head motion	Mechanoreceptor	Hair cells (semicircular cells)

Figure 2.3: *Modalities and associated receptors. Adapted from (Kandel, Schwartz et al. '00)*

But, why is it important to understand how the sensory pathways work? The answer is quite simple: unconventional computer interfaces may not work the “traditional way”. Within a conventional human computer interface, only receptors are stimulated to convey information towards the user. However, with some unconventional user interfaces, this is changing. In order to create a sensation that results in a perception one does not necessarily need to stimulate a receptor via an external source. Since the sensory information consists of electrochemical energy, one can also directly stimulate the nerve system or even the brain via electrical stimuli. In the following sections, multiple kinds of interfaces will be shown that bypass the normal sensory pathway, by not stimulating the receptors at all. Similar effects can be seen on the motor side of human information processing. These kinds of interfaces have a large impact on how we can observe interaction processes – a discussion on this issues follows in the section 2.2.6.

2.2.1 Vision

Perception
The visual system is the most important sensory system of the human body and consists of three parts: the eye, the lateral geniculate nucleus and the thalamus, the sensory processing area in the brain. The visual system senses electromagnetic energy with properties of waves and particles, called photons.

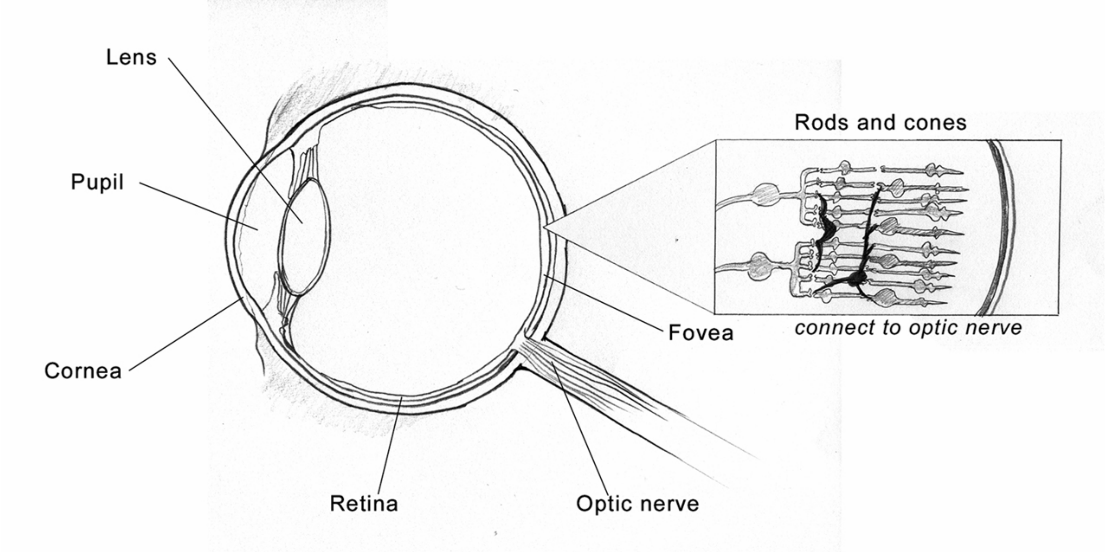


Figure 2.4: *The visual sensory system: eye and optic nerve.*

The human eye (Figure 2.4) senses these photons as light. This light is radiation with wavelengths between 370 and 730 nanometers. The different wavelengths of this radiation are interpreted as different colors. Other radiation, like infrared or ultraviolet light, cannot be observed by the human eye. When light passes the cornea and the lens in the eye, it stimulates the visual receptors, the cones and rods. These photoreceptors trigger electric signals, via the neurons into the visual area of the brain. The visual system generates stereo vision with depth perception up to about 6 meters and can focus on object as close as 20cm and up to infinite depth. The eyes observe within a range of 210 degrees (the human field of view). Observed objects are perceived with different form and size, color, depth, brightness and motion. The detail the human eye can perceive is limited by the number of sensory cells, which number about 120 million (Salvendy '97).

How the human eye perceives the outer world has been the source of a large number of (contradicting) theories. An overview of these theories can be found in (Goldstein '02). Visual perception affects a large number of HCI related factors, including the focusing of attention, the ability to group elements, and the registering of feedback. A discussion of these factors can be found in most HCI books, including (Shneiderman '98; Preece '02).

Hardware interfaces

There are a huge number of traditional visual output systems available, ranging from TV or monitor-based up to meeting room projection wall systems. Within the field of augmented and virtual reality, most of the systems used make use of some kind of head-coupled device or wall display systems consisting of one or multiple walls. An overview of most of these systems can be found in chapter 3 of (Bowman, Kruijff et al. '05). A selection of unconventional display systems can be identified and will be provided hereafter.

General display advancements

The first group is roughly characterized as displays that provide and visual information to the human eye, via a body-external source such as a screen, therefore, being advancements or derivates from general display devices. Some display systems adapt general display principles, for example by projecting on non-solid screens, like the Fogscreen (Fogscreen '05) which projects on fog curtains or different forms of screens. Other approaches focus on creating high-resolution display systems that get close to the actual "resolution" of the human eye. Finally, there are several directions focusing on the creation of spatial imaging that are independent of any glasses (such as passive or active stereo glasses), including holographic (Figure 2.5, (Actuality '06)) and autostereoscopic displays (Perlin '00). Especially holographic displays are still rather exotic and can, therefore, certainly be regarded as unconventional visual display technology.

Figure 2.5: *Holographic display.*
Courtesy of Actuality

Non-invasive retinal display

The second group of visual output devices directly focuses on stimulating the human visual system. An example of a non-invasive interface is the Virtual Retinal Display (VRD), developed at HITLab (Tidwell '95) and now sold by Microvision (Microvision). The VRD makes use of photon sources to generate light beams (laser) in order to create rasterized images on the user's retina. With this method, every color needs a separate source. The light beams are intensity-modulated, hence allowing both fully-immersive and see-through display modes. The display can potentially create full field-of-view images that are both bright and high resolution, and it is highly wearable.

Invasive displays (visual prostheses)

Next to non-invasive methods, there are a couple of developments that make use of invasive techniques in order to create purely artificial vision. These interfaces operate within the user’s body and stimulate the retina, optic nerve, or the visual cortex. Next to these purely digital approaches, there are some developments combining neural cells and photoelectric devices, so called biohybrid implants. These devices are predominantly developed for disabled people, but may find applicability in future display systems for non-impaired people.

The seminal work on visual prostheses started in the late 60ies with an 80 electrode system implanted into the visual cortex of a volunteer, which was fed by transcutaneous signals (Brindley and Lewin '68). These kinds of prostheses elicit small, limited subjective sensations of light (phosphenes). The aim is to produce patterns of phosphenes that resemble low-resolution mental pictures, allowing information from a camera to be transferred to the visual cortex. The principle of electrically stimulating neural tissue has been extended over the last decades and has resulted in minimal vision systems. Probably the most advanced system has been demonstrated by Dobelle (Dobelle '00). Dobelle makes use of a 64 microelectrode array that is fed by a belt-mounted signal processor receiving images from a digital camera. The system has been successfully “installed” in a couple of people that can see shades of gray that can somehow be interpreted by the brain.

Next to the electrical stimulation, some researchers experiment with magnetic stimulation (transcranial magnetic stimulation, also known as TMS). Using a coiled device, magnetic field impulses are generated that stimulate the visual cortex. The method has the advantage of being non-invasive, but is still in its infancy (Gothe, Brandt et al. '02).

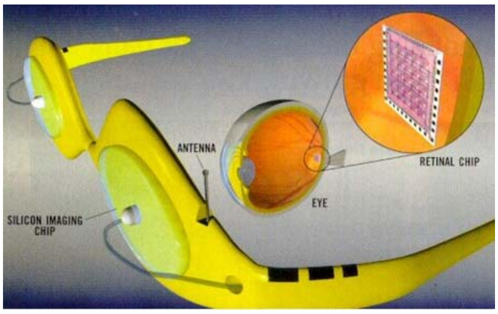


Figure 2.6: *Retina prosthesis with retinal chip.*
Development at University of California at Santa Cruz.
Courtesy of W. Liu

Stimulation of the optic nerve with electrodes has been probed within a limited amount of experiments. Veraart et al (Veraart, Raftopoulos et al. '98) demonstrated rudimentary phospene mapping through so called “spiral cuff” electrodes placed close to the optic nerve. Results of these experiments were limited, but promise to go into the same direction as the visual cortex stimulation. A particular advantage is that the implant can be easily inserted medically.

Developments like those performed in the Boston Retinal Implant Project (BRIP '04) or Retinal Prosthesis Project (RPP '04) focus on creating implants that either replace the photoreceptors in the eye (sub-retinal approach) or directly communicate with the ganglion cells (epi-retinal approach). These interfaces depend on photodiodes that are used as photoreceptor cells (sub-retinal approach) or electrodes (epi-retinal approach) to create cortical potentials. A first successful usage of a retinal implant, such as illustrated in Figure 2.6, was reported in (Palanker '06).

Vision substitution
Finally, experiments have started that use other channels than the human eye to convey visual information. One example is the tongue-stimulating system that makes use of electric pulses coming from an array of electrodes (144) to trigger touch receptors in the tongue. The pulses are actually patterns from images captured with a camera. As such, the tongue sensations through training result in some sort of perception of shapes in space (Kaczmarek, Weber et al. '91). Currently, the device developments foremost seem to focus on providing not directly visual information, but to create vestibular sensations (see section 2.2.5) (Wicab '05).

Application
It is rather difficult to do something really unconventional with visual interfaces. Both display systems and techniques that make use of the human visual system are heavily explored. Nonetheless, some directions can be given that are still open within the field of 3DUI or new fields that may result in the development of 3DUIs. These developments can be roughly divided in three categories, which are afterwards described in more detail:

- *Enhancing the visual system*: providing techniques or technologies that provide the user with the means to do things that are not “supported” by the human visual system itself
- *Replacing the visual system*: currently only focused on providing a means to make blind people see, it offers some provocative possibilities for future interfaces
- *Higher level sensations*: visuals may affect human subjective sensations, like emotion or tricking out the human with “false beliefs” (like the feeling of presence, in which a user is tricked into believing she is in another, digital world)

Taking a look at this from a psycho-physiological perspective, many enhancements can be foreseen. Many of these enhancements focus on the adaptation of the image normally perceived of the world. Many of these methods have already been probed in the field of visualization techniques, both in 2D and in immersive environments. Nonetheless, several directions can be stated that move 3DUIs towards unconventionalism. Most of these directions would probably involve the blending in of visualization techniques with the sensation of the real world through a mixed reality medium.
First of all, objects or information that lie *outside the visual spectrum* (370-730 Nm) can be visualized, for example by using a night vision camera capture. Adapting the visual spectrum (color, brightness, etc.) by adapting the image of the outside world is another field that might be explored further. Real-life tasks with different light conditions might profit from color (spectrum) adaptations – an example is an outdoor AR assembly scenario in a dark environment, in which light-intensified imaging is blend in.

Next to the adaptation of the color spectrum, the *modification of viewpoint* or viewpoint-independent viewing offers interesting possibilities. Limited by a single, fixed viewpoint, viewing methods can be developed that go beyond simple multi-viewpoint representations as provided by 3D modeling tools. These methods can theoretically be driven to extremes, like the provision of a different viewpoint on every eye. The usage of different viewpoints has been probed in airplane (military) technology, by using head-up displays (HUD), but within these scenarios, information is normally not conflicting. The information provided to the HUD normally fits to the visual image received by the non-covered eye. The combination of two different viewpoints can be both interesting as well as leading to a cognitive overload – the possible problems of combining multiple viewpoints is a well-known phenomena in wayfinding scenarios (Darken and Cevik '99a).

Related to multi-viewpoint display is the *blocking and de-blocking of objects*, enabling users to see through objects or making objects invisible. A possible method to achieve this is the real-time merging of footage from a normal camera, combined with a camera that can detect different heat ranges. Another possibility is, as is often done in augmented reality applications, to mix in digital content in video footage from a real scene.

Support for *extremely wide field of view* by encompassing the complete view of the user is still largely unsupported, though may yield considerable advantages for 3DUIs. One measured effect is the support of feedback via peripheral vision, like motion cues in wayfinding applications. In addition, new ways of providing feedback might be found that do not block the central vision, by displaying visual cues in the user’s peripheral vision area (see chapter 7 in (Bowman, Kruijff et al. '05)).

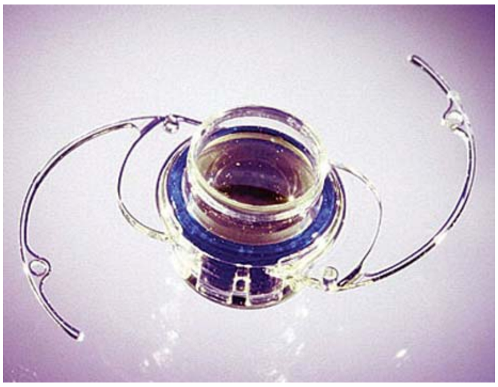


Figure 2.7: *Implantable miniature telescope.*
Courtesy of VisionCare Ophthalmic Technologies, Inc.

Experimentation with *different lenses* can provide interesting possibilities for 3DUIs, surpassing earlier experiments like the 2D Magic Lenses ((Bier, Stone et al. '93)) or its 3D equivalent by Viega et al. (Viega, Conway et al. '96). Similar to the two different viewpoints from different angles, two viewpoints with different lenses (like normal and close-up) can be rendered, for example to aid in specific manipulation actions. The Implantable Miniature Telescope by VisionCare (Figure 2.7) gives some idea of the possibilities of working with different lenses. Implanted in the eye, it provides magnified central vision for people with viewing disabilities. A similar enhancement could be provided by digital means. Even though a digital connection to such a lens construction would be difficult, similar functionality could be useful in a 3DUI, for example, to digitally enlarge remote or extremely small objects in outdoor Augmented Reality applications.

Finally, just like the miniature telescope, future *invasive display technologies* (implants) can possibly provide a new breed of 3DUIs. Irrelevant of the current display quality (the devices currently just provide some sensation of blobs of light), invasive devices could completely replace any normal real-world vision, thus delivering a completely artificial and, therefore, possibly fully immersive sensation. It can be expected that such interfaces are highly wearable and are therefore well suited for outdoor scenarios as currently covered by AR applications. Nonetheless, the usage of invasive technology for the visual channel raises a large number of questions of ethical and medical nature. The replacement of the "real" sensation of the world may bear considerable danger when used for other purposes than in assisted technology (technology for disabled people). A discussion on some ethical factors can be found in section 3.5.

2.2.2 Audition

Perception

The auditory sensory system (Figure 2.8) is the second most important human input channel. The auditory system consists of the outer, middle and inner ear. The outer ear collects the sound and directs it to the eardrum, which vibrates, mimicking the sound waves it receives. The eardrum produces a reaction in the three bones of the middle ear (Malleus, Incus, Stapes), which causes the fluid in the inner ear to move. This affects the hair cells, the receptors, in the cochlea (inner ear). The hair cells bend back and forth, thereby sending electric signals to the acoustic/auditory nerve, which carries them to the brain.

Sound waves are not only sensed by the ear. Organs like the stomach and the lungs also sense sound waves and vibrate at specific resonances, especially in the lower frequencies. As such, they can also sense audio and deliver sensations like "feeling" low frequencies.

The ear is able to process frequencies between approximately 20 and 20.000 Hz (the higher frequencies decreasing with age), and detects loudness, pitch, and timbre. Sounds that are not audible to the ear (infrasonic and ultrasonic sounds) may not be detected by the human ear, but may have other effects on the human body, including nausea. Especially when sound comes from the frontal hemisphere, humans are good in localizing sound sources through the intensity difference between the ears. Sound is rapidly detected, can have an alerting functionality (which can be powerful as feedback mechanism), and we can focus on specific sounds, blending out others (Goldstein '02).

Sound can have an affective function on the humans' emotional state. The wrong usage of infrasound (low frequency) can provide humans with uncomfortable feelings, up to a noxious state. On the other hand, music like Mozart's can have a relaxing effect, as numerous experiments have shown, or might even support spatial reasoning (for a further discussion, please refer to the application subsection).

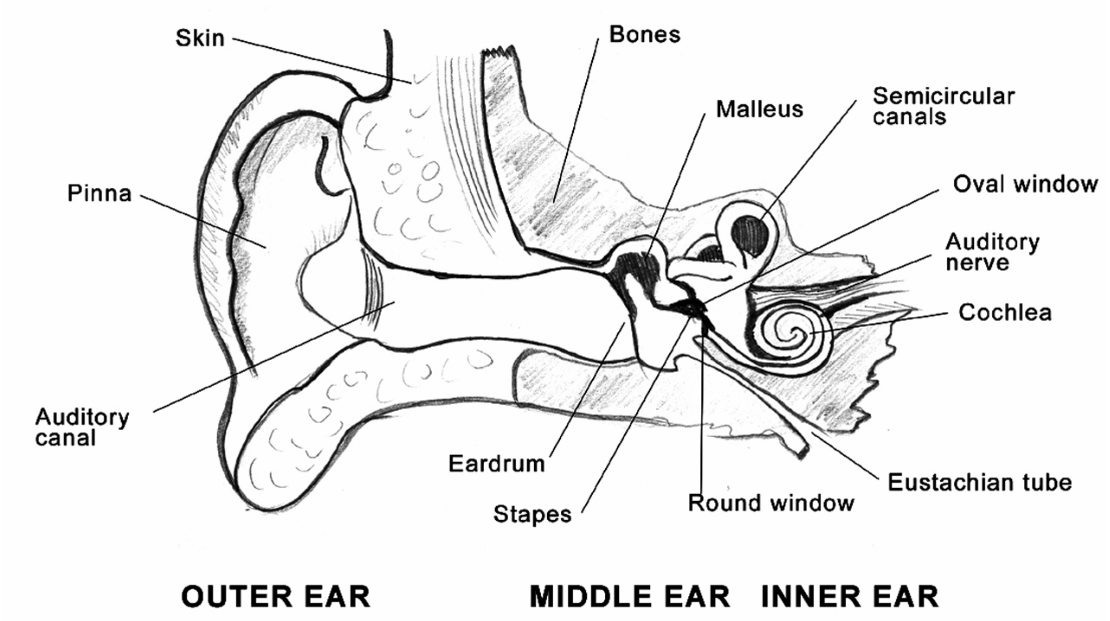


Figure 2.8: *The auditory system.*

Hardware interfaces

General display devices
Similar to visual displays, a huge number of audio displays exist. The major group consists of loudspeakers in different forms, like floor speakers or headphones, or new developments like large(r) arrays of loudspeakers as used in wave field synthesis (Berkhout '88). Most of these developments, such as Yamaha’s commercial device shown in Figure 2.9, make use of some kind of vibrating surface to produce a sound wave and can be unconventional in the huge number of speakers used.

Figure 2.9: *Commercially wave field synthesis device for usage in smaller space systems. Courtesy of Yamaha*

Ultrasonic based sound generation
Recently, new developments have been made to produce sound waves by using alternative methods. One of these developments makes use of ultrasonic sound to produce localized sound that can only be heard at one single spot. A sound wave

normally consists of a small pressure wave, going up and down when traveling through the air. During this non-linear movement, the air itself causes the sound wave to change, thereby producing new sound frequencies. These sound frequencies can also be caused by ultrasonic sound waves that create these new frequencies in the air. Ultrasonic energy is highly directional and produces a kind of column in front of the emitter, therefore producing a directed sound volume, not a spread one like with a normal loudspeaker. Since ultrasonic sound cannot be perceived by the human ear, a user will only hear the sounds produced by the ultrasonic waves within the conical volume. Examples are of these devices are (ATC '04; HRL '04).

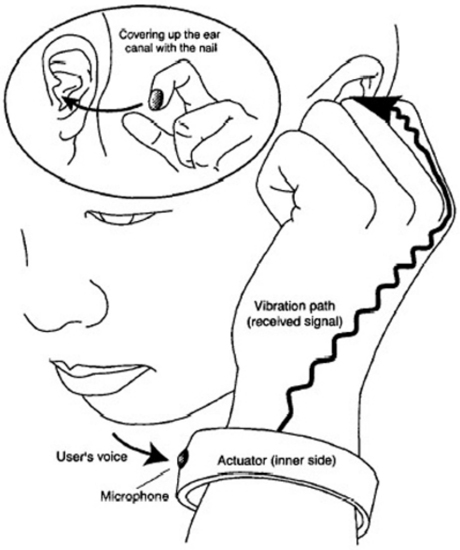


Figure 2.10: *Earbone vibration interface.*
Courtesy of M. Fukumoto

Vibration and conductance based sound generation
Using earbone vibration, Fukumoto et al’s Whisper (Figure 2.10, (Fukumoto and Tonomura '99)) produces sound by sticking a finger into the ear canal. Whisper is a wrist-worn input device, consisting of a small microphone and an actuator. The actuator translates the received voice signal into vibrations that travel via bone conduction in the arm and finger into the user’s ear. As with normal auditory production, the ear bones start to vibrate and let the inner fluids move the hair receptors, hence producing an auditory sensation. The audio quality is said to be of good quality, since the finger-in-ear method blocks out much noise. The device is now sold by NTT Docomo under the name Fingerwhisper (NTT '04).

The Whisper development, to some extend resembles research in cutaneous communication, in which actuators are attached to the skin to transmit auditory information. This kind of communication, though, is meant for people who are severely hearing impaired or deaf. Using this method, no real auditory sensations are produced – communication takes place over “buzzing” feelings. Hence, they are actually vibrotactile sensations.

Invasive sound displays

In addition to non-invasive audio generation, cochlear and middle-ear implants are available that produce audio sensations through invasive methods. In comparison to visual implants, cochlear implants are well-proven and used by large numbers of people. The cochlear implant (Figure 2.11, (ABC '06)) makes use of a combination of externally worn hardware and implanted components. Sound is picked up by a directional microphone and sent to a sound processor. The sound processor digitizes the sound and transfers it to a small transmitter antenna, which sends the coded signals via radio frequency to the internal implant. The implant converts the signals into electrical signals that are sent to the electrode array of around 20 electrodes that are placed in the user’s cochlea. These electrodes stimulate the nerve fibers, creating an auditory sensation.

In addition to cochlear implants, middle ear implants that amplify a recorded sound and transfer the sound via bone or skin conduction to the inner ear are available. Examples are the Vibrant Soundbridge (Gerbert-Hirt, Hauser et al. '04) or the TICA (Bowditch, Cox et al. '03), a fully implantable device which is does not seem to be produced anymore.

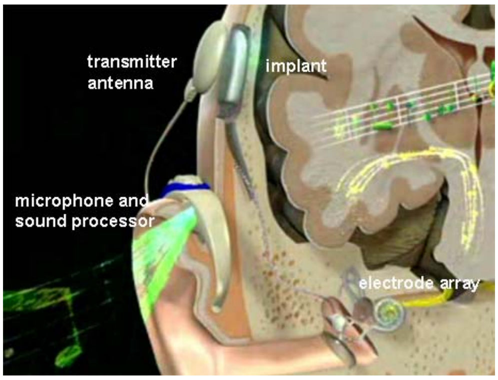


Figure 2.11: *Components of a cochlear implant.*
Courtesy of Advanced Bionics Corporation

Audio displays for non-sound generation

Finally, some hardware interfaces are available that produce sound waves, but no direct auditory sensation, since they are in the infrasonic or ultrasonic sound range. Infrasonic sounds are generally produced by subwoofers or large pipe-like constructions, whereas ultrasonic sound is produced by human speech or musical instruments. The devices can be used in VR setups, as will be handled in the next section and section 4.2.

Application

Within VEs, audio is mostly used for general audio generation purposes like playing some music or background sound or providing basic *feedback*, like a simple plong when an object is selected. Even so, most of the display setups have a rather minimal sound setup, which is often not in line with the high costs spent on the visual display. Hence, based on personal observations, it can be stated that audio is rather undervalued.

In order to use the full extent of the human auditory capabilities the setup of truly *spatial sound displays* can be considered. In this regard, the generally used 5.1 or 7.1 surround sound systems provide only a basic number of sound generation possibilities that can be upgraded by using true spatial sound setups. Such setups can make use of a large number of speakers, situated in a possibly regular pattern around the user's hotspot in a display system. Personal experiences, with around 20, up to 50-speaker systems, indicate a better immersion support: a user can be better pulled into the actions of an application (especially story-based applications).

One of the key advantages of using a speaker array is that sounds can potentially be connected to the appropriate objects (Caulkings, Corteel et al. '03). When objects have sounds connected to them, the sound can actually come from the same direction as the visual location of the object. Locatable audio cues can be useful as wayfinding aid, for example by coupling particular sounds to landmarks such as a railway station.

Related to object-sound connections, the creation of more advanced sound feedback mechanisms need to be considered. Such feedback mechanisms may vary widely, ranging from actual sound simulation (the realistic simulation of actual sounds that are changed through actions in a VE) up to the before mentioned usage of sound for *non-sound purposes*. As can be seen in section 2.2.3 on somatic perception, as well as in one of the case studies (section 4.2), sound can also be used to produce haptic sensations.

Using advanced spatial sound generation software, audio-only spatial environments can also be considered. For example, the LISTEN project (LISTEN '05) produced several *soundscapes* that would change through user interaction with the environment. In this case, interaction was based on a storytelling engine that presented sounds to the user based on location and path through a spatial environment.

Finally, an issue still open for discussion is the audio effect on user performance. In the nineties, a large number of tests were made on the so-called "*Mozart effect*". After an initial study by Rauscher and Shaw (Rauscher and Shaw '93), there were indications that Mozart's music would have positive effects on user performance, especially providing short term IQ bursts and improved spatial reasoning. A whole industry came up making profit by selling this still rather under-researched issue. Other researchers could not find any effect by playing Mozart (Newman, Rosenbach et al. '95), but all together, researchers seem to agree that with Mozart specific areas of the brain were quite stimulated, possibly leading to finer motor coordination and "other" higher thought processes. As such, sound could also have an effect on spatial performance in a VE, and effect performance of foremost manipulation and wayfinding techniques. A larger test environment would be needed to prove this statement, and even then, applicability should be well chosen.

2.2.3 Somatic and kinesthetic

Perception

The somatic and kinesthetic systems handle the sensations that relate to force and touch. The somatic system perceives cutaneous (skin) and subcutaneous (below skin) sensations, whereas the kinesthetic system senses mechanical sensations in the joints and muscles. These sensations are also generally known as haptic feedback and relate to the communication of information on geometry, roughness, slippage and temperature (touch), and weight and inertia (force).

The kinesthetic system senses the position and movement of limbs. The system has one submodality, which is regularly focused on in 3DUI research, namely proprioception (Mine, Brooks et al. '97a). This submodality refers to all sensory inputs from the musculoskeletal system.

Skin and muscle sensations are received by several receptors, namely thermoreceptors, nocireceptors, mechanoreceptors including proprioceptors and chemical receptors. As a result, humans are able to perceive: light touch and pressure, vibration and surface discrimination, pain, temperature and “harm” (noxious sense) via the somatic system. The kinesthetic system senses limb position, movement and force.

The quality of somatic and kinesthetic stimulation depends on both the resolution of the stimulus and certain ergonomic factors. Resolution refers to both the spatial and temporal resolution of a haptic stimulus. Spatial resolution deals with the diversity of areas in the human body that can be stimulated: some parts, like fingertips, are more sensitive than others. Temporal resolution basically refers to the refresh rate of the stimulus, which can be up to 1000Hz (Massie '93).

Somatic or kinesthetic stimulation can lead to pain. When going over a specific threshold, which is normally different for every human being, a stimulus can overly stimulate a receptor, normally triggering the noxious sense (Goldstein '02).

Hardware interfaces

A large number of devices exist that stimulate the somatic or kinesthetic senses. Over the last decade, research in haptic devices has become increasingly popular. In the nineties, haptic research was still dominated by rather complex devices that were mostly derived from research highly similar to robotics (like exoskeleton devices)(Biggs '02). Many of these developments still have an experimental look-and-feel due to the often strange robotic constructions used. Both miniaturization of mechanics and alternative approaches for creating haptic stimulations have resulted in smaller scale and more diverse devices. The devices can be best differentiated based on their actuators.

Figure 2.12: *A haptic device combining a body-coupled construction with a ground referenced device.*
Courtesy of Immersion Corporation

Body and ground-referenced devices

The majority of available devices focus on force feedback via body or ground-referenced hardware or a combination of both (Figure 2.12). Force reflecting devices mostly make use of pneumatics or hydraulics mounted in devices to put force on hand-coupled devices such as a joystick, pen, or steering wheel or more complex constructions using a multitude of moveable pins that adapt a flexible surface, like the Feelex (Iwata, Yano et al. '01). A second group exists that makes use of string-based (Ishii '94) approach, leading to both small, grounded devices, as well as larger devices that allow haptic feedback in larger projection systems such as a CAVE^{tm} (Buoguila, Ishii et al. '00). A special group of haptic devices consists of different kinds of motion platforms, like the Gaitmaster, a device simulating locomotion onto a staircase (Iwata '01). In addition, *tactile feedback* devices exist that provide cutaneous and subcutaneous feedback to a subject. Examples are inflatable air-bladders, small rapidly inflatable elements, or alloy shapes that are flexible and deform under heat influence (Bullinger, Bauer et al. '97). Strongly related are *vibrotactile feedback* devices that make use of small vibration elements like pager motors. An example is the SmartFinger, a nail-mounted vibrator that simulates textures while tracing the surface, using a photo detector (Figure 2.13, (Smartfinger '05)).

Recently, a separate group of force reflecting devices is getting more attention, even though its principles are known for a longer time. These devices are based on new “smart” materials or elements like piezopolymers, magnetism (Mignonneau and Sommerer '05), or shape-memory-alloys. A good overview of “smart” materials can be found in (Fletcher '96), whereas descriptions of most general haptic and tactile devices can be found in (Burdea '96).

Figure 2.13: *SmartFinger nail-mounted tactile device.*
Courtesy of JST & University of Tokyo

Wearable haptic devices

Moving away from traditional exoskeleton devices, *miniaturization* allows for smaller and more wearable and embeddable devices. Application is wide and can range from simple joysticks to light, up to possibly wearable, interfaces using a string-based

approach. The area of embedded devices seems a promising field for unconventional interfaces. Built as small blocks like the physical widget (Phidget) approach (Greenburg '01), or hidden completely, minimal haptics can empower scenarios like driving aids (haptic feedback during emergency situations to quickly warn a driver) or highly flexible toys for playful scenarios, similar to LEGO Mindstorms (LEGO '05).
Most of the actuators used in miniaturized haptic interfaces are based on vibrotactile feedback mechanisms, using small pager-like motors. A practical example of using pagers is the tactile vest by Lindeman (Lindeman, Sibert et al. '04) which is similar to NASA's Tactile Situation Awareness System (NASA '05). These systems are, among others, used for providing directional cues in navigation (wayfinding) tasks. A related tactile approach for wayfinding that should also be mentioned is the tactile map, a dynamic map that can be felt by blind people (Jacobson '96). The tactile map makes used of raised contours in a mechanical device and, as such, can also be interpreted visually.

Props and tangible interfaces

The field of passive haptic devices, generally known as *props*, brought forth numerous unconventional interfaces. Generally, props are real-life devices embedded in a VE, thus easy to use by a user. The classic interface is Hinckley's medical application using a plastic plate and a doll's head to perform cutting plane actions (Hinckley '94). Following Hinckley's idea, many interfaces have been developed. Probably the most influential direction based on the props idea is known as Tangible User Interfaces (TUI), started by Ishii and Ullmer (Ishii and Ullmer '97) under the name *tangible bits.* Most of the TUIs are used for manipulation actions and are quite powerful, since many actions in a VE can be represented by an object known from real-life. TUIs can be made extremely robust and, as such, frequently find their way in public space installations such as DisneyQuest's Pirates of the Caribbean, which uses plastic cannons to interact in a pirate game (DisneyQuest '06). TUIs show resemblance to the Phidget approach mentioned before due to their building-block like characteristics.

Electromuscular stimulation

Focusing on muscular contraction and release, electro-muscular stimulation provides an interesting way of providing haptic feedback. Using small electric currents, muscles contract through impulses provided via small pads on the skin. The muscular contraction can potentially be used to create biomechanical changes in the pose of a person. A more detailed investigation on the usage electromuscular stimulation can be found in the case study handled in section 4.3.

Non-haptic somatic feedback

There are several devices focusing on the stimulation of somatic sensations that do not necessarily result into force perception. *Heat elements* provide a way of communicating temperature to a user, either by body-coupled devices such as a heat pad, or by using a heat blower. Similarly, *air movement* can be created by using a fan, or some other kind of wind machine (Deligiannidis and Jacob '06). Finally, different levels of *moisture* can be generated by using an evaporative humidifier.

Non-body coupled haptic stimulation or substitution

Finally, side-effects of audio and wind can be used for providing haptic and pseudo-haptic (substituted) sensations. Human cavities, like the lungs, start vibrating when stimulated via certain high volume audio frequencies, resulting in a sensation that most people know from a (rock) concert. The usage of a small burst (or punches) of wind can be used to tickle or even slightly deform the human skin. For more information, please refer to section 4.2 for a detailed description of these feedback methods.

Application

Most actions in a virtual environment do not follow any of the physical rules that apply in the real world. Most manipulation actions do not provide any physical feedback to the user, and, generally, navigation is done in “floating mode”: the user is standing and pointing in a specific direction. In most systems, information is provided to the somatic and kinesthetic system by using sensory substitution. Sensory substitution is the usage of an alternative human input channel to communicate similar information (section 3.3). Neglecting physical feedback does not necessary lead to bad interaction, though adding haptic cues does provide advantages in specific circumstances. More complex manipulation tasks benefit from haptic cues by provision of more precise interaction feedback cues. It has been shown that real motion cues provide better wayfinding in virtual environments (Usoh '99). Haptics have also shown to be effective in medical scenarios (Steffin '98). One particular interesting experiment was the usage of plastic spiders as “tactile augmentation” in phobia treatment scenarios (Hoffman '98). Finally, the addition of haptics can increase the realism of an application and increase the sense of presence (Carlin, Hoffman et al. '97; Biggs '02).

But, what are the actual technical or human innovations to be expected in the haptic research area? In the following paragraphs, a closer look is taken on new directions or possibilities for haptic feedback. A detailed discussion on general usage of haptic devices and their effects on interaction can be found in (Burdea '96).

Figure 2.14: *Pos.t wear feedback mechanism.*
Courtesy of POSTECH

A field still open for further research is *full-body haptic feedback.* Theoretically, the human body can be encapsulated in an exoskeleton structure, but for the majority of scenarios, this is unacceptable. Currently being probed is the usage of arrays of vibrotactors, for example mounted in clothes like the before mentioned tactile vest. A vest-like construction focusing on truly full-body feedback is Pos.t wear (Figure 2.14, (Yang, Jang et al. '02)), based one multiple ring-like levels mounted with around 60 vibrotactors. Current full-body feedback devices are still characterized by their small spatial resolution, since they only provide rather rough feedback, both due to the limited amount of actuators and the bodily zones they stimulate (the different parts of the body have diverse sensing levels).

Another challenge is support for *group-based haptic feedback*. Even though passive feedback, using props, has been probed in multi-person setups (see section 4.2), active haptic feedback is still largely under-developed. Due to technical limitations (it is rather hard to hook up every single person with a haptic device) there is currently no real group-based haptic interaction outside those setups that make use of multiple desktop-based workspaces. A first attempt for rough group-based haptic can be found in section 4.2. By using audio shockwaves larger groups of people can be stimulated with a limited amount of haptic sensations.

Finally, one of the more dubious kinds of somatic feedback is the infliction of *pain*, caused by haptic feedback that passes a specific threshold. Inflicting a user with pain via physical methods does not, in most cases, make any sense at all, but it can be imagined that for medical experiments, it could be useful. Within the games area, the infliction of pain has been tried out at least once. The rather popular PainStation (Figure 2.15, (PainStation '05)) combines multiple devices, including a heat pad, electroshock, and a whip to create painful sensations in users. As such, it adds an extra level of game play to a user. The ethics of using such installations are dubious – still, people play voluntarily the PainStation.

Figure 2.15: *Users playing PainStation.*
Courtesy of www.painstation.de

2.2.4 Chemical

Perception

The human chemical system (Figure 2.16) handles both smell (olfaction) and taste (gustation) sensations. Both systems are functionally linked – taste sensations are more difficult to register by a human being when no olfactory information can get through. The reception of olfaction and gustation is handled by chemical receptors that are stimulated via chemical substances. An olfactory sensation is triggered by molecules in the air, which are released by substances in our environment. Inhaled through the nasal cavity, nose, and mouth, they stimulate different receptor cells. Humans have about 1000 different receptor cells, each sensing a different chemical bond. Many people have tried to classify the range of odors, but no successful or generally accepted categorization exists. Not all of the chemical receptors are activated in every human being – clear differences, for example between races or as a result of training can be noticed (Menashe, Man et al. '03). To discriminate a scent, multiple receptor cells are active. A particularity of these receptors is that they are directly connected to the brain. As such, the perception cycle is much faster than with other human input systems.

In addition to the sensation of scents, so called trigeminal sensations can be perceived. The trigeminal system connects to (noci) receptors at multiple places, including the lips, tongue, tooth, oral, and nasal cavities. Thus, humans can also sense burning sensations (like hot pepper) or temperature via the nose and other body parts, like itching sensations on the skin (Goldstein '02). It should be stated that these sensations are more closely related to tactile sensations and, therefore, cannot necessarily be categorized as purely chemical.

Practically, scent is foremost used for smelling food or for detecting danger (like fire). Due to the speed of perception, smell is especially well-purposed for danger detection. Both unconsciously and consciously, smells affect our emotions, for example when we smell another person, and a regularly connected to memories or experiences. Overall, humans tend to detect anywhere between 2 and 1000 smells, and quickly adapt to an individual smell in about less than a minute. Therefore, complex smells are easier to remember (Kaye '99).

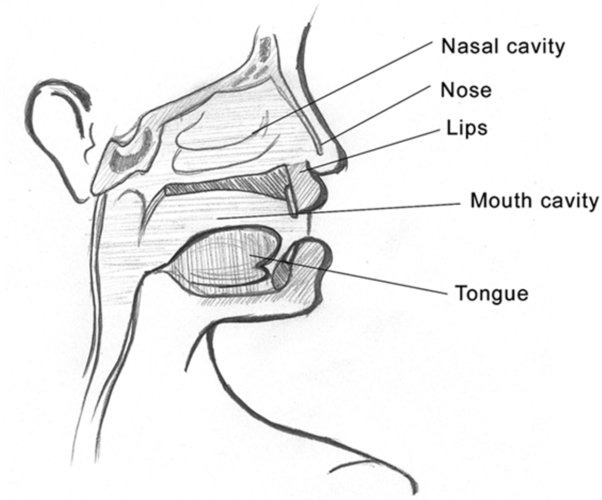


Figure 2.16: *The chemical system.*

The perception of taste is performed via around 10.000 taste buds in the tongue, mouth cavity and throat. The taste buds can adapt extremely fast to new tastes and are most active in the temperature range of 22-32 degrees Celsius. Just like with smells, there is no standard classification of tastes, even though the rough distinction between sweet, sour, salty, bitter, and unami (savory or meaty flavor, coming from a Japanese categorization) seems to be somewhat accepted (Goldstein '02).

Hardware interfaces

General system methods

Over the years, a number of different smell devices have been developed. Almost exclusively, these devices evaporate some kind of smell to arrive to a user via air. Such devices can make use of un-powered (like a membrane) or heat-induced methods to evaporate liquids, electrostatic evaporation of gels, or micro-encapsulation of scents that can be released mechanically or via heat. One of the basic problems of all these systems is to time the actual duration of scent, affected by the time to the clean the scent from a space after it has performed its duty. Another problem is to actually create an appropriate scent, as pure scent, or a mix (Hamnes '02) (Kaye '04).

Smell gun
A device that can well be used within VR setups is Yanagida's smell gun (Figure 2.17, (Yanagida, Kawato et al. '03)). Using a camera to track the user's eyes to define the location of the nose, it can direct smells towards a user using a small air cannon. The device can load multiple kinds of scents from small containers, enabling diverse smell generation. The amount of scent propelled to a user is limited and as such, duration of small is short, unless one keeps propelling small air shots towards the user. On the other hand, due to the limited amount of scent propelled, different smells can follow each other shortly after each other.

Figure 2.17: *Nose-tracked personal olfactory display.*
Courtesy of Y. Yanagida

Frequency pulses
In 2003 and 2004, Sony obtained several patents (US patents 6,729,337 and 6,536,440 (USPTO '05)) entitled "Method and system for generating sensory data onto the human neural cortex." The patent texts provide a basic explanation on using dual transducers to project sound patterns using low frequency pulses onto the neural cortex, to stimulate effects that would (according to external sources) include the generation of smells in TV sets. Due to the informal reports from different sides (web logs and news papers) it has to be seen how far this technology goes.

Tasting and drinking simulators
Hardware stimulating taste sensations are very rare. At SIGGRAPH 2003, Iwata et al. (Iwata, Moriya et al. '03) showed the food simulator, a biting-force haptic device that can generate both taste and smell sensations and appropriate audio effects (like the breaking of a cracker). Chemical sensations of taste are generated using a micro injector, whereas a vaporizer delivers smells.
Focusing on the "edible", Maynes-Aminzade introduced two so called *Edible User-Interfaces* named BeanCounter and TasteScreen (Maynes-Aminzade '05). BeanCounter consists of rods filled with jellybeans that can be opened using electronically controlled valves, thereby dropping jellybeans in a small bucket, to be eaten. Using this system, a

hardware memory monitoring, as well as a network monitoring application were built that would release jellybeans when a system would be optimized. TasteScreen makes use of flavor cartridges mounted on a LCD panel that can drip-wise release transparent fluid over the screen forming a liquid residue that can simply be licked by the tongue. Both interfaces seem to be not much more than a good joke, but demonstrate the possibilities of introducing taste in a user interface.

Finally, an interface foremost focusing on the experience of drinking is the *strawlike user interface* (SUI, Figure 2.18) developed at University of Electro-Communications in Tokyo (SUI '05). The interface combines sample data of actual pressure, vibration and sound produced from drinking from a straw and generates sensations accordingly using a valve and a speaker. Associating the drinking action with appropriate visual images, a sensation that is somehow related to the sense of taste is created, though the haptic component of drinking clearly dominates.

Figure 2.18: *Drinking interface using a straw simulation. Courtesy of INAMI laboratory*

Application

The application of scents and tastes is rather limited in VR installations. Scent seems to be predominantly used in public spaces, like malls, to create an interesting atmosphere via an ambient smell. Only a few installations, like the Gyeongju VR theatre (Park, Ko et al. '03), have actually used smell in VR productions. Most of the smell applications seem to come from the art area, such as Plett and Haque's Scents of Space installation (Haque '05).

The usage of smell in a VE can have an impact on the *sense of presence and immersion*. That is, environments become more vivid and life-like, and, based on personal observations, people tend to get pulled more into the actions of a VE. Projects like the Sensory Environment Evaluation project (SEE, (Morie, Iyer et al. '03)) also focus on the usage of smells to create more "emotional" environments, and first tests, like (Dinh, Walker et al. '99), seem to prove that the addition of smell indeed improves the sense of presence. Nonetheless, more evaluations need to be performed in order to prove the validity and application.

Some research indicates that olfactory stimulation can have an effect on *human learning and memory*. Washburn (Washburne and Jones '04) states that olfaction could increase information processing, reduce response times, reduce errors, increase recall, and enhance physical performance. Hence, olfaction can be coupled to tasks that have a higher memory load, like wayfinding. Quite related, scents potentially can help in *training scenarios*. Due to the strength of the chemical system to sense danger, training scenarios can be designed to specifically train users on dealing with smell in hazardous situations. As such, smell can also be used as wayfinding aid to *direct attention*. The previously mentioned test by Dinh et al. not only focused on presence effects, but also on memory issues. By connecting smells to specific objects in space, users would associate smells to specific locations and would, therefore, possibly find their way faster. Smell, like the coffee odor located near to a virtual coffee machine, thereby become a landmark, a technique especially usable for less experienced navigators (Thorndyke and Hayes-Roth '82).

The application of taste is both technically and functionally rather difficult. Like with smell, it can be imagined that taste can enhance the sense of presence, but since we also do not taste much more than food in daily life, application will probably be extremely scarce. One of the exceptions is the usage of taste in more playful scenarios. It can be imagined that taste could be used in kids' applications, for example to stimulate kids in a learning application (as bonus when passed a test) or to train them to recognize tastes. Taste could also be usable in art installations, possibly in some kind of wild application (like "eat the painting" to reveal what is behind it). Hence, it can be concluded that the application of both smell and taste needs to be further researched to understand its effects.

2.2.5 Vestibular

Perception

The vestibular sense can be best understood as the human balance system. Its structure (Figure 2.19) is situated behind the ear and consists of the otolith organs and three semicircular ducts. The otolith organs (the utricule, also called vestibule, and the saccule) contain small hairs that are stimulated by movement of fluid within the organs, thereby delivering information on the direction of gravity and linear acceleration. The three semicircular ducts are placed in three perpendicular planes and, just like the otolith organs, also contain fluid that can stimulate the hair cells inside. The semicircular ducts provide information on angular acceleration and deceleration. The vestibular nerve system consists of seven nerves in the otolith organs and semicircular ducts and is part of the 8th nerve, also known as acoustic nerve.

As can be concluded, the otolith organs deal with linear movement, whereas the semicircular ducts register rotational movement, thereby contributing to the human sense of balance. The system also provides cues for bodily motion, having both a static and dynamic function. The otolith system monitors the position of the head in space, contributing to the control of the human posture. The semicircular ducts track the rotation of the head in space, affecting the so called vestibular eye movements. These movements allow people to keep the focus fixed on an object while the head is moving. All together, the vestibular system affects the control of the eyes, neck, and trunk/limbs (Goldstein '02). One of the side effects of the vestibular system is its influence on motion sickness. It is believed that motion (or simulator) sickness is caused by the mismatch between vision cues and vestibular cues. Methods for avoiding motion sickness via the contribution of vestibular cues can be found in (Harris, Jenkin et al. '99).

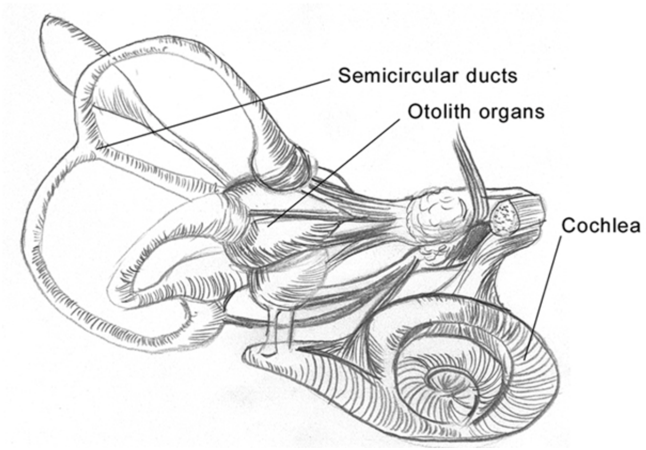


Figure 2.19: *The vestibular system.*

Hardware interfaces

Hardware interfaces that provide or support the acquisition of vestibular information can be subdivided in devices that either support natural or artificial movement.

Natural movement support

Devices that support natural motion enable users to move around freely, by monitoring the user via some kind of tracking system. The range of movement varies between systems. To support a high amount of freedom for moving around, either wide-area tracking systems, such as HiBall (Welch, Bishop et al. '99) or other vision-based methods, are needed.

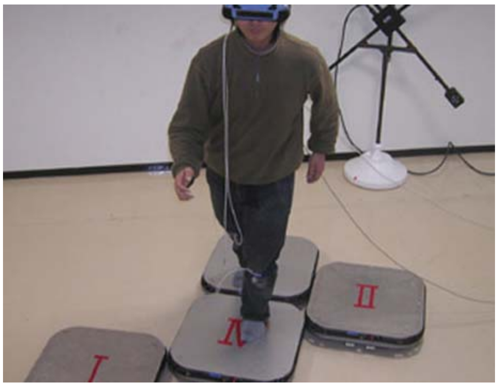


Figure 2.20: *CirculaFloor walking interface. Courtesy of University of Tsukuba*

Artificial movement support

The biggest group of artificial motion support devices is motion platforms, ranging from simple single person platforms, large motion platforms (as used in flight simulators), up to specialized devices, such as the Gaitmaster (Iwata '01) or the CiculaFloor interfaces (Figure 2.20). The latter is a truly innovative interface that allows for omnidirectional movement on plain level by making use of robotic floor tiles that move around, sensing the user’s direction (CirculaFloor '04). Another possibility is to make use of the wide range of sport devices, like bikes, up to stranger constructions resembling hang gliders (Soares, Nomura et al. '04). Some of these interfaces provide only limited motion cues, though. A comprehensive overview of locomotion devices can be found in (Hollerbach '02).

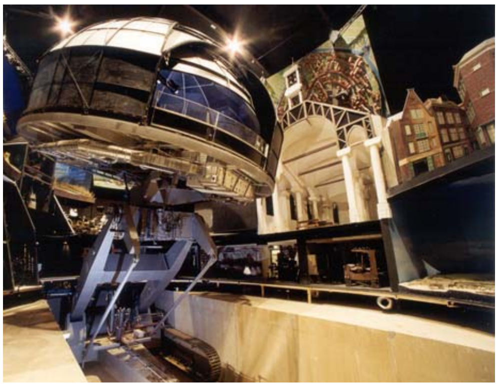


Figure 2.22: *HollandRama installation.*
Courtesy of Rexroth Hydraudyne

Though traditionally coming from the field of simulation systems, artificial motion systems find wide applicability. Especially interesting are those installations in which motion platforms find their way in completely new application areas. Particularly interesting is the HollandRama installation, which makes use of a massive motion platform supporting around 160 people (Figure 2.21) to move museum visitors along a multitude of real panoramic scenes built in a dome-like building. Though not a mixed reality system, the system applies interesting unconventional human input methods, ranging from fine acceleration simulation resembling flying in a balloon, surround audio and visual perceptual illusions, as can be seen in Figure 2.22 when looking at the architectural reconstruction (church interior) directly right of the capsule.

A smaller group of devices, with an increasing amount of research, is based on electrical stimulation of the 8th cranial nerve, through which the vestibular data flow. Most often, the surface above the mastoid bone, close to the ear, is stimulated. Multiple devices exist, ranging from the (no longer produced) Motionware device, or the galvanic stimulation interface like (Harris, Jenkin et al. '00; Maeda, Ando et al. '05), as shown in Figure 2.22. Another possibility is to make use of a vestibular implant which works similar to the cochlear implant described in section 2.2.2 (Wall, Merfeld et al. '02).

Figure 2.22: *Galvanic vestibular stimulation interface.*
Courtesy of M. Jenkin

Application

Vestibular stimulation predominantly affects navigation interfaces. They greatly enhance *traveling* by moving away from passive movement techniques like pointing, especially when physical movement or exertion is part of the task. Such tasks include military training scenarios and sports applications. Nonetheless, reliability of physical motion devices, like a treadmill, is not very high, and the devices are extremely costly. On the other hand, wide-area tracking is becoming available at lower prices and seems reliable, making this technique interesting as long as the physical space moved through is no limitation.

It has to be stated that the truly vestibular component of natural motion does not directly affect the quality of motion, but foremost affects *wayfinding*. Multiple tests (Chance, Gaunet et al. '98; Klatzky, Loomis et al. '98) have shown positive effects of real motion on spatial orientation, since important information on depth and direction is provided.

As a result of the addition of vestibular cues, it can be stated that the sense of presence may increase, since the actual performance of tasks mimics the real world in a much better way (Usoh '99) – the inter-sensory conflict between visual movement and bodily movement can be partially or fully broken, thereby also (as stated before) reducing motion sickness.

2.3 Human output

Human output primarily refers to actions that are afforded by our joints and muscles (Figure 2.24 and 2.24). These different actions are response actions to a human input (stimulus). As can be read in section 2.1, the peripheral nervous system triggers effectors, via electric signals, that can result in both voluntary (motor) and involuntary actions.

Most of the human output can be defined as a *control task*, previously also labeled as interaction tasks such as manipulation or navigation. These tasks couple the body to some kind of controller, thereby creating a control-body linkage between the human and a computer input device. This control-body linkage can be based on physical contact or by ways of monitoring.

The control task can be characterized by its accuracy, speed and frequency, degrees of freedom, direction and duration. Task characteristics directly affect the human output channel. Some tasks may be performed with only a certain body part, whereas with other tasks this may be possible via multiple body parts (Figure 2.23). The ability to perform a task with alternative body parts can lead to *control substitution:* exchanging one body output channel with another one, under the premise that the task can be mapped to both output channels. As can be seen in the next sections, control substitution has been the basis for a large number of unconventional interfaces. Hence, a closer look will be taken at the different possibilities of the joints of the human body. It is of importance, though, that control actions can also be mapped on other human systems, such as the eyes or brain. These systems, known under the name of biocontrol or biofeedback systems, allow for the performance of tasks that are *not* necessarily based on motor actions.

	shoulder	forearm	wrist	fingers	hip	knee	ankle	toes	mouth
flexion	▨	▨	▨	▨	▨	▨		▨	
plantar flexion							▨		
dorsiflexion							▨		
extension	▨	▨	▨	▨	▨	▨		▨	
abduction	▨		▨		▨				
adduction	▨		▨		▨				
medial rotation	▨				▨				
lateral rotation	▨				▨				
pronation		▨							
supination		▨							
opposition				▨					
inversion							▨		
eversion							▨		
depression									▨
elevation									▨
retraction									▨
protraction									▨
gliding			▨					▨	

Figure 2.23: *Movements afforded by joints.*
(selection of joints in human body)

Figure 2.24: *Overview of possible movements of joints. Adapted from (Goldstein '02)*

The overview of both general muscle-based and biofeedback systems will only handle some basics. For a more detailed overview of human anthropometric mechanisms, refer to (Salvendy '97; Goldstein '02). For a general overview of control issues, please refer to (Bullinger, Kern et al. '97).

An important issue is the separation of different motion types allowed by the biomechanical system. As can be seen in Figures 2.23 and 2.24, the joints afford three major kinds of movement: namely gliding and angular movements, rotation, as well as some special movements. From the perspective of control substitution, Figure 2.23 shows the exchangeability of specific control tasks between different body parts. Some conclusions can be made when looking at the table:

- Moving an object in a plane is rather easy
- Moving or controlling mounted devices in possibly multiple degrees / planes can be done by multiple body parts
- Moving a non-mounted object in multiple planes is restricted to just a couple of body parts
- Fine-grain actions are only possible with the hands, due to the dependency of many techniques (or devices) on the opposition possibilities of the fingers.

Besides the possibilities of the joints (and thus body parts), the possibilities of the muscles need to be taken into account when examining a control task. Muscles have three main characteristics, namely force, velocity, and duration of contraction, building a bone-muscle relationship that works as a lever system. Some control tasks cannot be substituted due to muscle limitations, since the load of a task cannot always be transferred to certain muscles.

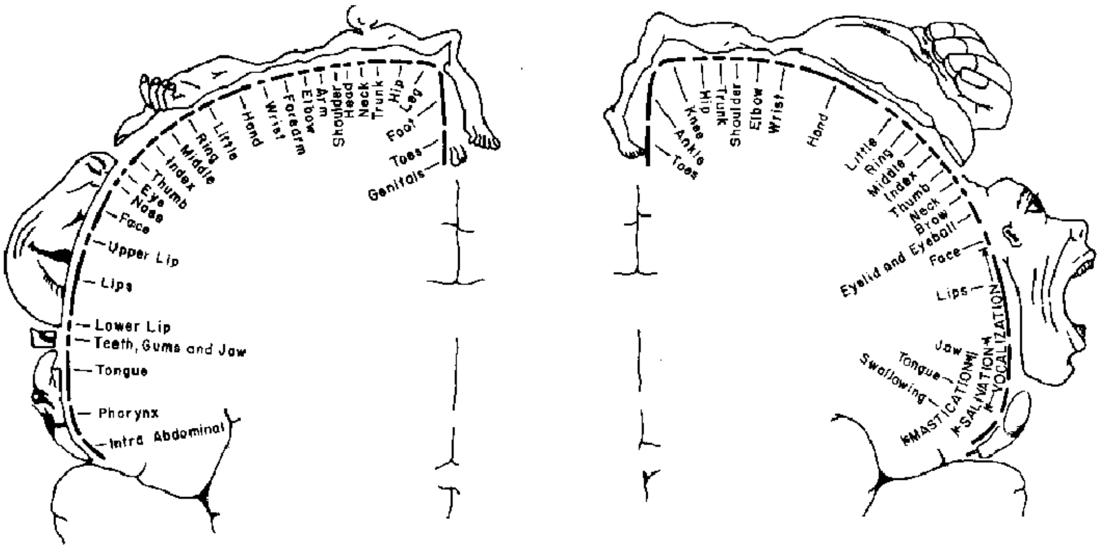


Figure 2.25: *Penfield and Rasmussen's homunculus.*
Courtesy of Penfield and Rasmussen

The sensory-motor distribution of the cortex is of importance for the performance for the different body parts. The mapping of the different body parts to the areas in the cortex was pioneered by Penfield and Rasmussen, in a study published in 1950 (Penfield and Rasmussen '50). One of the outcomes is the famous *homunculus* (Figure 2.25), a strange looking figure that represents the mapping of the body parts to the cortex. In interaction studies, the map is generally used to show the possible “precision” with which tasks can be performed: the larger the body part is presented, the more likely it is that more precise actions are afforded.

In the following sections a similar approach is taken as in section 2.2, by discussing the psycho-physiological background characteristics (control), technology (hardware interfaces) and usage (application) of the human output channels.

2.3.1 Head – mouth and throat

Control

The head offers a multitude of output channels, including the mouth and throat, the eyes, the face, and the head itself. In this section, the mouth and throat are described. The next section (2.3.2) describes the head and eyes.

The mouth and throat (Figure 2.26) consists of the oral cavity, lips, jaw and chin, teeth, tongue, and the vocal cords with the related larynx and pharynx. Interestingly, the mouth and tongue occupy about the same amount of sensory and motor cortex as the fingers and hand, hence potentially allow for precise interaction (Figure 2.24).

From an HCI perspective, the mouth and throat are mainly used for speech production. Speech is produced as air is pushed from the longs towards the mouth, being shaped into patterns of air pressure by actions from the different structures of the vocal tract (like the tongue, nasal and oral cavities, and teeth) (Goldstein '02). Furthermore, teeth, jaws, and tongue support biting and licking. It should be stated that the tongue is a strong muscle and highly flexible. Finally, mouth and throat are used for respiration, spittle production, and lip-oriented tasks like kissing, and pressing something (like a straw) between the lips.

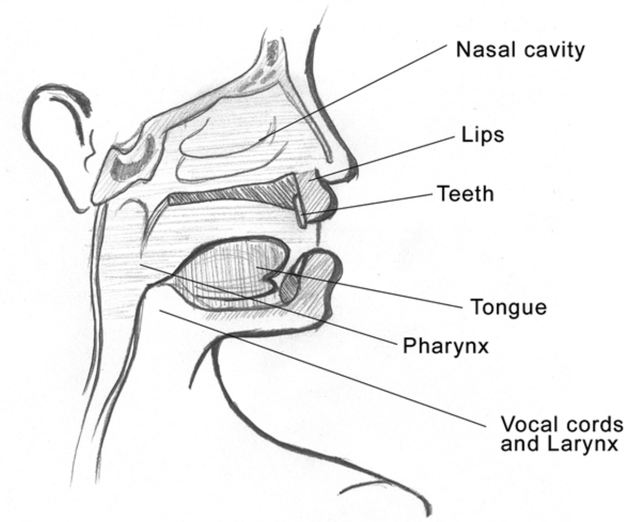


Figure 2.26: *The mouth and throat.*

Hardware interfaces

Speech systems

The most common interfaces used with the mouth and throat are definitely speech systems, consisting of both speech recognition systems, and the more advanced spoken dialogue systems. A large number of microphone-based speech systems exist, including IBM ViaVoice, BBN Hark and Speechworks (for an overview, see (McTear '02)), with rather high recognition rates.

Some new directions in speech systems exist, including the usage of electromyography (EMG) to detect the speech producing organs, placing detectors on the lip, cheek and under the jaw (Manabe, Hiraiwa et al. '03). Such systems can be used for “speech-less recognition”, since one does not actually need to produce hearable speech for the recognition engine.

Figure 2.27: *Pipe interface using wind pressure sensor.*
Courtesy of G. Scavone

Respiration sensing

Rather unconventional is the sensing of respiration, to detect blowing behavior like playing a flute or blowing out a candle. There are several ways to sense breath, including wind sensors (Hatana, Masui et al. '03), thermal conductors (Makinwa and Huising '01), the analysis of the speech spectrum (Iga '99). Also possible is the sensing of pressure caused by air, like done by Scapone (Figure 2.27, (Scapone '03)), in Fels et al’s two-person breath controller Tooka (Fels and Vogt '02), or with Sip ‘n Puff, a pneumatic switch (Therafin '05). Finally, in the Davies’ famous Osmose application, breath was detected by measuring the width of the chest (Davies and Harrison '96).

Mouth cavity and tongue sensing

The sensing of mouth cavity and tongue for other purposes than speech has been performed in different forms. Based on general vision-based up to using ultrasound scanning methods, several researchers have made applications that made use of the contour of the tongue (Vogt, McCaig et al. '02), or the shape parameters of the mouth (Figure 2.28, (Lyons, Haehnel et al. '01)) to control application parameters. A direct way of controlling an application with the tongue using a tongue-controlled joystick has been successfully tried by Salem and Zhai (Salem and Zhai '97).

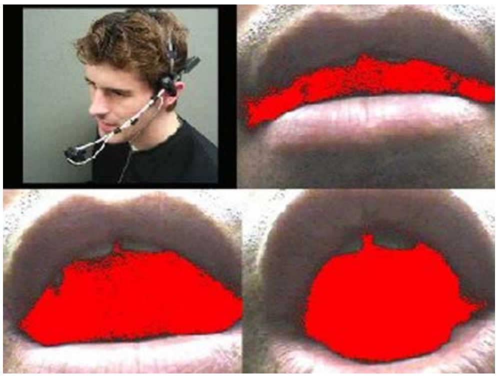


Figure 2.28: *Mouth interface using cavity sensing.*
Courtesy of M. Lyons

Chin joysticks
Chin-controlled joysticks, which are generally used as assistive technology, form a well accepted way for disabled people to control a wheelchair. Mounted on a plate close to the head, disabled people can rather easily move around the joystick to steer in the same way as using a hand-controlled joystick. Such joysticks may also hold several buttons that can be pressed by the chin (Pride '05).

Biting interface
Finally, biting, spitting and kissing have been largely ignored in computer interfaces. The previously mentioned Food Simulator from Iwata (see section 2.2.4) partly deals with biting, but kissing and spitting interfaces though have not really been tried out yet, even though simple sensors (like a touch sensor) could be used for these. This may be due to social or ethical issues.

Application
Even with their long existence (well over 10 years), *speech* interfaces still seem to be a bit unconventional, or at least exotic, since they are not often applied. Mostly, they seem to be used in *multimodal interfaces*, combined with another modality, like gestures (Bolt '80; Billinghurst '98). Direct speech input, with a limited vocabulary can be used for simple commands in a system control interface, but for more complex tasks, it is generally discarded. Some newer interfaces like the one by Latoschik (Latoschik '01) add more “intelligence”, by means of semantic models to the interface that may provide better performing interaction with more complex applications.
The range of head or tongue-controlled joysticks can also be used as a navigation interface. Navigation may not be as fluid and precise as by hand, even though training will increase performance (Salem and Zhai '97). Breath-based navigation, as used in Osmose (Davies and Harrison '96) is necessarily two-dimensional, since breath only has two parameters: breathing in and out.
Breath can also be used for other *linear movement tasks*, like selecting an item from a list, even though this may seem tedious. On the other hand, a pneumatic switch like the previously mentioned Sip ‘n Puff can be easily used for selection purposes, either in manipulation or system control tasks. Nonetheless, it can be expected that breath will most likely continue to be used predominantly in art of music-oriented applications, by enabling rough interaction for non-demanding tasks or to mimic musical instruments, like Scavone’s Pipe.

Finally, as Vogt et al. (Vogt, McCaig et al. '02) and Lyons et al. (Lyons, Haehnel et al. '01) demonstrated, the mouth and tongue can be used to control any kind of *generic parameter*. Probably matched best to expressive tasks like musical control, they can also be used for other tasks like drawing or text entry in combination with another modality like the hand (Chan, Lyons et al. '03).

2.3.2 Head – face and eyes

Control

The human eyes are both an input and an output channel. Head and eye based control can include small gazes from the eye and large gazes from the head. There are several eye movements: brief eye fixations (about 600ms) connected to ballistic eye movements (also known as saccadic movements, 400 – 600^o per second, within a 1-40^o visual angle) and general eye movements between a range of 15-20^o with a duration of 30ms. Head movements can be up to 250^o per second, up to smaller arcs of 800-1000^o per second. The head and eye movements are strongly coupled to attention mechanisms.
The face consists of a large number of muscles, thereby allowing a highly diverse play with the skin. They lead to different facial expressions, used both to illicit emotion and in social interaction.
Finally, the head allows a wide variety of movements, including multiple kinds of extensions and rotation (Goldstein '02).

Hardware interfaces

Head tracking

Enabled by a large number of different tracking systems, from magnetic and optical to hybrid ones, head tracking is a common interface for VEs. A complete overview of tracking systems can be found in (Bowman, Kruijff et al. '05).

Eye tracking

Multiple techniques exist that allow the tracking of ocular (eye) movement (Richardson and Spivey in press (a)). Reflective light techniques measure the light reflected by the pupil or limbus, for example by using infrared light (Figure 2.29). Next to head-mounted systems, simple video camera based systems also exist that can be placed away from the user.
Another technique, labeled dual Purkinje image, measures the disparity between different reflections in the eye. The corneal reflection and the reflection of the rear of the lens in the eye are measured, adjusting a series of mirrors with servomotors until the two images are superimposed on electric photoreceptors. Multiple methods based on contact lenses exist, ranging from contact lenses embedded with small mirrors (reflection-based technique), to non-optical methods that make use of a contact lenses that include orthogonal wire coils tracked by a magnetic tracking system.
Finally, several electro-physiological methods exist (Sörnmo and Laguna '05), which make use of electrodes near the eye. Corneal-retinal potential is registered by electrooculography (EOG), or full-field electroretinogram (ERG), which measures the electrical potential generated by the retina during light stimulation. A more detailed discussion on electro-physiological methods, being part of biopotential interfaces, can be found in section 2.3.6.

Figure 2.29: *Skalar IRIS IR Light Eye Tracker.*
Courtesy of Cambridge Research Systems Ltd

Application

The usage of eye tracking can be found in several application areas. Predominantly, it has found its way in numerous psychological experiments that investigated the user's *focus and attention* on graphical user interfaces, or the hand-eye coordination issues of using different input devices (Smith, Ho et al. '00). An extensive overview of psychological-oriented applications can be found in (Richardson and Spivey in press (b)). For interaction purposes, eye tracking has been moist often used in assistive technology applications, aiding disabled people in controlling a computer in a hands-free manner (Cleveland '94; Istance, Spinner et al. '03). A large group of applications deal with eye-based typing, also called *eye typing*, using a virtual keyboard, shown on a display. By using the eye gaze to control a keyboard, people with motor disabilities can provide symbolic input to a computer. Coupled to a speech generation system, these systems can also be used for speechless communication. An overview of different eye typing systems can be found in (Cogain '05).

Eye typing basically controls a cursor on a two-dimensional plane. Hence, the eye tracking can theoretically be used for any task in a 3DUI that involves translation in two dimensions (2D selection task). Such tasks are normally *symbolic input*, but it can also be used for *system control*, for example to select items from a menu or to control a switch. Generally, cursor-based control tasks are coupled to a zooming technique in order to increase accuracy (Fono and Vertegaal '05).

Whereas the eye-controlled tasks can be found in several areas, the usage of the head has predominantly found its way in *navigation-oriented tasks*. Similar to eye typing, the head is generally used in combination with assistive technology. Most often, it is used for wheelchair control, to allow steering through natural environments by ways of head-joystick coupling. Using the same method, the head can touch a switch like Quantum Rehabs Switch-it Head switch array (Pride '05) or other mounted device.

Next to using a head-operated joystick, head gaze has been regularly probed in *gaze directed steering* techniques. Hereby, generally the orientation from a tracked head sensor is used to map gaze to direction of flight (Mine '95a). Studies have shown, though, that pointing outperforms gaze in specific travel tasks (Bowman, Koller et al. '97). In addition to steering, some manipulation techniques exist that base selection on the gaze of a user, for example in combination with hand pointing. An example is Forsberg's aperture selection technique (Forsberg, Herndon et al. '96).

Functionally, the mounted joystick can be used for navigation or any other task that can be mapped to the joystick (*two dimensional pointing*) in a VE, whereas switches or buttons can be used for one-dimensional system control. Most likely, control will not be as fine as with the eyes. The actual effects would require a comparison of head-to-mounted-control coupling with eye-based control in a formal evaluation though. For a more detailed overview of eye tracking issues, please refer to (Jacob '95; Wilder '02; Duchowski '03).

Going on step further, the field of *perceptual user interfaces* (PUI) explores multimodal and perceptive capabilities (monitoring a user's behavior via machine perception and reasoning) to create more “natural” user interface (Turk and Robertson '00). Eye gaze can provide focus cues, which can potentially lead to more advanced feedback methods, error reduction, or to dynamically adapt user interface functionality.

Finally, related to the usage of gaze for behavioral reasons is the tracking of *facial expressions*. Primarily used for facial animation, the face can also provide important information, for example to detect surprise (“why is my computer doing that?”), in order to deliver adapted feedback. For a complete survey on face recognition techniques, please refer to Zhao et al. (Zhao, Chellappa et al. '03).

2.3.3 Hand and arm

Control

The hand is the most dominant human output channel and allows one to perform a tremendous number of tasks, using highly diverse devices. The musculoskeletal configuration of the hand (Figure 2.30) consists of three bone segments: the wrist (carpus), the palm (metacarpus), and the fingers (phalanges).

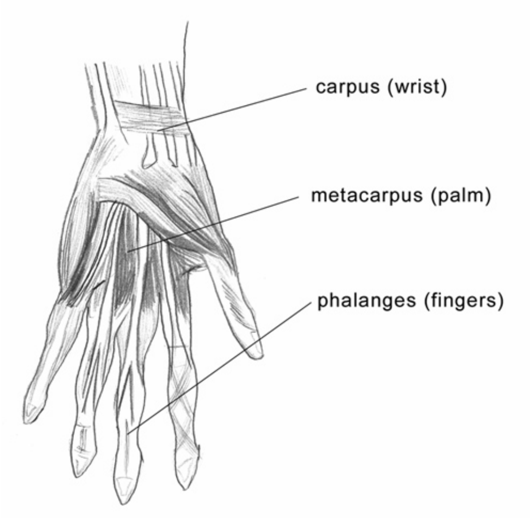


Figure 2.30: *The hand.*

The hand contains three muscle groups, found in the thumb, the little finger, and the palm. The musculoskeletal system affords numerous movements, including gliding movements of the wrist, angular movements (abduction, adduction, flexion, extension) of the fingers and the wrist, and “pinching” of the fingers via opposition. All together, the constellation of the hand, wrist, and arm is a lever system and allow for a large number of control dimensions. Nonetheless, not all the configurations are comfortable. Based on factors like force, duration and repetition, task performance may be limited (Marras '97). The hand allows both coarse and fine-motor actions. Depending on the hand-device coupling and the grip, humans can perform actions using a power grip of or a precision grip. A power grip is mostly performed by holding a device within the palm of the hand, whereas the opposition possibilities of the fingers allow fine motor control when a device, or a part of the device, is held between the fingers. A good source for the different musculoskeletal effects on the usage of (3D) input devices is (Zhai, Milgram et al. '96).

The hand can be used for unimanual and bimanual tasks. Depending on hand preference, one hand (the dominant hand) is more sensitive than the other hand. In bimanual action, the non-dominant hand forms the frame of reference for the dominant hand (Guiard '87). In our daily life, the hand is used for a virtually endless number of different tasks, including grasping and holding, squeezing or bending, pushing and pulling, gestures and postures, hitting or knocking, and throwing.

Hardware interfaces

General hand-operated devices

Providing a complete overview of all hand-used input devices is hardly possible. Generally used devices include: mice, keyboards, digitizing tablets, game controllers, and touch screens. For a large overview of available devices, please refer to Buxton’s list of input devices (Buxton '05).

Within 3DUIs, interfaces include 3D mice like the Bat (Ware and Jessome '88) or similar flying mice (like the Spacemouse (3Dconnexion '05)), wireless mice (like the Bug (Stefani and Rauschenbach '03)), or user-worn 3D mice (like the FingerSleeve (Zeleznik '02)), multiple kinds of gloves (Defanti '77) (Zimmerman '87) (Kramer '91), vision-based hand or arm detection (Leibe, Starner et al. '00), pen-and-tablet interfaces (Poupyrev '98a) (Szalavari and Gervautz '97), and haptic devices (Burdea '96). For more information, please refer to chapter 4 in (Bowman, Kruijff et al. '05).Mechanical controls such as push buttons, switches, wheels, and knobs are used as singular control, or combined to form some kind of device. These devices often are “garage design”-style devices. For an introduction, please refer to section 3.6.

Mixed devices

There are several more unconventional devices that combine different kinds of hand-operated devices into a single device. One example is Immersion’s haptic glove construction, which combines a ground-referenced device (a mechanical desktop haptic device that resembles the SensAble Technologies Phantom) with a haptic glove providing body-referenced feedback and human input (see Figure 2.12).

Another example is the Control Action Table (CAT, Figure 2.31). The CAT is a special kind of construction, in which a pen-tablet device is mounted in something like a circular tabletop (Hachet '03), demonstrating a high level of integration of different kinds of interaction technology. Within the tabletop construction, angular sensors and potentiometers are mounted. The potentiometers do not only sense force, but also translational movement. Hence, the CAT combines a 2D tablet with a fully 3D (6DOF) interface metaphor. Even though some rotations and translations may be hard to perform, the CAT supports constrained interaction by allowing singular DOF control. The CAT is now commercially available from Immersion (Immersion '05).

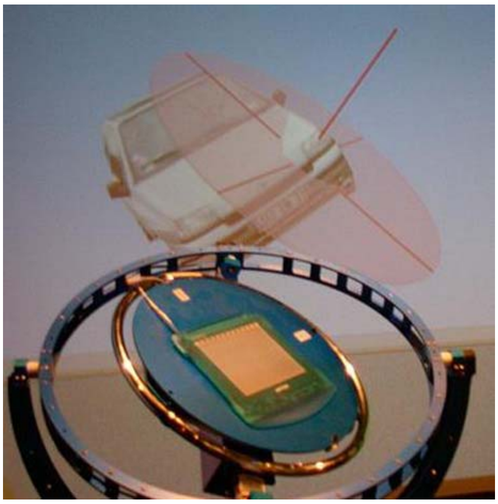


Figure 2.31: *Control Action Table.*
Courtesy of M. Hachet

Special purpose devices
Numerous special-purpose devices exist that perform a very specific function. The biggest group of these special-purpose devices is called *props* (Hinckley '94). Props are passive real world devices generally used in different kinds of tangible interfaces (Ishii and Ullmer '97). Examples are Froehlich and Plate's Cubic Mouse, a coordinate system prop using rotational and translational rods fit into a box (Froehlich '00) (see section 4.5), the CavePainting application, using small paint buckets and a brush (Keefe '01), and Pierce Voodoo dolls technique using real dolls for remote manipulation actions

Figure 2.32: *Bongos controller.*
Courtesy of Nintendo

The usage of props is very popular in the games industry, resulting in a large range of devices that mimic real-world devices. This can be as simple and wide-used as steering wheel interfaces to control car games, up to more exotic devices like Nintendo's Bongos for the Donkey Kong Jungle Beat game (Figure 2.32, (Nintendo '05)) or Taito's Real Puncher, a boxing interface (System16 '05).

Other kinds of special-purpose devices exist that do not necessarily have a real-world similarity. A good example is ShapeTape, a flexible tape that looks like a ribbon, consisting of fiber-optic curvature sensors. The device provides bending and twisting behavior and is used in 3D curve design and system control (Grossman '03).

Combining different controllers in a single input device allows for a wide range of task-specific devices. An example of an "expressive", and very task-specific controller is demonstrated by Merrill's Adaptive Music Controller (AMC, Figure 2.33, (Merrill '04)). The AMC combines controllers such as accelerometers, gyros and bend sensors in a single device to support "free-gestures" for musical control. The AMC thereby specifically focuses on the association between gesture and sound.

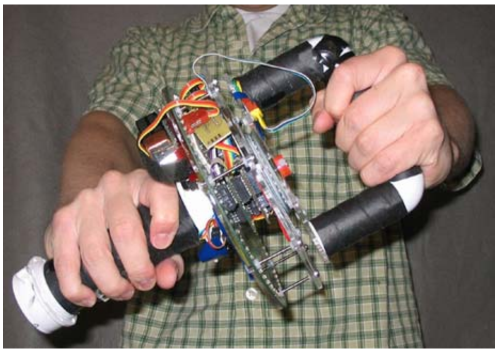


Figure 2.33: *Adaptive Musical Controller.*
Courtesy of D. Merrill / J. Paradiso

Gesture-based touch interfaces

Related to the usage of gestures and postures, there are several new(er) developments that make use of projection screens combined with touch sensitive surfaces. Within 2D projection (Figure 2.33), such screens have been used frequently (Rekimoto and Matsushita '97; Streitz, Geissler et al. '99), but within VEs, they are rather exotic.

The combination of screen and touch sensitive surface brings forth the possibility of closely coupling 2D interaction in a VE, even if this still poses several problems (see chapter 10 in (Bowman, Kruijff et al. '05)). An example of a workbench-like device with a touch-sensitive surface is Rekimoto's SmartSkin (Rekimoto '02).

Application

Throughout the years, a huge amount of research has been performed on hand-operated devices and their task-performance. This research has been predominantly carried out for desktop applications, but much also applies to 3DUIs. Some of the basic references are:

- *General literature on input devices*: devices in desktop applications (Preece '02) (Shneiderman '98), devices in 3DUIs (Jacob '96), nonconventional devices (McMillan, Eggelston et al. '97), design space of interaction devices (Card, Mackinlay et al. '90), and design issues of spatial input (Hinckley, Pausch et al. '94)

- *Human factors and evaluation of devices used in 3DUIs*: general literature on human factors in VEs (Stanney, Mourant et al. '98), hand-device coupling (Bullinger, Kern et al. '97), effects of muscles (Zhai, Milgram et al. '96), perceptual structure (Jacob '92), user performance, and 3D input device design (Zhai '98a)

The application of hand-operated devices shows a rather diverse image. There seems to be an off-balanced acceptance rate and resulting “unconventionalism” of some devices: some devices are highly accepted among a small amount of people, but highly unconventional for the larger public. An example is the field of haptic devices. Even so, several interaction-oriented areas can be identified that show unconventional approaches.

Most hand-oriented *manipulation* techniques have been accepted over time. Some directions, though, are still less accepted or developed or even unconventional. Tasks like delicate (fine) haptic manipulation, for example used for medical applications, are still quite uncommon, just like complex or more advanced modeling gestures and postures. Examples include SmartScene (Mapes and Moshell '95), or multimodal techniques combining gestural interaction and speech recognition (Latoschik '01). Several techniques take the approach of supporting unreal interaction (so called “magic” interaction) by mimicking comic character behavior, such as the Go-Go interaction technique inspired by the Inspector Gadget series (Poupyrev, Billinghurst et al. '96a).

Figure 2.34: *Griffin made with Topobo.*
Courtesy of H. Raffle

The field of *tangible user interfaces* (Ishii and Ullmer '97) keeps coming up with new interfaces, like those introduced in section 2.2.3, introducing “real” objects in VEs, up to complete toolkits like Kitamura et al.’s ActiveCube (Kitamura, Itoh et al. '01) or Raffle et al.’s Topobo (Figure 2.34, (Raffle, Parkes et al. '04)). These toolkits resemble the LEGO Mindstorms toolkits and can not only be used to make toy-like constructions, but also to couple *dynamic physical actions* into VE interactions. For example, the usage of such toolkits can greatly enhance interaction in molecular visualization environments by a possible real-world to virtual relationship between the physical molecular construction and its virtual representation.

Outside of the usage of hand direction for simple camera movements like the camera-in-hand technique (Ware and Osborne '90) or pointing, there are not many unconventional hand-oriented *navigation techniques*. Some exceptions are the grab-the-air (or "tug of war") technique used in SmartScene (Mapes and Moshell '95). Another one is the gesture-based technique resembling "finger-walking," which was intended to be used for control of robots (Sturman, Zeltzer et al. '89).

A field still serving as incentive for unconventional interfaces is *two-handed interaction techniques*. Based on earlier research, like from Buxton et al. (Buxton and Myers '86) (Kabbash, Buxton et al. '94) or Guiard's framework (Guiard '87), multiple unconventional interfaces have been created. Interfaces include medical interfaces using props (Goble, Hinckley et al. '95), Pierce's two-handed manipulation technique using a real doll called Voodoo dolls (Pierce, Stearns et al. '99), and two-handed haptic input based on the Iwata et al's Spidar system (Murayama, Bougrila et al. '04). Several studies have shown that asymmetric bimanual techniques, in which the left and right hand perform different, but dependant actions, can have a significant performance increase (Hinckley, Pausch et al. '97b; Balakrishnan and Kurtenbach '99b).

Besides general 3DUI tasks, several hand-controlled "special" tasks can be identified. These tasks include: the performance of sign language by making hand gestures (Fels '94) and a range of sports / active-movement oriented tasks, like boxing or hitting (see the previously mentioned Taito real Puncher), or throwing, as performed in the ImpactTV application. In this application, tracked (sport-related) objects, like a ball, are thrown at a large TV projection screen (Figure 2.35), functioning as a way of remote controlling TV channels (based on Mueller's Impact system (Mueller '02)).

Predominantly focused on hand-based control, a large group of multi-user environments exist that support *cooperative, distributed, and non-distributed interaction*. Even though multiple examples exist that support collaborative interaction, like Studierstube (Schmalstieg, Fuhrmann et al. '02) or MASSIVE (Greenhalgh and Benford '95), most environments duplicate single user interaction or simply add a videoconferencing component and allowing interaction via turn-taking mechanisms.

Figure 2.35: *ImpactTV throwing interface.*
Courtesy of F. Mueller

Some unconventional cooperative applications exist that demonstrate "dependant" interaction between two persons. One example is the Hubbold's collaborative stretcher carrying application (Figure 2.36, (Hubbold '02)), which uses multiple haptic devices to mimic the handles of a stretcher. Another field that may yield new interaction techniques is co-located interaction in which two persons interact simultaneously, using the same projection display, but with independent viewpoints (Simon '05).

Figure 2.36: *Collaborative stretcher carrying using haptics. Courtesy of R. Hubbold*

2.3.4 Feet and legs

Control

The foot (Figure 2.37) consists of three bone segments: the tarsus (7 bones), the metatarsus (5 bones), and the phalanges or toes (14 bones).

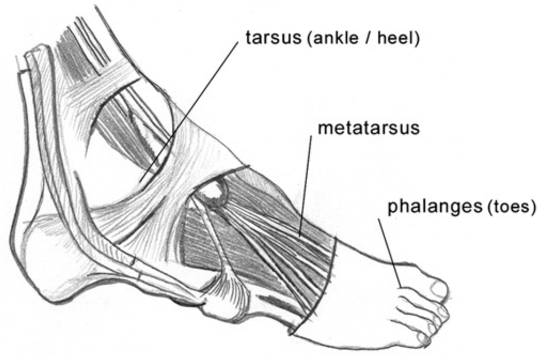


Figure 2.37: *The foot.*

Several muscular structures run over the feet, including the fibrous bands from the ankle, the dorsal muscle on top of the feet, and multiple layers of plantar muscles under the feet. The musculoskeletal construction of the feet (ankle and toes) allows several movements that partly resemble the movements of the hand: plantar flexion, dorsiflexion, inversion, eversion, flexion, extension, and gliding (see table in section 2.3). Hence, some tasks performed by the hands could potentially be (through control substitution) performed by the feet.

Not surprisingly, the leg-wrist musculoskeletal construction is, just like the hand-wrist-arm construction, a lever system. The main purpose of this system is to support a possible upright posture and the control of human motion. Besides walking, control-oriented behavior is predominantly coupled to the feet. Only incidentally, the knee or shinbone is used to push or hit a control mechanism. Generally, the feet only allow coarse control actions (Bullinger, Kern et al. '97; Goldstein '02).

Hardware interfaces

Desktop-device mimicking controls

The most basic foot controls are those devices that are more or less enlarged (desktop) controls. These include: foot-pedals (essentially a large button), large floor mounted trackballs that can be turned by the feet, and floor mat game controllers like the wireless Intec Dance Mat controller (Intec '05). Some devices demonstrate hybrid approaches, combining conventional button-like controls with unconventional input or output. An example is Mohamed and Fels' Pedal Panner for musical control, combining a foot mouse with vibrotactile output (PedalPanner '05).

Pressure-sensing devices

The second group of inputs is pressure-sensing devices. These devices integrate capacitive sensors to sense simple forces for location tracking (Leikas, Strömberg et al. '04). Some devices measure complex force distributions (weight, balance), by using arrays of sensors. An example is the Emed-x system, with up to 6080 sensors in a 475x320 mm area (Emed-X '05)). Other devices, such as the Z-tiles system (Figure 2.38, (McElligott, Dillon et al. '02)), are essentially floor tiles that can be coupled to create matrices of tiles, forming a complete floor (Srinivasan, Birchfield et al. '05).

Figure 2.38: *Z-tiles floor system.*
Courtesy of McElligott et al.

Motion tracking devices

Next to the tracking of location on floors, the determination of motion is performed by a large number of locomotion interfaces like the omnidirectional treadmill (Darken, Cockayne et al. '97) or the Gaitmaster (Iwata '01), and the rather unconventional VirtuSphere interface (Figure 2.39). Please refer to (Hollerbach '02) for a general overview of these devices. Apart from general locomotion devices that focus on walking, multiple sports devices are being used. Probably the most used sports device are cycles, such as the one used in the classic Legible City application from Shaw (Shaw '05), or newer versions using fitness bikes like by Välkkynen (Välkkynen, Heinilä et al. '01).

Figure 2.39: *VirtuSphere walking interface.*
Courtesy of VirtuSphere (VirtuSphere '05)

A device that needs to be handled separately is Rekimoto and Wang's Sensing GamePad (Figure 2.40, (Rekimoto and Wang '04)). Within the gamepad, two electrodes are hidden that can sense static electricity changes that occur when a user lifts her foot from the floor, thereby changing the capacity "loop" the human has with the ground.

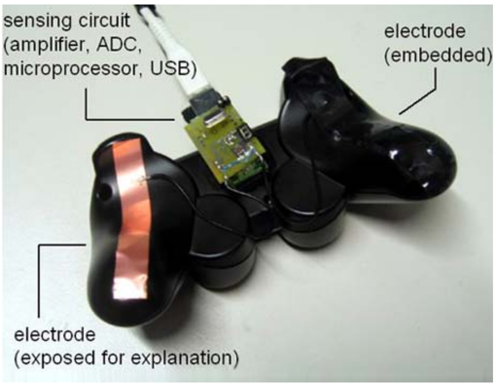


Figure 2.40: *Sensing GamePad.*
Courtesy of J. Rekimoto

The change is compared to the neutral electrostatic potential level in order to detect movement, filtering out specific known noises, such as those caused by clothes rubbing over each other. The method resembles body capacity methods used for full body interaction, as handled in section 2.3.5. Though only detecting footsteps and not complete motion patterns, Rekimoto and Wang imagine interesting applications possibilities, such as simulating a virtual pedal for system control or sports training.

Biometric behavior tracking
Analysing the complete biometric behavior of the feet is handled by Morris and Paradiso's *gait sensor system* (Figure 2.41, (Morris and Paradiso '02)).

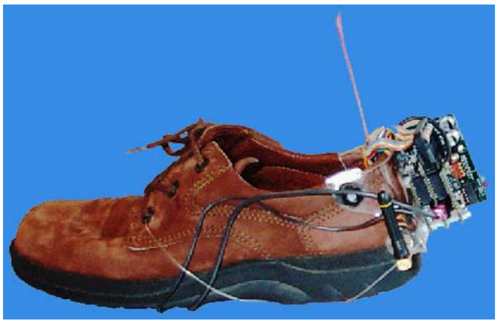


Figure 2.41: *Gait sensor system using multiple sensors.*
Courtesy of J. Paradiso

Mounted in the sole are several sensors: force sensitive resistors for striding time and weight distribution of the feet, heel and toe strike sensors, an electric field sensor for distance above ground measurement, and a bend sensor for sensing the flexing of the foot sole. Gyroscopes and accelerometers attached to the back of the heel measure angular velocity and linear acceleration, whereas sonar measures distance between the feet. Finally, a bend sensor around the heel senses dorsiflexion and plantarflexion of the ankle. A small circuit board handles the sensor data and transmits via an antenna.
Loosely related to the gait shoe are shoes from Adidas and VectraSense that include sensors and cushions (inflatable bladders) to dynamically adjust damping in the shoe during walking. This kind of product shows the potential of embedded computing in daily life products.

Kicking interfaces
Finally, several interfaces support kicking behavior. The simplest way of supporting kicking is not by tracking the leg, but rather the object that is kicked around. Examples are the previously mentioned Impact football system (Mueller '02) or the tracked ball system from Cairos technologies (Cairos '05).

Application
Obviously, the main action afforded by the feet and legs is movement (locomotion), whether it be walking, running, or cycling. As such, the feet and legs are generally used as a physical *travel* technique, either by walking around freely through a (tracked) space, by walking in place (Slater, Usoh et al. '95), or by simulating walking, via devices such as the Gaitmaster. Walking has the advantage of being the most direct

and natural way of traveling and provides vestibular cues important for wayfinding. Nonetheless, it often raises technical issues, like the possible need for long cables to connect to input devices, when wireless connections cannot be established.

Locomotion of a user can be used to perform analyses, like the biometric behavior of feet, or perform path analysis in wayfinding applications. As such, it can make wayfinding in buildings more effective, or can be used to effectively design large urban spaces (SpaceSyntax '05).

Sometimes, systems make use of foot pressure to deduct *posture* information in order to determine the (walking) posture of a person (Yin and Pai '03).

Using simple button-like controllers, one can also use feet for selection of items, up to performing simple *system contro*l actions, such as performed by LaViola et al’s iSlippers (LaViola '04). The iSlippers consist of conductive cloth patches attached to simple slippers that can generate button events via a modified Logitech Trackman Live interface. Different button events can be generated by making connections between different cloth paths, thereby resembling system control using the Fakespace pinch gloves (Fakespace '05).

Finally, a whole range of foot-controlled *exertion* applications (or games) exist. Probably best known is the Dance Dance Revolution application, available for game consoles, but also regularly played in public in game halls. Tapping on buttons embedded in a floor mat, expert users can play the game with incredible high pace, mimicking dance moves (actually directions) seen on a screen.

2.3.5 Body

Control

Full body interaction potentially combines all human body output channels, from head to feet. As such, it puts together all human output modalities as described in section 2.3.1 up to 2.3.4, added with the body torso. The output may consist of actions of separate body parts, which may also be coupled into compound tasks, up to actions that make use of the body as a whole. Hence, all joint movements identified in the introduction of section 2.3 can be performed. Most of the actions are performed in an upright pose. Due to the dependency on this pose, many possible actions will overlap with feet-controlled actions.

Hardware interfaces

Body configuration tracking

The largest group of full-body interfaces detects the body configuration of a user, using different kinds of techniques. The most widespread and easiest way is to make use of vision-based techniques to grab the body outline, possibly including the separation of main body parts. The basic method is well known from Krueger’s Videoplace application, which delivered the silhouette of the user and included this into an “artificial reality” (Krueger, Gionfriddo et al. '85).

Nowadays, most techniques make use of shadows from IR-illuminated scenes to detect so called “blobs,” determining the body shape configuration (Wren, Basu et al. '99) (Cohen and Lee '02). For a complete survey, please refer to (Gavrila '99).

Next to basic vision-based methods, multiple techniques are available that focus on the tracking of more precise body configurations. Most of these techniques apply general tracking techniques (like magnetic (Ascension '05) or optic (MotionAnalysis '05)) to detect multiple markers or sensors on the users body. Others, like the Gypsy system from Animazoo (Animazoo '05), make use of exoskeleton-like methods.

Figure 2.42: *ChairIO chair-based interface*
Courtesy of IM/VE, Universität Hamburg

Body supportive constructions
Besides the wide range of body tracking interfaces, specialized devices exist that base on some sort of body supportive construction. Any object that supports the body can be made into an interface, such as several chair-based interfaces, including (Tan, Slivovsky et al. '01) and (Beckhaus, Blom et al. '05). Tan et al’s Sensing Chair (Tan, Slivovsky et al. '01) makes use of pressure-distribution sensor sheets to detect pressure distribution for sitting posture recognition. On the other hand, Beckhaus et al’s ChairIO (Figure 2.42) allows for direct interaction, using the dynamic rotational and translational axes of the chair’s movement range for intuitive, hands-free travel.

Figure 2.43: *DigiWall climbing interface.*
Courtesy DigiWall Technology AB

Though, the idea of integrating technology for interactive purposes can by far surpass general purpose devices. Imaginable is anything which supports the human body in daily life, from digital floor tiles or walls that react to the human, up to integrating technology in sports devices that truly focus on full body activities, such as the climbing interface from DigiWall (Figure 2.43, (DigiWall '06)), which includes game-like installations with light and sound. Other more exotic examples make use of “hanging” constructions, to support the user floating in the air, including several virtual hang glider constructions (Soares, Nomura et al. '04), Humphrey at Ars Electronica (ArsElectronica '05) mechatronic installation simulating free flight, or Fels et al’s Swimming Across the Pacific (Figure 2.44, (Fels, Yohanan et al. '05)).

Figure 2.44: *Swimming across the Pacific.*
Courtesy of S. Fels

Body capacity sensing
Finally, there are several hardware techniques that make use of the distortion of electric fields caused by the body capacity of a user. The best known device, based on capacitive sensing is the Theremin. This device, actually intended as musical instrument, was developed by Leon Theremin in 1919 and makes use of an antenna to sense the proximity of a body part. Several interfaces have been developed that make use of capacitive sensors or a replication of the Theremin principle by using optical methods like a laser (Hasan, Yu et al. '02). Most of the electric field interfaces are oriented to musical installations, like the Brain Opera (BrainOpera '04), or the sensing of gestures for general purposes like navigation (Smith, White et al. '99). Some research is also performed on using electric fields for interpersonal communication and information change over so called Personal Area Networks (Zimmerman '96), but application is still limited.

Application
Full-body interaction can be seen as potentially supporting “unencumbered interaction,” interacting like we perform tasks in our daily life, not bound to any cables or devices. Environments based on vision methods indeed may support this kind of action, but most actions still seem to be bound to devices, especially when higher precision in task performance is required. Interfaces that truly make use of all human output channels (when the term “full-body interface” would be taken literally) seem too complex to currently develop, but may have great potential for the future.

The majority of full-body interfaces focus on supporting *navigation* through an environment. Techniques include simply tracking the location of a user via pressure-based methods or the user’s shadow up to interfaces that track the posture of a user. Some techniques solely focus on the torso, such as torso-directed steering or leaning, where the posture itself is analyzed to determine direction (Tollmar, Demirdjian et al.

'03). Some very specialized techniques like the previously mentioned swimming interface support navigation via sports oriented techniques (Fels, Yohanan et al. '05). Such sport oriented navigation techniques have the additional factor of being exertion-oriented (Mueller '02).

Manipulation techniques that focus on full-body input basically depend on coarse input and, therefore, only support approximate interaction. Based on contour or shadow input, these methods mostly focus on rough manipulation of objects. Most methods do not make use of information of the full-body – often, information from only the upper part of the body is used. Resembling sports-oriented interfaces, some of these interfaces are purely exertion-based and thereby largely fun-oriented. The most popular example is Sony's EyeToy device, with which a wide diversity of games can be controlled (EyeToy '05). Some manipulation tasks are used in social interaction scenarios to communicate with (virtual) characters. An example is Maes et al's ALIVE system, in which a user could play around with a dog using body language (Maes, Darell et al. '96).

Most of the *system control* interfaces that take the full body of a user into regard are rather basic. Interfaces such as "tool belts" place menus (virtual tools) at body-oriented positions, but are not regularly applied. Applications like the EyeToy make use of very simple desktop-like menu systems, in which user's body parts overlap with menu items to perform a selection. It would be rather easy to make use of the user's body configuration to issue commands – just think about the famous YMCA song in which artists formed body configurations resembling letters.

Finally, the usage of full-body output can be used for *analytical purposes*, in biomechanical studies. Application fields include anything from the analysis of gait, up to performing ergonomic studies on human loading boundaries (see (Marras '97)).

2.3.6 Biopotential

Control

Communication between sensors and effectors and the human "processing unit," the brain, generally occurs via small electric signals traveling over the nerve system. Such signals also convey information that can be "decoded" using biomedical signal processing to provide useful information on the biological systems. This information can be used to monitor a user's medical state or to trigger control actions. By using a biosensor, the decoded information extracted from the electric signals can be used to perform both voluntary and involuntary actions (see section 2.1).

Control is dependant on feedback from the human input channels. For example, when an object is grabbed, haptic and visual information is communicated that reports whether the operation is successful or failed. When humans acquire voluntary control of the physiological function by means of monitoring the electric signals, this is referred to as biofeedback or biocontrol. Without any sensory feedback, this kind of voluntary control does not work – the (bio) control feedback loop is simply different from the control feedback loop using musculoskeletal-based actions. Figure 2.45 describes how signals from the brain and effectors are used to provide human output, without using the motor behavior to control a computer input device (compare with Figure 2.1).

Some biological systems convey information that can hardly be controlled voluntarily. Such information can communicate a user's status (for example, stress level) that can be used to control parameters in an application. This kind of control is referred to as involuntary control.

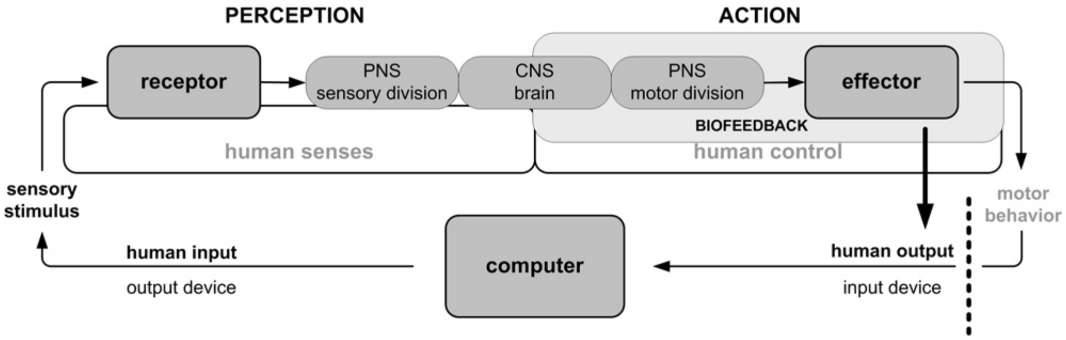


Figure 2.45: *Biofeedback observed from an information processing point of view.*

Electric signals are measured at the surface of the user’s body and are related to the ionic processes that arise from a specific group of body cells. According to (Sörnmo and Laguna '05), a cell is encapsulated by a membrane possessing permeable properties. When a cell is stimulated by a current, the membrane potential changes under effect of specific ionic substances (like sodium and chloride). The membrane potential generates a signal referred to as *action potential*. The ability of excitable cell membranes to generate action potentials causes a flow in the tissues surrounding the cells. Tissues are a conductive medium - hence the collective electrical activity of cells can be measured on the body surface. Different action potentials have different shapes, ranging from spiked waveforms of a nerve cell up to more extended waveform of a cardiac cell. However, recorded signals are often masked by noise and interference, possibly from another body process. Therefore, noise reduction and careful analysis methods are highly needed, especially when “hidden” action potentials like brain waves are sensed.

Different bio-electrical signals can be sensed, namely those in the brain, heart, muscles or body temperature. In addition, respiration (breath) is sometimes seen as biofeedback method, even though sensing does not necessarily include the measurement of electric signal. Next to obvious voluntary and involuntary actions that often deal with some kind of motor control, the electric signals can also convey information on psychological processes like stress or excitement (Wilson, Lambert et al. '99; Wang, Prendinger et al. '04; Healey and Picard '05).

Hardware interfaces

Biocontrol sensors

The core of the hardware interface is formed by a biosensor. A biosensor is a small analytical device including a biological or biologically-derived sensing element either integrated within or associated with a physicochemical transducer. The usual aim of a biosensor is to produce either discrete or continuous digital electronic signals (Turner '96). Biomedical signals are acquired by placing two or more biosensors (electrodes) on the user’s body, or close to the body (QUASAR '05). These electrodes can be connected to different kinds of hardware, namely stationary hardware using body coupled devices or portable biofeedback devices that mimic the stationary versions. Monitoring devices that are completely external to the user can be applied, but are not widely used.

Different kinds of sensors for registering different kinds of signals can be identified. Based on the nature of the transducer, biosensors include optical, electrochemical, electrical, gravimetric (mass), pyroelectric (heat), and piezoelectric (force voltage) sensors (Guiseppi-Elie '02). These sensors can be used not only for user medical diagnostics, but can also be used for other purposes, like ecological monitoring or the control of food and beverages.

Biosensors are not necessarily used to focus on a single psycho-physiological phenomenon. Most often, different sensors are combined using the same system platform. The combination of different sensors regularly leads to recognition problems, since different signals (like EOG and EEG) may disturb each other. The corpus of data may lead to new insights not possible to measure in uni-modal mode.

A specific piece of hardware deserves special merit, since it may be very practicle to wearable 3DUIs, namely the SmartShirt from Sensatex (Figure 2.46, (Sensatex '05)). The SmartShirt has been developed at Georgia Tech under the name of Wearable Motherboard and records heart rate (EKG), respiration, and body temperature through an undershirt interweaved with special fibers (WearableMotherboard '05).

Figure 2.46: *SmartShirt psycho-physiological monitoring garment. Courtesy of Georgia Tech / Sensatex*

Non-invasive and invasive systems

For usage in 3DUIs, portable biofeedback devices seem to be most appropriate to limit the amount of cables between user and computer. Both invasive and non-invasive portable devices exist. The largest group consists of compact, wireless devices connected to skin-placed sensors (Biocontrol '05; Mohsani, Najafi et al. '05; PowerLab '05; Toumaz '05). Additionally, some approaches exist that connect a PDA for wearable biophysical monitoring (g.tec '05). Many of these system are based on telemedicine principles, focusing on health care and sharing of medical knowledge over a distance using telecommunication means (Pattichis, Kyriacou et al. '02). Some more recent examples are the web-based platforms developed by Lau et al. (Lau, Churchill et al. '02) and Lamberti et al. (Lamberti, Montrucchio et al. '03).

Different kinds of invasive wireless and tethered biosensor and stimulation systems are available, most of them have currently only been probed on animals. Similar to other prostheses, like the cochlear implant, these developments have been made possible by advancements in micro electro mechanical systems (MEMS) and nanotechnology that contain a biological component (Vo-Dinh '04). Because of possible infection, costs and possible pain, wireless systems have become increasingly popular. Systems like those described by Wise et al’s (Wise, Anderson et al. '04) or Sawan’s (Sawan '04) make use of arrays of mostly silicon-based probes patched to a telemetry chip. As can be seen in Figure 2.47, this chip contains a small A/D converter, some control logic and power management (sometimes including a voltage regulator receiving power via an external power amplifier), and a transmitter to send the signals to an external receiver unit. These systems can both read biosignals and provide microstimulation. For more information on the impact of nanotechnology on implants, please refer to (Fromherz '03).

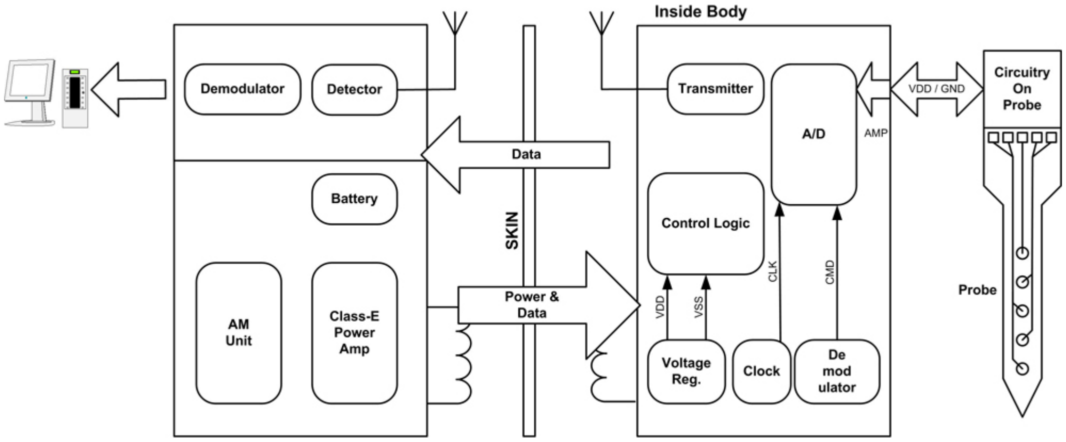


Figure 2.47: *Block diagram of wireless implantable micro system. Adapted from (Wise, Anderson et al. '04)*

Heart beat sensing

One of the most basic measurements is the registration of the *heart beat*. There are basically two methods to measure the heart beat, using an electrocardiogram (ECG) or a blood volume pulse (BVP) sensor.

The ECG detects the electrical energy of the heart by placing electrodes on the chest, arm or leg. With every beat of the heart, an impulse goes through the heart, which governs the rhythm and rate in which the ventricular muscle contracts. The same measurement can also be performed by placing sensors in the heart, which produces a so called electrogram (EG).

The BVP sensor is a photoelectric sensor that measures the reflection of light off the skin via a process called photoplethysmography. When the heart contracts, blood is forced through the vessels. The vessels swell up and, thereby, change the amount of reflected light detected by the photo sensor.

Body temperature sensing

There are two general sensors for sensing *body temperature*: thermocouples and thermistors. A thermocouple makes use of two metal wires that are welded together at the end – the wires generate a unique thermoelectric voltage that changes according to the temperature difference between the two wire ends. A thermistor is a thermally sensitive resistors consisting of semi-conductive material. This material shows a large change in resistance in proportion to a small change of temperature. Thereby, thermistors have a smaller temperature range but are more sensitive (PowerLab '05; Thermometrics '05).

Next to sensor-based recognition of body temperature, thermography can be applied, by making use of an infrared-based thermovision camera (also called a Pyrometer). Most cameras just have a low resolution, delivering a detailed temperature distribution image of about 76.000 individual measured points (in case of a 320 x 240 resolution).

Galvanic skin response

Loosely related to the measurement of temperature is the analysis of *skin conductance* via galvanic skin response (GSR). GSR makes use of chloride electrodes to measure skin gland activity and pore size. These activities reflect changes in the sympathetic nervous system and, thereby, provide an insight in the change of emotional state of a person, like the level of stress (Ark, Dryer et al. '99). An example of a GSR device is Picard and Scheirer's Affective Jewelry (Figure 2.48, (Picard and Scheirer '01)).

It should be stated that besides using GSR methods, emotion can also be detected by vision-based (facial tracking) or voice pattern analysis (Sun, Sebe et al. '04) (Busso, Deng et al. '04).

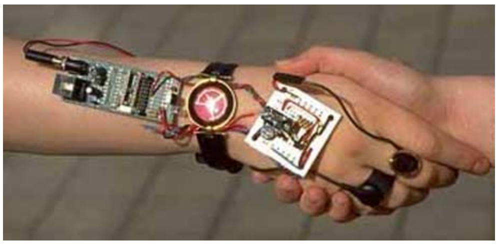


Figure 2.48: *Affective jewelry: arm-connected GSR device.*
Courtesy of MIT MediaLab

Respiration sensing

In order to sense *respiration*, several methods can be applied. These methods, as have also been discussed in section 2.3.1, include using wind sensors, thermal conductors, the analysis of the speech spectrum, the sensing of pressure, and measuring the chest by using elastic material.

Eye-based control devices

Also described in section 2.3.2, there are several methods to track the user's eye. Besides vision-based approaches, there are two approaches that make use of biosensors, namely those that generate an electroretinogram (ERG) or electrooculugram (EOG). The ERG is used to analyze the electrical potential of the retina during light simulation by placing an electrode encapsulated in a contact lens onto the user's retina. EOG is used to read the horizontal and vertical movements of the eye. Due to a voltage difference between cornea and retina, the eye movement generates bioelectric signals. By placing electrodes close to the eyes (Figure 2.49), these signals can be read. EOG and ERG based system include the MonEl2 from Metrovision (Metrovision '05), or the Cambridge Research Systems Eyesense platform (Cambridge '05).

Figure 2.49: *EagleEyes EOG-based eye tracking.*
Courtesy of Boston College

Electromyography
By using electromyography (EMG), *muscle activity* can be detected. During muscle contraction and release, electrical signals are generated by the "firing" of motor neurons, which can be measured in two ways. The easiest and most direct way is to place sensors on the skin overlying the muscle (surface EMG) – since the tissue below the skin forms a conductive medium, one does not necessarily make a connection directly with the muscles. Furthermore, intramuscular methods are available that insert needle electrodes through the skin into the muscles to directly detect the biosignals.

Figure 2.50: *EMG-based control of an airplane*
Courtesy of NASA Ames Research Center

There are several studies that specifically explored EMG for usage in computer interfaces, including VE studies with the Biomuse platform from Lusted and Knapp (Lusted and Knapp '96), Jorgensen et al's study on "joystick control" for flight simulations (Figure 2.50, (Jorgensen, Wheeler et al. '00)), or the reading of facial muscles for voiceless speech recognition by Ninjouji et al. (Ninjouji, Manabe et al.

'03). Many commercial solutions are available, including (Biocontrol '05; DelSys; Thoughttechnology '05).

Brain-computer interfaces

Finally, there are several invasive and non-invasive methods to track brain activity generated by the different processing areas in the brain (refer to Goldstein (Goldstein '02) for more details). The most widespread, non-invasive method is to make use of an electroencephalogram (EEG).

The EEG makes use of electrodes placed at the skull to detect the electric biosignals (Figure 2.51). The usage of an EEG to control computer applications is generally referred to as Brain-Computer Interface (BCI, (Wolpaw, Birbaumer et al. '02)). Next to the EEG, one can make use of Positron Emission Tomography (PET, measuring cerebral blood flow), functional Magnetic Resonance Imaging (fMRI, measuring blood oxygen level), or Magnetoencephalography (MEG, sensing the magnetic signals generated by brain tissue), but these methods are not widely used for interaction purposes. Examples of commercially available EEG interfaces are: (Cyberlink '05; MindPeak '05; Nolan '05). There are also multiple developments at universities, such as the Cerebus at Medialab Dublin (Lalor, Kelly et al. '04).

Figure 2.51: *Berlin Brain Computer Interface.*
Courtesy of K. Mueller

Several invasive brain activity measurement methods are available, fitting under the general name of electrocorticogram (ECoG). ECoG makes use of electrodes directly implanted in brain tissue to read brain activity. Till now, these methods have mostly been used on primates (Serruya, Hatsopoulos et al. '02), but several experiments have been performed on humans, including Cyberkinetics' trials with the BrainGate system (Fofonoff, Martel et al. '02; Cyberkinetics '05) and the brain implants from Dobelle Institute (Dobelle '00).

Application

The usage of biosensors to control parameters in a VE is highly diverse. Some of the application areas focus on the usage of biosignals for direct control of an application (voluntary control), but it is possible to identify a wide range of involuntary actions too. There are several ways for using biosignal-based *manipulation* actions in a VE. The most direct way is to capture muscular potential to mimic actions normally performed with a physical input device. Thus, by using myography, muscular activity can be sensed and mapped directly on movements (Ferguson and Dunlop '02; Trejo, Wheeler et al. '02). With its origin in the control of prosthetic devices (Guger, Harkam et al. '99), electrodes can be placed at different parts of the arm to detect motions for fine (hand muscles) and coarse motor (forearm) actions. Some detailed studies have focused on the relation between EMG and different kinds of grasps (Ferguson and Dunlop '02) and the movement of the arm itself (Rask, Gonzalez et al. '99).

Though predominantly developed for navigation purposes, the system developed at NASA by Jorgensen et al. (Jorgensen, Wheeler et al. '00; Wheeler '03) shows how an EMG can be used for control purposes too. Sensors are sewn in a fabric, which is placed around the wrist to sense muscular activity, resembling the Sensitive Skin construction of Lumelsky et al.(Lumelsky, Shur et al. '01). In the system, the hand mimics the movement normally performed with the joystick and, as such, the interpreted signals are being used. In the presented application, NASA showed how a flight simulator (Boeing 757) can be controlled via an EMG. It can be easily imagined that using the same principle, objects can be moved through space.

Next to using an EMG, experiments with tapping brain signals have shown to be successful for very basic tasks. Systems like the Berlin Brain-Computer Interface (Krepki, Blankertz et al. '03) support *1D and 2D-control* tasks like the movement of a cursor over a screen. A BCI detects the potential related to the movement preparation, or even the motor imagination, and maps this using a variety of analysis methods to a control output (Trejo, Wheeler et al. '02; Krepki, Blankertz et al. '03). The basic principle behind this interface could be mapped to 2D oriented manipulation techniques used in an immersive 3DUI, like the Pierce image plane interaction techniques (Pierce, Forsberg et al. '97).

It should be mentioned that two different kinds of BCI's exist: those dependant on muscular activity (like gaze, via visual evoked potentials) and those that make use of oscillatory brain activity (Wolpaw, Birbaumer et al. '02). Either of them is rather easily recorded using inexpensive equipment and cheap, or even open source, software (OpenEEG '05). Unfortunately, they deliver a low level of performance: tasks can currently only be performed at slow speeds.

As can be derived from the previous section, biosignals can also be used for *navigation* actions and related studies. EMG-tracked motor behavior mimicking joystick control, like demonstrated by Jorgensen et al (Jorgensen, Wheeler et al. '00), can be well mapped to a movement control. Similar to cursor control, EEG data can be used for simple steering tasks. Even though it is currently only suitable for simple navigation (steering either left or right), several tests have successfully shown brain-controlled based on imaginary movement control (Bayliss and Ballard '00; Leeb, Scherer et al. '04). Next to the performance of travel tasks, EEG can also be used for *analyzing brain activity during navigation* (Ekstrom, Caplan et al. '05).

There are several ways of using biosignals for providing *symbolic input*. A whole range of mostly EOG-based eye-typing systems exist, in which the user's gaze is matched to using a virtual keyboard, including EagleEyes (EagleEyes '05) and EyeWriter (Lileg, Wiesspeiner et al. '99). These systems are similar to other gaze-directed typing applications that do not use EOG – see the Cogain website for a complete overview (Cogain '05). Next to EOG-based input, one can also make use of EMG-based approaches that track fine motor actions, as described in (Wheeler and Jorgensen '03) to perform symbolic input.

Being medical monitoring devices, biofeedback hardware has regularly been used for setting up analysis or supportive-based computer systems, for other than general monitoring purposes. Examples include life support using EOG-input (Kato, Yamaki et al. '02), sleep analysis systems (Bieliková '02; Biosomnia '05), and the treatment of disorders like attention deficit disorder or epilepsy. For an overview of different disorders treated via biofeedback methods, please refer to (EEGSpectrum '05). Several studies also focus on making use of VEs, including the aid of physically disabled people (Lusted and Knapp '96) or reducing chronic pain (Steffin '05). Furthermore, several studies have focused on stress, including Healey and Picard's study on stress during driving (Healey and Picard '05).

Related to the medical issues, several monitoring approaches have been used to observe the users mental processes related to attention. The most general approach is to make use of eye gaze to detect different attention factors like spatial attention, scene recognition, and memory related issues (Richardson and Spivey in press (b)). Other tests have used EEG recordings to detect mental load patterns and to adapt applications accordingly (Wilson, Lambert et al. '99; Chen and Vertegaal '04).

Figure 2.52: *User playing the Journey to Wild Devine game.*
Courtesy of Wild Divine

As has been discussed in the control part of this section, biosignals can also be used to *detect the emotional state of a user*. The state of the user has found its way frequently into different games. The games actively make use of the biofeedback loop to generate game content dynamically with respect to the affective state of the player with goals like excitement or relaxation. However, it is difficult to determine whether the state of arousal is positive or negative – some players might enjoy stressful games whereas others do not (Sykes and Brown '03). Also, there are a huge number of emotions (Morgan and Heise '88), so that the state detected is not always correct. These kinds of games are generally known as affective games (Gilleade, Dix et al. '05), derived from the term affective computing (Picard '97). Examples of affective games are Journey to the Wild Divine (WildDivine '05) and Relax-to-Win (Bersak, McDarby et al. '01).

The Journey to the Wild Divine (Figure 2.52) is a commercially available package with a console (the "Lightstone") consisting of a handful of GSR sensors that should be placed around the finger tips. The game is an adventure specifically focused on solving small tasks that will possibly help the user to relax or to create excitement (by challenging the user). Personal observation has shown that the game-mechanisms takes quite some time to get used to and are not always very effective, but users regularly seem to be interested or fascinated by the game play style.

Biosignals have regularly been used for musical expressions, using EEG, EMG or BVP measurements. A range of experiments have been performed by Atau Tanaka, using a wireless gestural EMG-based interface (Dubost and Tanaka '02). An experiment called REGEN (regenerative music) focused on using different biosignals like heart beat, respiration and brain waves to create musical patterns. Hereby, it was also tried to make use of collective readings. In the DECONcert, 48 people were monitored simultaneously to generate a collective soundscape (REGEN '05). Finally, in the 2Hearts system, users wear BVP sensors and hear music that is linked to their heartbeats, thereby changing in harmony, rhythm, and tone as heart rates change over time (McCaig and Fels '02).

2.4 Human Behavior

Throughout this chapter, issues have been handled that do not necessarily fit within a certain input or output channel category, but nevertheless affect the I/O channels as a whole. This section provides a short discourse on some of the higher level factors that affect the human I/O loop.

This section deals with behavioral issues that adapt the way a user performs an action or that affect the user’s understanding of feedback. These issues generally are identified when performing a task analysis (Luczak '97). Issues like skill, training, intelligence, age, and sex of a user define how functionality is designed (Kieras '97). It is necessary to create a cognitive (mental) model of a user of the system in order to adaptively react when a user is interacting with a system.

Hence, it is important to take a deeper look at the behavioral aspects of a human being and its effects on user interaction. Looking from a behaviorist (or also cybernetic (Wiener '48)) perspective, action sequences are defined by five major components. Except spiritual processes, all seem to apply to interaction processes (Huitt '03):

- Cognitive: perception, storage, processing and retrieval of information
- Affective: emotional component that affects perception and thoughts before and after they are processed
- Conative: management of input and output
- Behavioral: output of the user

The model (Figure 2.53) shows that it is important to observe human output as a process, in which conation binds knowledge and affect. Conation is seen as consisting of both covert (user defined action) and overt (the controlling of environment parameters that have an impact on the user’s action) components. Hence, next to the user herself, the (social) environment, ranging from the direct environment around the user up to society, needs to be included as parameter when analyzing task performance. The environment itself can also be defined as regulatory element that can take affect on a user, without covert actions taking place. For the purpose of separating perception and memory, perception is seen as separate mechanism beside cognition. Furthermore, the human capabilities have been added, forming the backbone of the human I/O loop and, as such, also define their action parameters. The model is highly abstracted, though: for a deeper reflection on behaviorist factors, please refer to (Huitt '99).

The key to providing an “intelligent” user interface does, therefore, not only lie in the analysis and usage of human capabilities (like skill, intelligence) and how they change dynamically (including attention and memory activity), but also on the analysis of the hidden mental state of a user that defines conation. Secondly, for a 3DUI it is important to take the environmental effects into account, being either the simple natural environment parameters or a social structure, like another person in the same (work) environment. To understand the effects of human behavior on the performance of actions, perceptual, cognitive, affective, and motor information needs to be captured and reflected on within the framework of the capabilities of the user. It should be clear that there is a strong interplay between the different mechanisms and the human capabilities – the human potential is basically formed by the different properties.

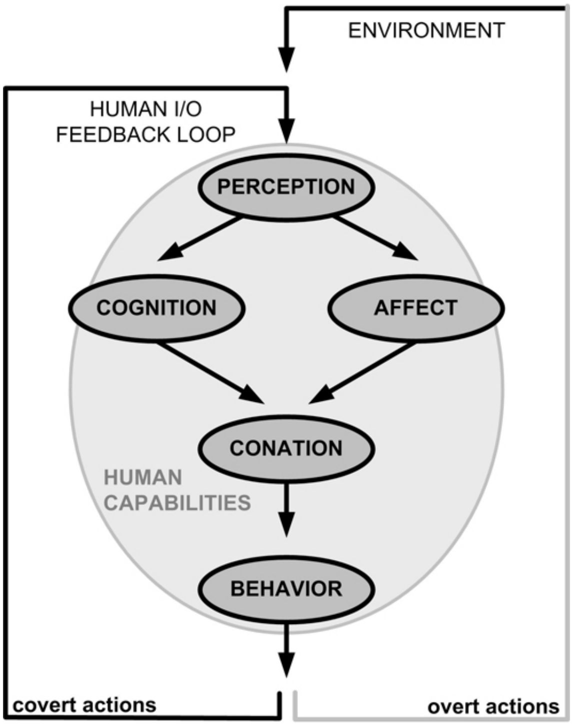


Figure 2.53: *Abstract human behavior diagram in relation to human capabilities and interaction. Adapted and extended from (Huitt '99)*

The different parameters can be used at different levels and different combinations. Developments, like affective environments, can sometimes make use of a single kind of data, whereas truly adaptive interfaces will require full integration of cognitive, perceptual, affective, and motor information for a task reasoning architecture. Examples of architectures that integrate the different kinds of information are Duric et al’s system that focuses on more general adaptive system mechanisms (Duric, Gray et al. '02) or Wood’s behavior-based architecture for military command and control (Wood '04). Other useful areas are the reduction of mental load, especially in safety critical areas or the true support for habit-based interaction.

Observation of single mechanisms has been probed in multiple test environments and are already partly handled in the previous sections. They include (Duric, Gray et al. '02):

- The usage of eye tracking to analyze the user’s attention and focus or cognitive overload
- The monitoring of different brain areas to sense cognitive activities

- The analysis of motor activity in user interaction, like the observation of key strokes (Gunetti and Picardi '05)
- The usage of facial expressions (Pantic and Rothkrantz '00), speech patterns, skin responses (GSR) and (voluntary and involuntary) motions (Bobick '97) or postures to predict the user's emotional state

The different types of information can be used in direct and indirect ways. They can be used both as way of monitoring and subsequently changing the internal state of the interaction mechanism or state of the computational environment, or to make use of the data to allow the user to control an application. This might be best seen in the field of affective computing. Here, it can take the form of computer-controlled influencing of the user's mental (emotional) state (i.e. human input) for example in ways as being done in storytelling environments. A hierarchical finite state machine can be made, in which emotional states with triggers can be traversed, depending on what the user is feeling. So, if a user is bored, she would receive specific emotional triggers. On the other hand, affect can be used to control an application directly, like is shown in several examples of emotional control of virtual characters or toys. Examples include the sympathetic interface (Johnson, Wilson et al. '99) and SenToy (Paiva, Prada et al. '03) for avatar control, or the Affective Tigger, an emotionally reactive toy (Kirsch '99).
Looking from system perspective, a field that has been taking care of adaptively changing interaction by interpretation of the user's mental state is intelligent user interfaces (IUI) (TechnologyWatch '05). In the IUI field, much effort is spent on the effect of decision-making processes on interaction (Stephanidis, Karagiannidis et al. '97; Jameson, Grossmann-Hutter et al. '01). As one example, it has been shown that interaction mechanisms can be habit-forming: users tend to consciously or unconsciously perform actions and sub-actions in a specific order (Koritzinsky '89). An IUI can take over by changing user input dynamically or to present information to a user in a more direct way. Nonetheless, an IUI often focuses on automation of user interaction by partly taking over at the command-level. For a 3DUI, this is not always appropriate, since automation of user interactions is basically the opposite of direct manipulation, a principle on which most 3DUIs are built upon. Hence, a more complete model of the user needs to be defined, bringing together multiple research fields.
The effects of the environment on the user's actions are partly easy to capture. If we observe the environment as a social structure, the influence of further persons is clearly definably. For example, sequences of actions can be viewed as the product of different human I/O channels, as is the case in collaborative environments. Furthermore, the effects of the natural environment on the perception of the whole range of possible setups in the mixed reality continuum can be clearly observed, although not frequently made use of. For instance, think about an augmented reality display that analyses the light conditions of the natural environment, thereby automatically adapting the display levels (like brightness and contrast) to achieve the best perceptual result for a user.

To conclude, behavioral factors are still an open and not completely understood issue, which needs to further explored. Some statements can be made, though:

- *Manage cognitive capabilities*: identify and deal with cognitive issues, such as overload to avoid decrease of performance, for example by changing (improving) feedback.

- *Affect may matter:* currently still largely unexplored, the emotional state of a user may provide useful information on the mental condition. It may be used with the purpose of changing their emotional state as a subjective effect (such as making computer games more fun), or to see if the emotional state is caused by cognitive factors, such as overload (like a stressful user) that may be improved to increase performance.

- *Support user-specific behavior*: as a result of cognitive and affective factors, a user may have specific problem solving capabilities (habits) that can be made use of to improve the effectivity or ease of a user interface

- *Do not overlook environmental issues:* social (other users) and indirect environment effects, such as noise or blending lights, can have a clear effect on the demands of an interface, both on the software and hardware side.

Additional research will hopefully further clarify the different factors and interdependencies involved.

2.5 Summary

In this chapter, the human I/O system has been described, consisting of a basic psycho-physiological description of the input and output channels, examples of unconventional hardware, and new technological directions, coupled to application possibilities and tendencies. Additionally, a small discourse on higher level interaction principles has been provided, focusing on behavioral factors affecting the process of interaction.

Dominating sensory systems gain most research attention and are well understood. Vision and audition make up the majority of perceptual resources and have been / are far more focused on in research than the other sensory channels. Nonetheless, unconventional directions such as invasive technology still allow for a wide portfolio of future research for spatial interfaces.

Haptic research is expanding and rather well understood, whereas smell and taste interfaces are still far behind and holding many unsolved problems. Interest in haptics is expanding, with results increasingly finding their way into commercial products. Smell and taste interfaces are highly complex and only more recently gaining some attention. Both fields still hold many unsolved research problems, such as the generation of correct chemical sensation and the separation of different feedback sensations. All three fields can be regarded as (highly) experimental and bring forth a range of unconventional interfaces and applications.

Research on vestibular feedback is especially important for navigation. The vestibular system, sometimes regarded as a sensory system at its own, has a large impact on navigation techniques. It can provide real-motion cues important for spatial understanding and supports the reduction of motion sickness.

Hand-oriented output still dominates. Predominantly because of its fine-grain action possibilities, most actions are still performed with hand-based interaction devices.

Related to precision, the hand also holds the largest sensory-motor distribution of the cortex and, therefore, is also well suited to control.

Non hand-based output is often a result of control substitution. Human output based on another output channel than the hand is often a result from control substitution. In this process, an action normally performed by the hand is mapped on another body part. A field that has greatly supported control substitution is assistive technology for the disabled.

Biopotential interfaces are rather complex, but allow for a wide variety of new kinds of interfaces. Sometimes also a result of applying control substitution, biopotential interfaces allow for the performance of basic interaction tasks, and are especially usable when other control channels are blocked or unavailable because of disability. Whereas eye-based techniques have a long and successful history, other techniques such as those based on brain or muscular activities have many open research issues (including hardware demands or filtering of data) and currently only allow for the control of very simple actions. Biopotential interfaces change the human interaction loop because of lacking direct motor behavior, which especially affects feedback mechanisms, and also lead to currently not well known interaction effects.

Full-body interfaces are hard to achieve. Even though all human I/O channels can be connected to, a full body interface is hardly possible to achieve. The quality of interfaces is often low, and wearability is often not guaranteed. Currently, full body interfaces are mostly concerned with supporting control actions performed by different body parts in a single interface.

Behavior adapts the way a user performs an action or understands feedback. From a behaviorist perspective, user behavior is affected by cognitive and affective factors that are bound by a process called conation (management of input and output). Conation is seen as consisting of both covert (user defined action) and overt (the controlling of environment parameters) actions. A better understanding of factors involved may lead to more “intelligent” interfaces.

Following the discussion on human potential, the next chapter focuses on the design and development of interfaces, by looking at the idea-finding, design, development, and evaluation of unconventional interfaces.

CHAPTER 3
Designing unconventional interfaces

This chapter deals with the design and development of unconventional human input and output methods. As can be seen in Figure 3.1, the different sections of this chapter are ordered according to a basic development process. The chapter will start with an overview of related fields of research, followed by an overview of human potential analysis factors. Both sections aid the idea-finding process. Subsequently, a short discourse on a main design method for developing unconventional interfaces is provided, focusing on sensory and control substitution. Guided by the design methodology, an overview of interface factors is given that affect the final system specification, focusing predominantly on interaction flow and feedback. The next stage is comprised of two sections focusing on development and integration of unconventional interfaces and dealing with application and transfer factors. Included are some experimental methods for creating hardware, called garage interface design. The chapter will be concluded by some pointers on evaluation techniques.

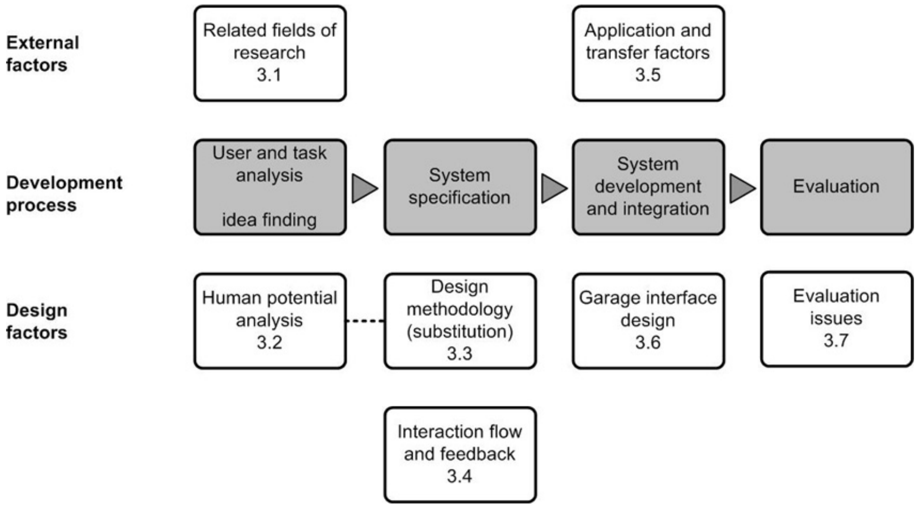


Figure 3.1: *Overview of different sections in chapter 3.*

This chapter deals foremost with special factors that are specific to the design of unconventional techniques. For information on the following topics, it is recommended that the reader refers to other resources:

- Standard user and task analysis (requirement analysis) methods and development methodologies. For this purpose, a source like Luczak (Luczak '97) can be used, as well as general HCI literature.

- General human-interaction factors, unless specifically applicable to unconventional interfaces. Such factors can be found in sources like (Shneiderman '98; Preece '02; Bowman, Kruijff et al. '05).
- Software toolkits in which interaction techniques can be built (Just, Bierbaum et al. '98; Tramberend '01; Schmalstieg, Fuhrmann et al. '02)
- General usage factors of unconventional interfaces, since they are largely the same as general mixed reality interfaces. An overview of factors can be found in chapter 10 of (Bowman, Kruijff et al. '05).

3.1 Related fields of research

There is no single best formula for designing and developing unconventional interfaces. A huge number of factors may initiate the design of a new interface, like a new piece of hardware, or the crazy idea of someone drinking a beer in the bar.

To start with, it is recommended to get a better idea on idea finding processes, for example by looking into specific methodologies such as design research (DRS '06; JDR '06). Design research is largely based on an idea evolving out of trying to understand the behavior of specific phenomena. Emerging theories can be formed that can be further focused on. The theory slightly collides with the design of unconventional interfaces, since most of the time, it is oriented towards problem-solving: Therefore, one must be aware of a specific problem. Even though many unconventional interfaces may address a specific problem, it may not always be the starting point of design. It often occurs that there may just be an idea, without understanding its exact purpose, or at most some kind of fuzzy problem. For example, as can be seen in section 3.2, a specific idea can be formed out of a possibility of a human I/O channel, without knowing its exact purpose.

Taking one step back from problem solving, there seem to be two main approaches that can be identified that focus on the overall process, namely the artistic or the scientific approach (chapter 10 in (Bowman, Kruijff et al. '05)). The artistic approach is based on intuition about users, tasks, or environments and more or less common-sense oriented. Aesthetics or simply “being different” may fuel the idea of a new interface, whether it be the adaptation of an old one or the creation of a completely new breed. On the other hand, the scientific approach is based on formal design, by performing strict user and task requirement analyses and evaluation. The two design paradigms may seem competitive, but can also be used complementary – for example, an idea coming out of some hunch can be worked out formally. However, designing a 3DUI is still rather difficult. In comparison to desktop environments, basics like formalized guidelines or extensive toolkits largely fail. Hence, the process of developing an unconventional 3DUI will often be highly experimental up till the stage that the goal is reached – be it performance (scientific approach) or excitement and beauty in an artistic development. The starting points, though, of both directions may be the same: they can both be human-driven or device-driven.

When designing a new, possibly unconventional interface, it is of great importance to observe related research directions, as provided by Figure 3.2. Looking at other fields can save time and limits potential failures when designing a new interface.

The most obvious related field of research is of course 3DUIs (Bowman, Kruijff et al. '05), applied at the full range of mixed reality applications. Partly a subfield of 3DUIs, many graspable (Fitzmaurice, Ishii et al. '95) and tangible interfaces (Ishii and Ullmer

'97) have been unconventional interfaces themselves or form a great inspiration for new interfaces.

There are numerous developments that are based on the initial thought of embedded computing (Alonso, Blair-Smith et al. '63). Embedded systems generally focus on the hiding of circuitry in everyday objects. This is already long done in most household appliances standing around us. One can basically put a sensor or actuator in every kind of object, in order to sense user input or create output or simply to track its location. Using the same principle, Weiser (Weiser '91) introduced ubiquitous computing (overlapping with the field called pervasive computing). He focused on making multiple computers available in the physical environment around the user, but making them effectively invisible. This view contrasts with the classical immersive-VR view, in which people are "encapsulated" in a digital world, but with the many variations of mixed reality installations. Information placed inside the direct environment of a user match well to a 3DUI of many current developments. Other directions, like ambient intelligence (Ambience '05; MITOxygen '05) make use of the same principles, and are highly sensor-oriented. Fitting inbetween, the field of wearable computing (Mann '96) makes use of small computers, either handheld or built in everyday objects (like clothes), and highly affects wearable AR research. Several directions exist that mix some of the developments mentioned, for example context-aware or embodied interaction, which combine ubiquitous and tangible interaction (Dourish '01).

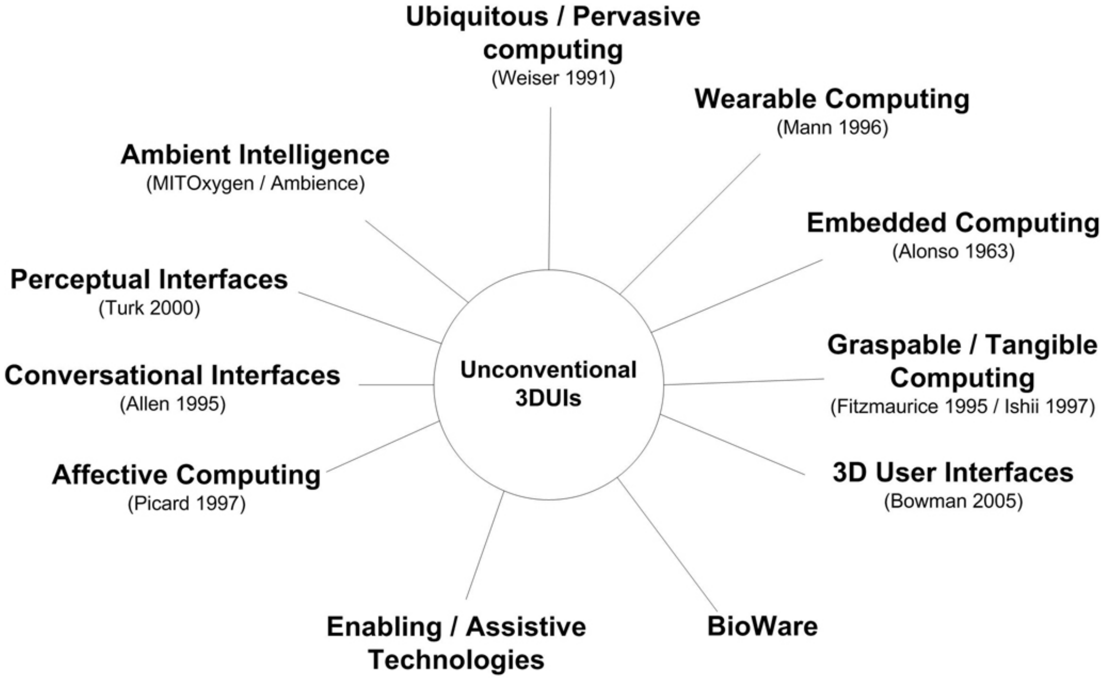


Figure 3.2: *Related fields of research.*

Some developments that apply at a wide variety of UIs can also be used well in unconventional 3DUIs. Examples include: perceptual interfaces (monitoring a user's behavior via machine perception and reasoning, as handled in section 2.3.2. (Turk and

Robertson '00)) and conversational interfaces (Allen '95). Both directions are based upon conversational metaphors of communication.

Finally, there are several biophysical or psychological oriented directions that influence the development of unconventional 3DUIs, namely affective computing (see section 2.3.6 and 2.4, (Picard '97)) and the fields of enabling and assistive technologies, that partly overlap with the field of Cyberware. The term Cyberware is mostly known from science fiction books and games, and generally (and critically) associated with "Cyborgs". This field may seem unscientific, but a large amount of research in biotechnology, more specifically in biomechatronics, makes this a real-life topic. This can be seen on the large number of implants in the previous chapters, and specific examples of "Cyborg" projects. An example is the project performed by Warwick et al. at the University of Reading (Cyborg '05), in which technology (a micro array tube sensing nerve fibers) was implemented in a person (Warwick himself) who had no physical deficiencies.

3.2 Human potential analysis

In this dissertation, the analysis of human potential (the focus of this section) forms the main starting point to design a new and possibly unconventional technique or device. This section will start with a small discourse that clarifies the difference between multisensory and multimodal processing, since some sensory factors can greatly affect how an interface functions (3.2.1). After this introduction, a discussion will be presented on the limits and possibilities of human potential (3.2.2), which will be followed by a model that explains different levels of human potential (3.2.3). Finally, the last section presents a basic classification of unconventional techniques (3.2.4).

3.2.1 Multisensory versus multimodal processing

One of the keys to developing unconventional interfaces promoted in this dissertation is the analysis of human potential. Predominantly a human-oriented direction, analyzing the human input and output channels is a very good basis to get an idea of a potentially new and useful way of interaction. When doing this, the first important step is to get rid of the traditional multimodal way of looking at interaction.

When focusing on the combination of multiple sensory modalities, the general conception of how modalities are handled is the "multimodal" point of view. Found throughout all general human-computer interface literature like Preece (Preece '02) or Shneiderman (Shneiderman '98), different modalities are seen as separate modalities. These modalities can be coupled, but were still to be handled as separate processes (Shimojo and Shams '01). Even though multimodal interfaces are mostly concerned with the users' output to a system (combining more than one output modality), it is more important to understand that the perceptual (human input) systems are generally seen as distinct units that do not affect each other.

Multiple studies have indicated that this view is incorrect (Shimojo and Shams '01) (Pai '03). Multisensory processing, in which sensory modality can affect each another, is proven to be valid and occurring more often than is regularly believed. The processing theory builds upon the integration of sensory signals inside of "multimodal association areas" within the brain. Several of the most common phenomena that support the multisensory processing theory are the, so-called, ventriloquist effect (spatial location of sound and visuals are correlated, like on TV when someone speaks) or the McGurk effect, in which vision alters speech perception (Pai '03).

The research on multisensory factors still needs to advance in order to fully understand its importance. Nevertheless, some effects can be identified from the previously mentioned studies. These effects can be labeled as "cross-modal effects" and have the following characteristics:

- *Cross-modal bias*: stimuli from two or more sensory systems can differ and affect each other leading to modified or even incorrect perception
- *Cross-modal enrichment*: a stimulus from one sensory system can enhance the perception of another sensory system
- *Cross-modal transfer*: stimulation of one sensory system may trigger perception in another system

A deeper discussion on these multisensory factors is provided in section 3.3, in which the issue of sensory and control substitution is handled.

3.2.2 Limits and possibilities of human potential

Before starting the analysis of human potential, a first question to ask is: why do I need to make a new, possibly unconventional interface? Reasons may range from increasing the feeling of "usability," when a technique performs well (or better), to the element of surprise, when human sensory systems are tickled. A wide variety of unconventional interfaces can be imagined, with a range of different goals.

When thinking about unconventional techniques from a human potential oriented direction, a distinction needs to be made between so called "magic" and "natural" techniques. The two technique categories provide different ways of designing and developing interaction techniques, but both have the same starting point: human potential. The main dissimilarity is that magic interaction techniques do not mimic normal human performance – they enable actions that are impossible through real-world physics. Designing magic techniques essentially makes use of an algorithm or device to *amplify* or "trick out" human potential. A well known example of a magic interaction technique is the Go-Go interaction technique, which allows a user to select and manipulate far away objects by stretching the virtual arm to unnatural proportions, using non-linear mapping of arm movements (Poupyrev, Billinghurst et al. '96a).

Whatever goal intended and independent of either natural or magic directions, before actually designing a new interface, some questions need to be asked.

- *Does one really need an unconventional technique?* Unconventional techniques should not always be used for the sake of unconventionalism or innovation. For many task situations, the usage of general hand output and visual feedback works perfectly well. Using alternative hand output or visual input techniques or techniques that make use of other I/O channels can make interaction particularly hard. The type of task or behavior, its function, and the sequence of activity might not fit the application of an unconventional technique. The technique might require specific skills of a user and specific hardware that cannot always be applied. Furthermore, the nature of the task might be unnecessarily hardened, by increasing the amount of actions or its complexity, leading to problems like increased error, or even hazardous usage. So, at the end, performance counts. Depending on the goal of the application, this may range from speed and accuracy up to safety or fun and needs to be matched well to the used technique.

- *Which psychological and physiological limits need to be regarded?* Consider the limits of the human potential and to what extent it makes sense to go up to this limit. This question is a basic task analysis issue, defining which task characteristics need to be satisfied, and is constrained by human and environment boundaries. In some cases, unconventional techniques can even go "beyond" human potential to allow for specific task control. For example, think about using digital magnifying lenses to allow for detailed visuals in order to control a very fine-grain task.

- *To which level can the human body be extended by devices*? The potential of a human being can be extended through artificial aids. Devices can be used that basically read human action (normal input devices), but a range of devices exist that truly expand the abilities of the human being by restoring potential (implants) or by extending their capabilities to artificial levels. The latter can also be achieved by implants, or by specific devices like a robot arm that might, for example, produce a higher force on a natural object than would ever be possible by a human. Such extensions always have a direct linkage to the actual human potential. Hence, the psycho-physical limits stated in the previous point also count, next to other factors ranging from workplace criteria (health inflicting rules), social, ethical and cultural, up to political issues.

- *Does amplification affect other sensory systems*? Amplifying human potential often also has its limits, which is regularly caused by another sensory system than the one that is amplified. For example, think about the Go-Go interaction technique. A user cannot endlessly extend the virtual arm to manipulate objects, since the human visual system will not be able to register changes anymore when the hand is too far away. Next to psycho-physical limits, the before mentioned cross-modal effects should be considered, in which the perception of a sensory system is modified by another system.

Using sensory and motor capabilities to their full extent, or even surpassing them, is directly related to the core discussion on the "ultimate" goal of Virtual Reality. Going back to one of the earlier articles defining Virtual Reality and its interfaces, Steuer explored the dimensions of experiencing Virtual Environments (Steuer '92). Even though the presented view has changed slightly due to the wider focus on mixed reality, Steuer explores several issues that are of great importance when investigating interfaces that make use of the potential of a human being.

Steuer defined the level of experience in a Virtual Environment (level of presence) as being composed by two components: *vividness* and *interactivity*. Steuer described vividness as:

> *Vividness means the representational richness of a mediated environment as defined by its formal features, that is, the way in which an environment presents information to the senses.*

Vividness is composed of breath (a function of the ability of a communication medium to present information across the senses) and depth (depth of the sensory information available in each perceptual channel) and basically refers to applying human input channels to their full potential. Interactivity is defined by the speed with which a user can interact with an environment, the range of different kinds of actions, and the possibilities with which the system can dynamically map the actions to controls. Whereas both vividness and interactivity are just basic factors that refer to HCI issues, the combination of both leads to a field of research that by far surpasses the general field of multimodal interfaces.

Even though vivid and highly interactive techniques *(vivid interaction)* have been envisioned from the start of VR developments, most applications are still stuck in the multimodal corner, focusing predominantly at the combination of visual and auditory system output and mostly just hand-based input, without haptics. Even though it has been mentioned over and over again, the simulation of all human input and output channels *at once* or at least within a single application has not been reached yet.

The development of vivid interaction techniques comes close to what some researchers have called a full-body interface (see section 2.3.5 (Tollmar, Demirdjian et al. '03)). The full body interface predominantly focuses on human motion: how can a user make use of all body parts to control an application. On a higher level, though, it is much more interesting to couple the full body interface view with the vividness and interactivity point of view: to make use of potentially all human input and output channels at once, simultaneously or in a serial way. Combining the vivid interaction and full body interface views, several statements can be made that are rather straightforward. Both directions focus on designing techniques that potentially make use of the full sensory and motor capabilities of a user. Hereby, the focus is moving away from traditional multimodal techniques in the direction of multisensory interfaces that differ at the level of human information processing. During the design of full-body interfaces, one of the predominant approaches followed is sensory and control substitution, handled in detail in section 3.3.

3.2.3 Four stage analysis model

When taking a closer look at human potential, four different levels of analysis need to be taken into account. Following the model from Gopher and Sanders (Luczak '97), task variables, processing stages, energetical mechanisms (mechanisms that foremost focus on the effort to plan and perform a task), and cognitive resources can be identified. Figure 3.3 provides an overview of the different levels and is built upon a similar information processing viewpoint as introduced in chapter 2. The model provides a clear separation between processing and energetical mechanisms, important when developing unconventional techniques. The top stage, evaluation, provides the meta-level mechanisms grounded in cognitive processing. The interplay between arousal, effort and activation relate to attention-controlling mechanisms. All together, the model provides an overview of efficiency of performance oriented at the level of energy that can be allocated for a task.

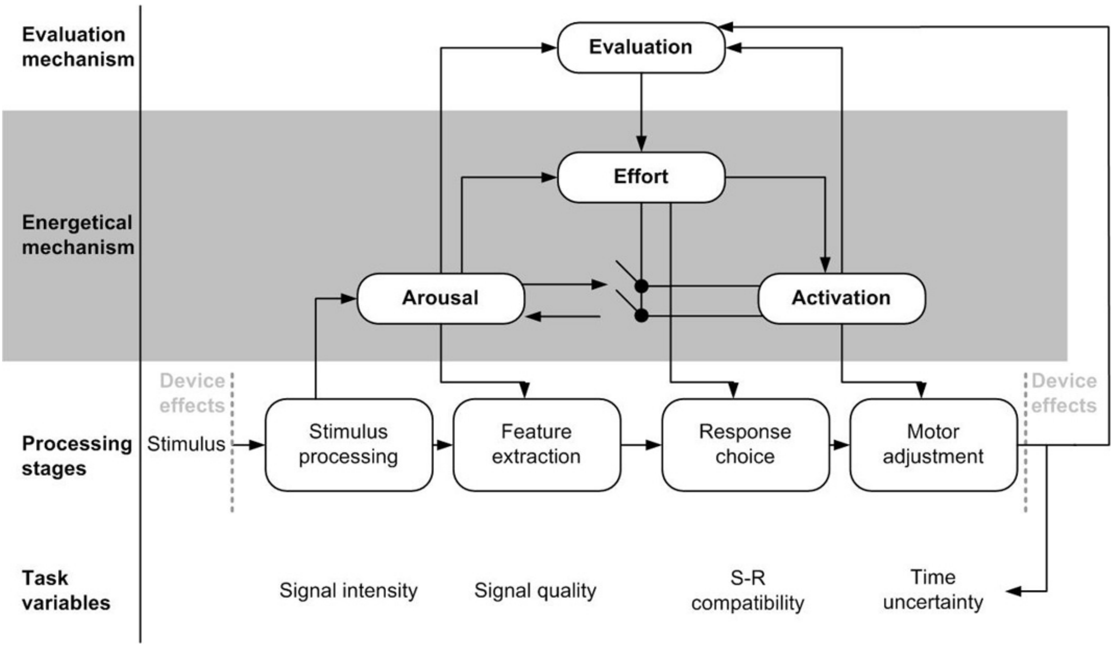


Figure 3.3: *Cognitive energetic linear stage model*
Adapted and extended from Gopher and Sanders (Luczak '97)

The energetic model can be coupled to the capabilities and, therefore, also limitations, of a user. Furthermore, it provides a way of interpreting and weighting possible effects of sensory or motor substitution (section 3.3).

Starting with the analysis of the potential of stimuli, the model provides several insights into the potential of using different sensory systems. A stimulus can trigger any of the body receptors and is defined by both the intensity and the quality of information it can provide. Here lies one of the first keys of human-driven interaction techniques: which stimulus can deliver which kind of information in which quality? Or, better said, from which receptor can features be extracted to deduce information to perform a task, and in which intensity do they need to be provided? Hence, to which extent do device factors affect the perception? Both the informational quality and the energetic effectivity can be deduced by comparing different sensory systems that provide similar amount and quality of information, through substitution methods. The effort needed to perceive the stimulus and subsequently trigger an appropriate output action can be investigated, after which conclusions should be made on its suitability. Hereby, sensory blocking (for example, an auditory channel could be blocked in a loud environment) or impairment plays an important factor for coming to the correct conclusion. When a sensory channel is useless in a specific task-user-environment setting, an effort comparison becomes obsolete.

A further issue is the maximum level of the stimulus intensity, depending again on user, task and environment, in which the limitations of the user come into play once more. When the intensity of the stimulus can not be matched by the user’s capabilities in order to extract the right amount of information, it may be unusable for the task at hand. Thus, in order to create unconventional output to a user, the information quantity and quality needs to be matched by the perceptual system, avoiding possible overload at the cognitive (evaluation) level.

When analyzing human output, similar factors need to be dealt within the perceptual side of the model. To create a suitable response and motor action, a user needs to spend a specific amount of effort. The created motor action is directly affected by the perceived stimulus and creates a closed action-feedback loop that needs to match the task at hand. Different output methods, in their dependency to a coupled stimulus, can be compared to derive a performance oriented model of task performance. Observing the action-feedback loop from an energetic point of view provides detailed clues on speed and accuracy and their related cognitive and motor load. These models can be directly related to performance studies applying Fitts' Law (Fitts '54).

A second issue that comes into play is the motor system-task compatibility and the control structure and ergonomic changes when exchanging motor systems to perform a specific action (control substitution). The model provides a direct view into the effort needed to perform the task with the specified motor system, including possible speed and accuracy effects, as mentioned before. Hereby, ergonomic considerations need to be taken into account in order to reduce physical stress on the user's body. Furthermore, a close look needs to be taken at any effects on posture when exchanging the motor system. When the task is performed using a different body part, as a result of the changed control-body linkage (labeled "device effects" in Figure 3.3), the biomechanical configuration of task performance changes. Thereby, of course the device coupling is an important one: which kind of movements does the device support with how much effort? All together, this ultimately feeds back to human potential – to what extent can the physiological system be used for the task at hand? Such an analysis goes through a multistage investigation, starting with the anthropometric characteristics of the human body, through force-related biomechanical values, up to the psycho-physiological limits of the user. Clearly, regulations and standards affect all levels by defining rules for specific task-environments, rules that limit psycho-physiological costs of exchanging motor systems.

A particularly interesting point that can be retrieved from the energetic model is the variety of ways of stimulating a user and retrieving signals for output purposes. Looking at the Figure in section 2.2.6, biofeedback mechanisms may trigger and retrieve information at neural level, or the brain. Clearly, the energetic action-feedback loop changes by using biofeedback systems, though, the premise of the model still is valid: how much effort needs to be spent? For example, the usage of a brain-computer interface may be ergonomically apt, since it puts only little extra force on the biomechanical system to perform an action by just needing to support the cables coming from the head. However, it may require great effort at the cognitive level, as experiments have shown (Krepki, Blankertz et al. '03). Thus, the stages of processing information change, and specific psycho-physiological limitations become obsolete.

Reflecting on the creation of multisensory interfaces, the cognitive energetic stage model provides a great aid in analyzing effectiveness of combining techniques by ways of addition or integration (see section 3.3). This may predominantly occur at the level of analyzing the perceptual or motor capabilities mentioned before, but there are several issues that play a role in the more cognitive-oriented levels. One such issue is decoupling, in which an additional input channel is used that differs from the main interaction channel, for example to provide feedback. Sharing capacities between different modalities may increase performance, though in some cases it also leads to a decrease. One example is the usage of speech, which can be used as an additional input in multisensory interfaces, like the well-known multimodal interfaces that combine speech and gestures. Shneiderman (Shneiderman '00) noticed a clear problem with the usage of speech, especially for more complex actions. Speaking and listening make use of the same mental resources as problem solving, consuming precious cognitive resources. Thus, multisensory does not always lead to a decrease of cognitive load, as for example claimed in (Rosenfeld, Olsen et al. '01). Nonetheless, the combination of multiple sensory or motor systems can lead to error reduction and correction, especially in environments that are troubled by noise (Oviatt and Cohen '00). Not only may the user retrieve multiple sources of information that can lead to the correct perception of the world, computer systems are also greatly helped by providing multiple sources of human output for cross-comparison. This is especially the case for all techniques that are based on recognition engines. Finally, the perceptual structure of the task at hand may support flexible and complementary behavior, by letting the user perform the same task, via different output modalities (Grasso, Ebert et al. '98) (Jacob '92).

3.2.4 Classification

From a control-oriented perspective, most of the unconventional human input and output techniques can be classified into a small number of categories (Figure 3.4). These categories provide a starting point for deciding which direction can be taken, when designing and developing unconventional interfaces, based on user, task, and environment characteristics for the case at hand.
There are basically three ways of providing *human input*. The difference between the different techniques is based on the level in which a stimulus is provided into the sensory processing mechanism of the human body. Receptors, the nerves, or brain areas are stimulated, using a variety of non-invasive or invasive methods. All these methods have been extensively handled in chapter 2 – for more information, please refer to the section of the appropriate sensory system.

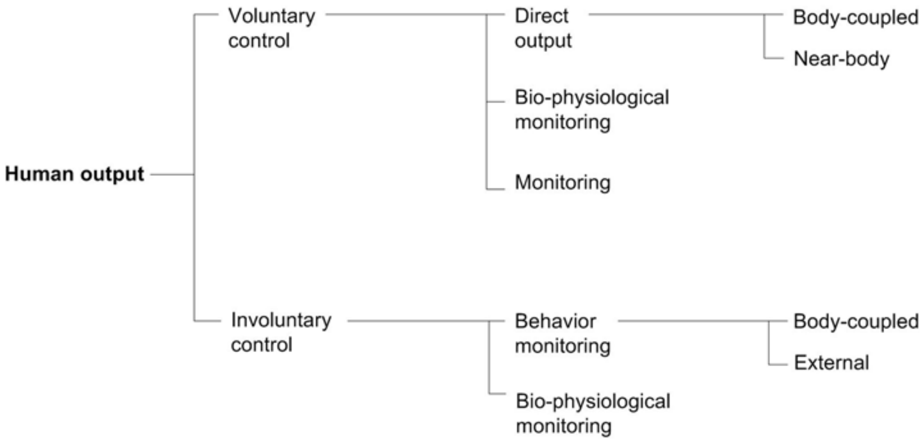


Figure 3.4: *Unconventional human control technique classification.*

On the *control side*, the classification of techniques is slightly more complex. In the first stage, one should differentiate between voluntary and involuntary control, mainly defined by the level of consciousness with which control is performed (section 2.3.6). Whereas voluntary control is mainly based on the intent of the user to perform a specific action, involuntary control refers to the body-internal mechanisms of keeping the human “alive.” It should be stated that voluntary and involuntary actions are sometimes interwoven, as will be explained later. As a result, it can be stated that not all control actions may clearly fit within the proposed classification.

The most common control action is based on direct output, in which a user makes use of body-coupled or near-body devices. Most of the unconventional techniques described in chapter 2 fit into this category. Psycho-physiological monitoring refers to the breed of techniques that come out of the corner of biocontrol devices, as described in section 2.2.6., thereby, focusing on activities in the human somatic system. Finally, monitoring techniques make use of external (ubiquitous) devices like cameras or sensor networks, to track the body configuration (like posture and gesture) of a user. Though related to direct input methods, obviously no devices are used to detect the posture or gesture, even though techniques using direct coupling exist. One example is a data glove for tracking gestures.

The category of involuntary techniques makes use of similar techniques to those used in voluntary control, but with another purpose. Behavior monitoring makes use of body-coupled or external devices to track behavioral aspects of user's actions, as have been dealt with in section 2.3. Finally, psycho-physiological monitoring can also be used for tracking involuntary information, like stress or other emotion, possibly overlapping with behavior-oriented techniques.

Moreover, there are unconscious, though voluntary actions, which are not fully noticed by the user, since they are not fully intended like the conscious voluntary actions. Examples are effects of habits on interaction and moving ("re-posing") particular body parts due to the force put on them during longer work sessions. These actions may not necessary lead to a control action, though may be used to adapt these actions. Through human factors interpretation methods, these unconscious actions can be interwoven with involuntary control actions, thereby closing the gap between the two subcategories monitoring and behavior monitoring.

At the end, what is the control classification good for? Apart from matching user, task, and environment characteristics, some factors need to be highlighted that can be derived from looking at the different kinds of techniques. One factor that affects both voluntary and involuntary control is the role of *intent*. The interpretation of the user's intent is rather hard and gets more difficult when unconscious (involuntary) activities need to be considered. Thus, before using involuntary control activities, one should create a clear picture to determine if it is truly useful to employ these techniques, since there is a large chance of "misbehavior" of the triggered actions by the computer system, in ways that actions get misinterpreted. The field of neural networks, the basis for most biocontrol interfaces, may have progressed quite well over the last decades, but using involuntary control to manage direct or indirect interaction is currently still an open issue.

The second factor that plays a dominant role in especially involuntary control is feedback. Without correct feedback, there is no way a user can notice what or how a computer has reacted to the unconscious data analysis, which may lead to an increasing amount of mode errors, frustrating the user, when the computer is acting on its own will. Thus, within the next section, a more detailed look will be taken on feedback mechanisms.

3.3 Sensory and control substitution

Returning to the human potential view, the cross-modal effects handled in section 3.2.1 gave a hint of what is probably the most dominant design methodology applied, when designing unconventional interfaces. This design method is known as *sensory substitution*. The method originates in the designing of compensory aids for people with sensory loss, resulting in so called assistive technology. Examples include the usage of speech recognition and eye-tracking.

The term “sensory” substitution should be seen in a wider perspective though - in the case of assistive technology, much work also focuses on the substitution of control methods for people with physical impairments. Hence, in the further discussions a distinction between sensory and control substitution will be made. Furthermore, sensory substitution should not be mistaken for sensory correction or replacement, in which the sensory modality is quasi-repaired, using artificial implants (Loomis '03).

Thinking again about the cross-modal effects, which different kinds of “substitution” can be identified? Are all of these methods truly substitution? The answer is “no”: When examining assistive technology, it becomes clear that a distinction between methods needs to be made:

- *Substitution*: one sensory or control channel is functionally replaced by another channel,
- *Addition*: a sensory or control channel is added to the task performance loop, in which the channels are *not* directly coupled. Addition is a general phenomenon in multimodal interfaces,
- *Integration*: a sensory or control channel is added to the task performance loop, but now the channels are directly coupled and, therefore, affect each other.

Both addition and integration lead to what has previously been called cross-modal enrichment. However, especially integration can be affected by cross-modal bias, due to the direct coupling of sensory channels. This process is called multisensory binding (Spence and Squire '03; Weisenberger and Poling '04) and is defined by spatial relationship and a “moveable window” of time-relationship of two sensory stimuli (synchrony). The relationships define how two stimuli are affected by each other, up to the level of one stimuli dominating over the other one. The time-delay might be as small as in the case of a slight delay in a TV-signal transmission, in which the sound of speech is perceived later than the movement of the lips of a speaker are seen, up to the rather conscious mental coupling of seeing lightning and hearing the strike of thunder several seconds later.

How does the field of assistive technology relate to the development of unconventional interfaces, especially for those people without physical impairments?

The answer is quite simple: There are many situations, in which generally used sensory or control channels might not be used or in which it makes sense to apply multiple channels. Besides making vivid interfaces, some other reasons of using sensory or control substitution can be stated. First of all, *mental or motor workload can be reduced*. In some cases, a user’s sensory or control channel is blocked or overloaded. This may occur in applications applying two-handed interaction, in which both hands are used, and yet another action needs to be performed. Another example is cognitive overload, in which the brain is unable to process information from a certain sensory system. In this case, the usage of another sensory or control system to perform the same action might help. Secondly, there are multiple *limitations of VR technology* that can be overcome. Haptics is one field that has been influenced by sensory substitution. On the other hand, money or the size of a group also leads to using "other" methods, simply because standard technologies might not be applicable. Finally, using sensory addition and integration methods, *performance of interaction can be increased*. One potential field of interest is the improvement of techniques that deal with collisions of objects.

Having stated these issues, it also becomes clear that the majority of techniques presented in chapter 2 can be characterized as applying principles that strongly relate to sensory substitution, addition, and integration. Hence, the underlying principles are a powerful way to create new unconventional techniques.

In order to define the need for substitution, addition or integration, a good way is to reflect once more on the four stage analysis model presented in section 3.2.3. Looking at this model, some specific questions can be deduced, resulting in additional user and task analysis issues:

- *Performance limits*: are there any problems in performing actions, caused by mental (cognitive load) or physical (motor system overload, inapplicability, or impairment) limitations that endanger fulfillment of the user's goals?
  In the case of assistive technology, users' performance limits or complete inability to perform has lead to new techniques to enable them to perform actions in another, but suitable way.

- *Matching functional characteristics*: in case task performance or feedback actions cannot be provided at the necessary level, up till which extend can it be taken over completely or eased by adding or replacing it with another human input or output channel? In addition, how does the functional reallocation affect the performance of tasks and what is the psycho-physiological "cost" (effort) of the change?
  The stages of processing information with their perceptual and central processing needs have to be examined and changes identified: through substitution, addition, and integration, the decision making process is the foundation of the task performance changes.

- *Hardware issues*: is the hardware required for functional reallocation applicable in the work environment, and does it not intrude the user's intimacy? Some hardware might not be usable or restrict the user to unwanted levels. When this is the case, it is required to rethink the functional reallocation.

- *Subtask effects*: the usage of multisensory techniques can result in rather complex compound task structures and substructures, which need to be handled in order to guarantee the flow of action in an application (section 3.4.1). Hence, any cross-dependencies between (sub)tasks and input and output channels will need to be handled carefully. Finally, when substitution methods are coupled with dynamic allocation of functions, by combining old and new methods, the effects of possibly mixing these techniques need to be defined.

- *Feedback requirements*: functional reallocation regularly requires the reallocation of feedback and should be checked with regard to overload on a user.

Once this analysis is performed, a closer look can be taken at the actual creation of substituted input or output channels.

3.3.1 Human input substitution, addition and integration

When developing an unconventional interface, one can follow several ways of choosing an input channel (Figure 3.5). A rather prominent approach has been to make use of substitution methods. When dealing with sensory substitution in developing a new kind of interfaces, the term “substitution” is actually slightly ambiguous. As discussed by Lenay et al (Lenay, Gapenne et al. '03), when one exchanges ones sensory channel with another one, one is not simply making a change at the receptor level. The whole information processing loop is re-ordered: the central nervous system needs to learn a new mode of perception. The brain has a high level of *plasticity* to accomplish this (Shimojo and Shams '01), but some effects need to be noted. A clear example is substituting visual information for a blind person by using auditory information. The blind person needs to learn to “see by hearing,” and, thus, needs to create a new *cognitive model* of the world (application) she is dealing with. Hence, when exchanging sensory channels, one needs to deal with the metaphor of communication and the influences the new sensory channel has on the interpretation of the information itself. This matching process can be incredibly hard. When changing the information process needed to interpret the communicated information, user interface designers should take an increased *learning curve* into account: users will need time to adapt to the new kind of information processing. Finally, sensory substitution can also take place within the *same* sensory system.

By making use of another combination of receptors, one can substitute information within a specific sensory channel. The best known example of within-system sensory substitution is substitution within the somatic and kinesthetic system, which hold a multitude of receptors that work together at a cognitive level (section 2.1.3.). A good example to explain some of the effects is the usage of vibrotaction. The usage of vibrotaction is one of the best known examples of sensory substitution on spatial interfaces. Born out of the inability to provide haptic output to a user under all circumstances, one looked at other possibilities to simulate haptic feedback (Massimino and Sheridan '93; Kontrarinis and Howe '95; Cheng, Kazman et al. '96). Examples of circumstances in which traditional haptic feedback devices cannot be used, are those situations in which the user needs a lot of freedom without the restrictions of desk-bound or body-coupled devices or in which there are too many users to support haptic feedback on an individual level. Haptic feedback is sensed from cutaneous and subcutaneous and limb level (skin and muscle sensations), and provides information on both force and tactile level (section 2.1.3.). By using vibrotactile devices, at least some of the haptic information can be communicated to a user, mostly the tactile part (force information is exchanged through vibration). By adding additional sensory channels, more information can be supplied to the user, who can interpret the combination of different sources in order to make a decision. Ideally, this decision making process is similar to the situation in which the user is provided with haptic information. Of course, this will not work for every situation – in some situations precise force information will be needed that cannot be simulated by vibrotaction.

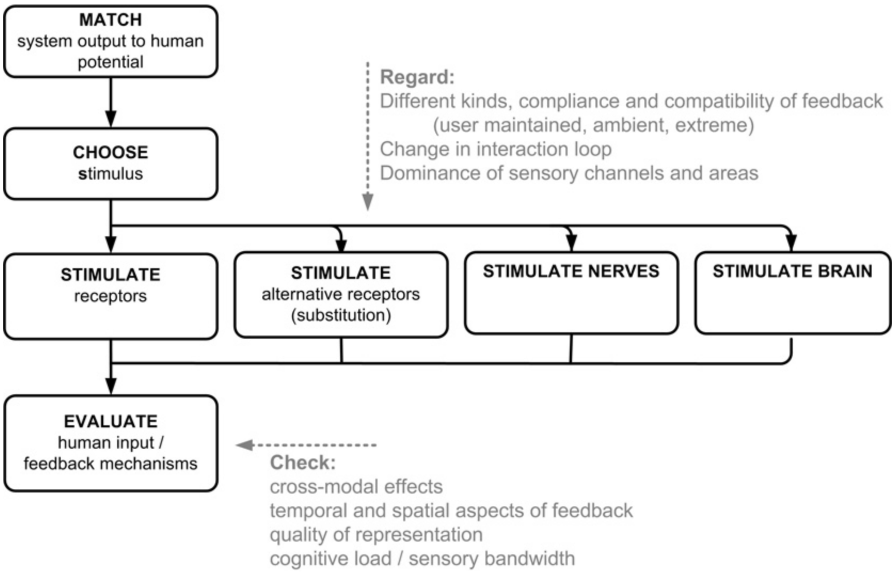


Figure 3.5: *Standard versus alternative human input aspects*

Current research efforts have shown that separate modalities can affect each other when integrated. When a user is confronted with concurrent stimuli, these stimuli can alter the perception of an event due to the plasticity of the brain. One of the studies that laid the basis for studying this plasticity was performed by Shimojo and Shams (Shimojo and Shams '01), which has lead to a number of psycho-physiological oriented human-interface tests. These tests have great validity, when adding or integrating sensory modalities. It may be that when adding a sensory modality, one is actually integrating modalities: cross-modal integration is believed to happen more often than previously been expected. In this section, the most significant results of these tests, with their effects on human-computer interfaces will be handled.

Most of the studies performed focus on cross-modal bias and transfer and have resulted in the following observations:

- *Vision alters other modalities:* already in the 70ies, tests have shown that vision adapts perception when stimuli of different sensory systems are combined. Probably the best known example is the McGurk effect, in which sound alters the speech perception of a speaking person: the sound “ba” tends to be perceived as “da,” when it is coupled to a lip movement of a person speaking “ga.” Another example is the ventriloquist effect, which is the direct coupling of a sound source to a specific visual cue. For example, the sound of speech always seems to come from a specific person when watching TV, even though the sound does not exactly come from that direction. An experiment by Lederman et al (Lederman, Thorne et al. '86) showed that visual feedback can overrule tactile cues, when observing the spatial density of a texture (also see: tactility alters vision).

- *Sound alters the temporal aspects of vision:* Tests have shown that the perceived rate or duration of a visual stimulus can be affected by related sound signals. Additionally, visual temporal resolution can be improved or degraded by sounds, depending on their temporal relationship (Shimojo and Shams '01).

- *Sound alters other aspects of vision:* Vision has a higher spatial resolution, which is why in spatial tasks, it will dominate. On the other hand, sound has a higher temporal resolution, which is why it will dominate in temporal tasks. Thereby, sound does not always only has an effect in temporal tasks. Some tests have shown that sound can alter the perceived visual intensity (Odgaard, Arieh et al. '04). An example, which shows great application in the area of collision detection feedback is the test by Sekuler et al (Sekuler, Sekuler et al. '97), which examined the effect of visual and auditory cues on visually moving objects. The test (see Figure 3.6) proved that, using an x-shaped trajectory, users would observe two objects pass each other without any effect when no extra feedback would be provided. Once a visual flash or sound would be supplied when the objects would cross, users would perceive the objects as bouncing off each other.

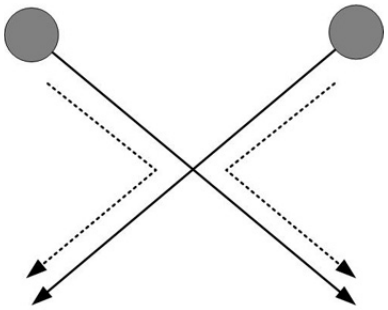


Figure 3.6: *Interpretation of an ambiguous visual motion event. After Sekuler (Sekuler, Sekuler et al. '97)*

- *Tactility alters vision:* As described above, Sekuler et al showed that sound can alter visual perception. In addition, Shimojo et al (Shimojo and Shams '01) proved that the role of the sound can also be taken over by vibrotaction. To which extent the haptic substitution of collision, by using tactile feedback, would affect this outcome was not shown. In another experiment, Blake et al provided indications that tactility can be used to disambiguate visual rotation information (Blake, Sobel et al. '04). Users could correctly interpret ambiguous spatial rotational information of a ball (globe), when touching a rotating globe. Finally, the previously mentioned experiment by Lederman et al (Lederman, Thorne et al. '86) showed that when users should focus on roughness perception when observing visual and tactile cues, tactile cues can overrule the visual ones.

- *Audio alters tactility:* Multiple experiments showed that sound can influence the perceived roughness of a texture (Weisenberger and Poling '04), but that, sometimes, the usage of auditory cues needed to be learned by the user. An example of the alteration was that a higher-frequency sound would result in the perception of a smoother surface. Yet another direction was shown by Bresciani et al. (Bresciani, Ernst et al. '04), in which a user was provided with simultaneous tapping feedback on a finger tip, consisting of tactile feedback and sound. When in a series of taps the tactile feedback would be left out, such that the user could only hear the tap, the user would still perceive tactility.

What can be presumed when looking at sensory enrichment, by adding or integrating modalities? Basically, one can observe the research outcomes from two directions: enriching information through disambiguation (Ernst and Banks '02) and biasing information. By adding a second or third sensory modality, the “correctness” of information can be increased. This means that, especially in more complex applications, performance of interaction can potentially be increased, while increasing a user’s ability to understand the data better.

To apply sensory addition and integration in an application, several approaches have tried to explain how sensory systems work under specific conditions. Spence et al (Spence and Squire '03) introduced the theory of *multisensory binding*, referring to the synchrony of multiple sensory feedback channels (section 3.3.2). It is believed that the brain makes use of spatiotemporal concurrences to identify which sensory stimuli are bound together. Thus, different sensory modalities can refer to the same perceptual event, when they fit within a close temporal and spatial frame. Clearly, this binding will become difficult in circumstances, where multiple events either happen in a small spatial area, or within short timeframes. Synchronization of feedback, thus, is of utmost importance – an offset can result in unwanted effects.

The second explanation of sensory integration is provided by the *modality appropriateness* hypotheses: it suggests that in a particular situation, the modality which best suits the task will be weighted most heavily (Weisenberger and Poling '04). But beware: contradicting hypotheses may invalidate this statement. For example, Shimojo and Shams refer to cross-modal integration in which the modality that carries a signal that is more discontinuous becomes the modulating modality (Shimojo and Shams '01).

To conclude, following guidelines for the application of sensory substitution, addition and integration can be provided:

- *Check temporal and spatial aspects of feedback*: when combining multiple sensory modalities, one should create a clear image of the interdependencies in both time and space in order to match these modalities or to avoid unwanted biasing.

- *Choose the right representation and check possible bindings:* as described by Loomis (Loomis '03), there are different kinds of representations perceived by the sensory system that need be taken into account. Abstract meaning refers to information that can be translated from one representation into another one, as is done in text to speech. Amodal representations describe abstract representations in the brain that can make use of multiple modalities, as is with spatial knowledge, but do not necessarily directly integrate these during perception (hence a case of sensory addition). Thus, using one of the modalities that is normally being employed can result in similar perception in comparison to situations in which all the modalities are combined. An example of this situation is the tactile map for wayfinding (Jacobson '96). Finally, there are isomorphic representations that integrate multiple modalities during perception in order to match these forms of information creating an unambiguous result. After selecting the right form of representation, the kind of binding should be mapped to the task at hand, seeing if any of the before mentioned biasing may occur.

- *Analyze sensory bandwidth*: when substituting or combining modalities, analyze the sensory bandwidth (capabilities) of the user that may change over time during interaction, avoiding sensory or cognitive overload. Furthermore, using a specific sensory modality may limit general cognitive capabilities, like shown by Shneiderman – speaking makes use of the same resources as problem solving (Shneiderman '00).

- *Keep interaction loop changes in mind:* as handled in section 2.3.6 on biopotential, by stimulating the nerves or brain, the interaction loop may change considerably. These changes need to be taken care of when designing the kind and amount of feedback.

In any case, sensory substitution, addition and integration always require careful evaluation, since, as can be can be learned from some of the contradicting research results discussed above, its phenomena are still not completely understood.

3.3.2 Human output substitution, addition and integration

Grounded in the assistive technology mentioned before, human output (control) substitution has found wide application. Similar things can be stated for control addition and integration – in many more complex applications, users combine multiple human output modalities, either serially or in a parallel way. A straightforward example is the usage of mouse and keyboard (*serial*) and the usage of a foot-controlled button and mouse (*parallel*). Parallel integration should be understood as two actions having a close to or identical timeframe – the foot-controlled button can be pushed, during mouse interaction. I can also be used in close relationship in compound tasks, in which a strong relationship between both device actions exists (the two devices are used to reach the same goal or perform a single command). The border between control addition and integration, which shows great resemblance to serial and parallel usage, may be difficult to separate – many actions can have a highly compound characteristic.

Additionally, control substitution and addition/integration can highly overlap. Think again about the foot-controlled button: the foot actually substitutes the hand by performing a button-press action, which is generally performed with the hand, but at the same time can be used simultaneously with the hand (mouse). Hence, in most cases, substitution forms the basis for addition or integration.

In order to see if a control can be substituted, a similar approach can be followed, as performed with sensory substitution: the syntax of a certain human output channel needs to be mapped to another channel (Figure 3.7). Actions can be matched by comparing different biomechanical characteristics of the human limbs (see section 2.2), or by using speech or biocontrol methods. In order to define the possibilities and limitations of control substitution, three major factors need to be analyzed: the capabilities of the user, the control-body linkage, and the control task itself.

The *capabilities of the user* are defined by both the anatomy of a user, and her practice or training with the different body parts. Different body parts can perform different kinds of movement (section 2.2), thereby posing specific ergonomic considerations on using them. A specific action can be performed in different ways by a body part, for example depending on the pose of the user. Sometimes, the substitution of a control can result in such a different pose, since otherwise it might not be performed ergonomically. Furthermore, sometimes control channels are blocked due to impairment or work situation (environment variables), limiting the possibilities of substitution.

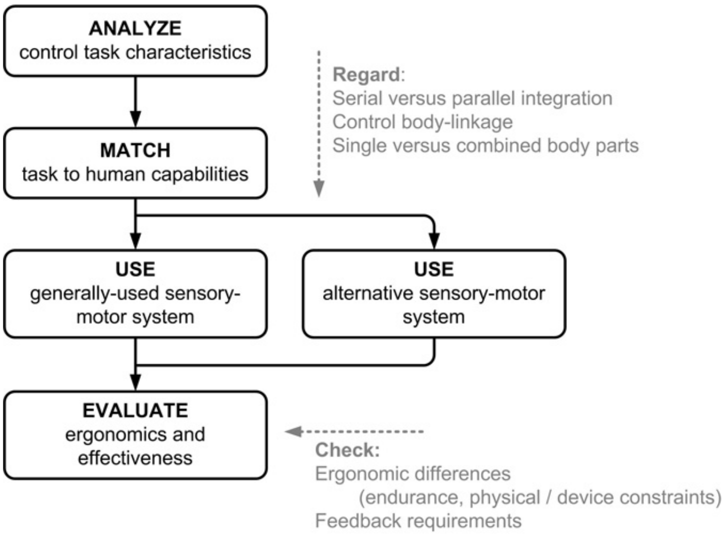


Figure 3.7: *Standard versus alternative human output aspects*

Different body parts will have different kind of *control-body linkages*, defined by device, extremity, grip, and whether dynamic or static coupling is performed. Control dynamics need to be taken into account, including linear and non-linear mechanical transfer characteristics of devices. Finally, the substitution possibilities are defined by the *control task* itself, which may be focused on issues like accuracy, speed or frequency, dimensions of usage, direction, and continuity.

The different possibilities of control substitution have been extensively handled in chapter 2 – all control sections hold examples of ways to perform actions other than by manual operation. The most common “pitfall” is actually the effectiveness of a substituted control. Combining the characteristics mentioned before, effectivity is defined by performance, endurance issues including stress, and may be limited by safety regulations. Furthermore, the information processing can change drastically: when a mechanical control is substituted by a biocontrol method or speech, this may lead to a decrease in biomechanical energy, but most likely increases the cognitive load on the user. As such, task performance may become much more ergonomic and effective from a mechanical perspective, but may tire the user mentally considerably.

There are a large number of studies performed on the usage of different kinds of controls by the user’s hands and feet (Bullinger, Kern et al. '97), but not so many with the other extremities or even biocontrol. It is sometimes difficult to see the relationship between different extremities during control integration. Hence, it is sometimes difficult to apply control substitution. Nonetheless, some directions have been investigated to a larger extent. Looking at the control syntax of integrated controls, much effort has been spent on interfaces that combine speech and gestures. Originating in Bolt’s seminal work (Bolt '80) “put-that-there,” the interdependencies between hand and speech-based output are rather well understood. Such interfaces have been used to provide a possibly more natural kind of interaction, but also to disambiguate human output.

Another area that received a considerable amount of interest is the field of *two-handed interaction*. While at first sight this may not be clearly recognizable a case of control addition, a two-handed interface basically combines two similar control modalities in a single human output method. Thereby, clear interdependencies can be stated between the two hands, defining its control syntax. The basis of most two-handed interfaces is Guiard's framework for bimanual manipulation (Guiard '87). Guiard identified three different kinds of manual actions: unimanual, bimanual symmetric (synchronously or asynchronously), and bimanual asymmetric. An example of a symmetric action is the scaling of an object by moving the hands apart at a similar rate. Hinckley's dolls interface (Hinckley '94) is a good example of asymmetric action, due to the different actions the left and right hand undertake. Focusing on asymmetric behavior, Guiard provided three principles:

- The non-dominant hand dynamically adjusts the spatial frame of reference for the actions of the dominant hand
- The dominant hand produces fine-grain precision movements, whereas the non-dominant hand predominantly is used to perform coarse actions
- Looking at the syntax of performance, the non-dominant hand initiates the manipulation action

All together, control substitution to create an unconventional interface is bound to the same rules as developing general "real-world" controls, such as a telephone or a coffee machine. Controls need to be coded and afford its functionality, and if this is not possible, the user should be informed in an alternative way, how to reach the functionality. Controls can be grouped to create provide a clearer overview or to support a specific order or relationship in processing a control action in a serial or parallel way. Controls should be accessible in relation to the used pose of the user, thereby taking ergonomic considerations into account. Finally, one thing that should be kept in mind is the usage of constraints. Well over a decade ago, Norman introduced the importance of constraining actions (Norman '90), but there are many possibilities that supersede traditional constraints. One such possibility is the usage of dynamical constraints, for example by using filtering of device output.

A final point that needs to be considered is the interdependency between sensory and motor systems during substitution. Changing a control will regularly result in a change of feedback too. The user-maintained feedback will change when a different extremity is being used, quite simply because the kinesthetic feedback will change. Furthermore, by exchanging modalities, different kinds of feedback might be needed to communicate the mode of action. For example, using a hand-based interface, showing a simple pointer might be useful to inform a user where she is pointing, whereas during full-body interface this is not possible.

In addition, the effect can work vice-versa: the substitution of a sensory channel can result in the needed change of a control channel. This will, for example, occur when visual output is exchanged with a non-visual output method. Many human output methods are based upon direction manipulation metaphors, where there is a close match between visual input and human output – when the visual input cannot be used, these methods cannot be used or need to be used in a different way.

3.4 Interaction flow and feedback

The usage of unconventional interaction techniques may cause distinct effects on the user's output structure, also know as the interaction flow. Differences to interaction loops of more traditional systems can occur that need to be dealt with. This section introduces the main factors involved in flow of action and takes a specific look at feedback mechanisms.

3.4.1 Flow of action

Once the new technique has been defined, a further step should be taken to define how the technique should be used within the application, identifying the interaction flow of the user. This interaction flow is often referred to as the "flow of action". Flow of action is a key issue in more complex spatial interaction environments, especially those that mix multiple devices for I/O purposes (Kruijff, Conrad et al. '03). Flow of action refers foremost to the structure of a user's output to a system, but the whole action-feedback loop is affected. Thus, flow of action is grounded in the information processing loop of a user.

The key issue in flow of action is the composite nature of tasks. Basically all tasks performed in a VE are built up of subtasks that are held together via a compound structure. This compound structure is the basis of the problem solving activity of a user and, thereby, directly affect operational effectivity (performance), including attention factors and issues, such as ease of use and cognitive load. One approach, which has found applicability in especially 2D interfaces, is Buxton's chunking and phrasing theory (Buxton '86), in which the compound structure is observed pragmatically as a dialogue consisting of small chunks that make up a phrase through human-computer interaction. Buxton came up with multiple subtasks, like selection, position, orientation, and specifying a path. These tasks can easily be compared to the universal interaction tasks identified in section 1.2: the close integration of selection and manipulation tasks is just one example of compound task structures in a VE. Whereas Buxton puts large value in the pressing of a button to combine different subtasks in a macro-level task, interaction in a VE is slightly more complex than that, especially when multisensory interaction comes into play.

One thing discussed by Buxton was the usage of gestures, in which different subtasks are combined in a continuous way. This statement is a crucial one when compared to spatial interaction, since most of the interaction in a VE is of a continuous nature (Doherty and Massink '99). The continuous characteristic of VE actions can ease the flow of action, since mode errors are noticed directly when the continuous stream is broken down, but some factors come into play that make flow of action a highly difficult issue. One of the more general problems noticed is the still under-developed level of many system control techniques. System control overload is a general problem in many VE applications, in which too much time is spent on issuing commands. In order to combine different subtasks, it is not always possible to both invoke and perform a task in one single action, like it is when using gestures or speech. System control actions regularly need to be performed that truly break apart the continuous nature of flow of action, due to their in-effectivity and poor (metaphoric) integration with other techniques. Some techniques, like those based on context sensitivity, seem to solve these problems better than those that make use of completely "externalized" methods, like menus or even PDA’s. Nonetheless, sometimes these methods are the only way to deal with more complex applications.

Especially when multisensory interaction comes into play, flow of action becomes complex. Due to the additional sensory and/or motor channels, the structure becomes multi-layered. Not only does the user need to change between different subtasks, but also between different input or output modalities. To sustain continuity in the flow of action, several issues need to be regarded in order to avoid mode errors, increasing ease of use and performance. The most important factors are as follow:

- *Cross-modal task performance*: when the user is allowed to use multiple modalities to issue a command, dynamical allocation of functions needs to be carefully investigated. The dependencies between tasks need to be analyzed in order to guarantee that related subtasks can be performed using the same modality or by a well-performing integration of several modalities. Hereby, small so called repeating interaction loops (Kruijff, Conrad et al. '03) need to be handled with: often, actions like minimal navigation are thrown in between larger subtasks. These small loops need to be supported in such a way that these loops do not disturb the performance of the larger subtasks. The trade-off of using a specific modality of device to perform such a smaller interaction needs to be regarded. Sometimes, the used modality used for the main subtasks may not support the small loop in the best possible way, though changing between modalities or devices may pose a much larger problem on the total performance of the compound tasks. Hence, switching should be avoided. One of the successful cross-modal task performance examples is the usage of combined gesture and speech actions (Bolt '80).

- *Cross-modal feedback*: when dealing with multisensory interfaces, one should always be sure that the user is able to register the feedback in a clear way. When multiple human I/O channels are used, it is best to make use of at least one unique sensory channel to communicate a basic amount of feedback to the user, to avoid mode errors. When feedback is scattered over multiple sensory channels in an incongruent way, the user will most likely get confused, unless the sensory information is coupled. That is, feedback can be enriched by adding another sensory channel, but replacement during the flow of action in a single compound task is not recommended. Under all circumstances, feedback needs to be compatible between sensory modalities, to communicate the same symbolic information throughout all modalities to the user. One of the most simplistic feedback methods in multisensory installations is to maintain a visual element in the corner of a display that communicates the current state of the user.

- *Cognitive overload*: sensory or motor overload can easily happen when a user is confronted with a large amount of information. For example, this may occur in large cylindrical displays when information is registered in the peripheral field of vision but not used actively; The user is not able to obtain an overview of all information at once and needs to turn her head continuously to get a clear picture. With complex data sets, this may put a high cognitive load on the user. When additional cognitive resources need to be used for further sensory or motor actions, it can cause considerable disturbances in the flow of action of an application. Hence, focus of attention in both the sensory and motor channels

> needs to be handled carefully. Furthermore, sensory information needs to be "tuned" in order to create a coherent information stream to the user.

There are multiple methods for investigating the flow of action inside an application. Cheap usability testing, for example by carefully observing a user, can already bring many problems to light. In more complex applications, though, it makes sense to log a user's actions during interaction, investigating the amount of time spent on different subtasks. When the timing is co-related to attention and cognitive load factors, pitfalls in the flow of action can be found.

3.4.2 Feedback mechanisms

As noted in the previous section, feedback is of utmost importance when dealing with the structure of information in an interaction loop. Every user interface needs appropriate feedback mechanisms to inform the user on the current state of the system. In so doing, feedback can be viewed from either a human or device-oriented view, that is they focus either on the sensory channels of a user or on the system (Bowman, Kruijff et al. '05).

From a human point of view, feedback can be categorized according to the human input dimensions, which reflect the different sensory channels. Just like any other human input technique, feedback can be uni-modal or multi-modal, possibly applying sensory substitution, addition or integration methods. Hereby, it is important to mark if the feedback mechanisms work in parallel or serial order, in order to define the consequences of cross-modal effects.

From a device or systems oriented point of view, three different kinds of feedback can be identified: reactive, instrumental, and operational feedback (Smith and Smith '87). Reactive feedback is user-maintained, self-generated feedback that results from operating the user interface. Instrumental feedback is generated by the control tools and elements, such as the vibration of a pen when a button is clicked. Operational feedback is the information received from the system as a result of the user's actions. The difference between these types of feedback can not always be made, but some useful considerations can be formed by focusing on the different kinds, separately or viewed as a whole.

A first, interesting remark can be made concerning self-maintained feedback. Initial studies, among others focused on the value of a button press, like the previously mentioned study of Buxton (Buxton '86) on the effects of continuous input for gluing different subtasks together. With the advent of full-body interfaces, there is a multitude of self-maintained feedback in comparison to the finger-oriented feedback in desktop interfaces. Users may receive a considerable amount of information that can be useful for spatial interaction. One of the first studies, which more or less played around with proprioception effects was performed by Mine (Mine, Brooks et al. '97a). This study has been (mis)used by multiple studies afterwards, when dealing with more or less physical-oriented interfaces, such as those using tangibility. Nonetheless, the power of self-maintained feedback, which can actually cover all combined sensory information maintained by the user's body, can provide useful insights in the state of the interaction.

Related to self-maintained feedback is what has been labeled passive feedback, in which props are being used to manipulate the virtual world. One of the first examples by Hinckley (Hinckley '94) has inspired many researchers, showing how instrumental feedback can be made available, based on the user quasi maintaining feedback herself by holding the object.

The second issue, which needs to be dealt with, is spatial and temporal compliance, in order to avoid the displacement of feedback. Feedback compliance is an important issue in multisensory interfaces: cross-modal feedback should comply between the sensory modalities used. Feedback should not conflict between sensory modalities, which would result in the degradation of performance (Smith and Smith '87). This becomes increasingly important when modalities are switched during interaction. The quality and metaphor of feedback should be maintained, either by using multisensory feedback, or by using a singular feedback channel that consistently provides the user with feedback, even during the switch between modalities. It may seem obvious to make use of the visual channel to sustain feedback and guarantee compliance – the visual channel is generally regarded as dominant sensory channel. Nonetheless, this is not always the case (section 3.3.1). Visual feedback can be overruled by another sensory channel (Shimojo and Shams '01), and there can be a competition for visual attention (Sellen, Kurtenbach et al. '92). This competition will become increasingly difficult to deal with, when feedback is not context sensitive. When the user will need to perform smaller or larger visual movement arcs to obtain correct feedback, this may decrease performance to a larger extend. Nonetheless, with more complex channels, this may be needed in order to provide the full spectrum of feedback (Kruijff, Conrad et al. '03). Furthermore, inter-sensory conflicts may occur that will create unwanted mental or even health effects. An example is cyber sickness, which is thought to be caused by the sensory conflict between visual, vestibular, and proprioceptive information (LaViola '00a).

Two special kinds of feedback should be handled separately: ambient feedback and “extreme” feedback. Ambient feedback is feedback that can be used in monitoring systems that provide peripheral feedback to a user. One example is a change of color in the light of a room to indicate that the temperature of the room has been changed, after the user has entered the room after a day’s work. However, there are many more of these kinds of feedback that still need to be explored (Ambience '05; MITOxygen '05).

On the other hand, extreme feedback is a rather uncommon way of providing information to a user. Under certain circumstances, users might be overloaded with sensory information or be loosing attention, which may call for high capacity kinds of feedback. An example is the usage of extremely bright or colorful feedback methods to trigger the user when she is visually overloaded. Another example is the usage of small electroshocks in a car-interface, resolving safety issues when the driver would fall asleep.

The bottom line of feedback is rather simple: users should have a clear idea of the current state of the system, and to achieve this, not too much of their cognitive capacities should be used for that purpose – clearly, evaluation of an application is needed to find this out.

3.5 Application and transfer issues

When developing unconventional interfaces, some issues need to be regarded that affect the actual integration in applications. This integration often is a case of porting the research results from a rather experimental environment, to a real-world application. This section focuses on some of the more prominent factors involved, thereby giving a closer look at social and ethical issues that often affect the usage of unconventional interfaces. Some of the issues listed below may result in a new development process loop. Hence, they also could be taken into regard when doing a user and task analysis.

- *Short-time innovation transfers, but long-term macro changes*: firms like Sony are showing that highly innovative technological components can find their way into products in a relatively short time. However, most of these innovations can often not be directly seen – they happen at micro level. At a macro level, drastic changes need to be made to general-purpose products, which need a longer time. Think about a Moloch like Microsoft, still dominating the desktop market: the company basically makes products that need to "serve all" from kid to granny, thereby slowing down possible more radical changes. The reason for this is quite simple: these products need to stay *accessible* for all users, and not all unconventional technology is suitable in that respect.

- *The need for evaluation*: many unconventional techniques are still in the cradle of their development, and will need considerable evaluation before they can hit the market. This also partly explains the previous point: small developments can make a faster market entry than those that require larger changes.

- *Integration in work processes*: unconventional techniques often require changes in work processes, due to the difference in performance. This provides difficulties especially in more complex work environments: for example, integrating new technology in the development process of an airplane takes up to 10 years. This point partly overlaps with the previously mentioned issue on macro level changes.

- *Understanding applicability*: with some of the currently available unconventional technology, it is not clear for which purposes they can be applied. This is particularly true for assistive technology. They may be great for supporting disabled people, but may yield little use for the general user. Furthermore, there have been many gadgets that hit the market fast, but are simply useless. It seems that not all inventions are made to be applied.

- *Law limitations*: unconventional technologies do not always comply with rules, or may need special permits. A clear case is the usage of implants, but there are other directions (including pseudo-haptics) that regularly run into limitations of what is allowed. An example of these limitations are the rules written down in ISO standards on work conditions.

- *Social acceptance*: as will be further handled in section 3.5, social acceptance highly affects unconventional technology. Factors include cultural background and age.

- *Conservatism versus experimentalism:* relating to the first point on macro changes, conservatism is still a point to consider. Especially with desktop systems, many users stay with what has been usable and known for a long time. New technology may not persuade them to change their habits. On the other hand, one can see a highly "experimental-friendly" community in young users, especially those that use game consoles. Hence, young users (literately the future users) will probably discard some of the conservatism that is applicable to current technology.

- *Interactive versus general usage products:* the pace of technological inventions is especially fast for general purpose products like display systems or mobile phones, but less for highly interactive technology. One example that stands out is the Sony EyeToy, a truly innovative interactive medium that has hit the mass market and has stayed there for a longer time.

- *Media integration:* there is currently a huge trend in the direction of integrative media, ranging from mobile phones with a camera and PDA up to so called MediaCenters that integrate desktop computer with video recorder, home cinema system, and internet capabilities. Integration of these media will require new interaction methods, mixing different kinds of input media – the market for truly hybrid devices is still open.

- *Interdisciplinary, multi-directional media transfer:* one can see multiple areas in which general technology is being coupled with (new) information technology in an interdisciplinary way. One such field is the medical area. The same is true in the opposite direction, where medical technology is finding its way to information technology.

It is not always easy to tell in which fields unconventional interfaces will be applied. One of the more straightforward usages is the application of assistive technology, which will certainly further find its way into the medical treatment and enabling technologies area for disabled people. The usage of techniques that have originated in the assistive technology sector will also be further investigated in other areas, whether it is in exploratory areas like arts, or in playful areas like games. Porting assistive technology to the desktop is not always easy – what may work well for disabled people may not work at all for a general user, even if the method shows great promise. One example of a technique that has found good applicability as an assistive input method is speech recognition. While it is well used by those who are unable to type, it has never really found its way into the desktop MR areas, even with the huge efforts through research and industry. It may still be due to performance issues of speech, but it may also well be the characteristics of the medium that make it hard to apply.

Originating in assistive technology, the area of man-interface integration (cybernetics) still sounds like science fiction. Investigating this area is often regarded as non-scientific, but true advancements have already been shown. What will certainly come by ways of assistive technology is a larger diversity of implants. The current usage of ear implants already shows the success of this direction. Such implants will not always work as assistive input method, but may well be used by “free will.” Thereby, *embedded* computing will be given a whole new meaning, but it is currently hard to say in which direction it goes. Warwick et al (Warwick, Gasson et al. '03) have experimented with invasive implants for interface purposes, but this is only one of the very few (known) examples. One imaginary direction may be the usage of user-embedded tracking methods in which the coordinates of a user’s body parts are always known and can be transmitted for multiple purposes. The user’s complete body will become an interface, without actually attaching any devices. Another, more likely, technique would be the embedding of medical monitoring devices to allow remote monitoring of patients, supporting a high level of mobility.

A field which has literally been a playground for unconventional techniques is the entertainment area. Ranging from current game console input methods like the EyeToy up to professional entertainment technology, as developed by Disney Imagineering, unconventional media have quickly found ways to excite or surprise users. It is also the games industry which may force the largest changes in the way we interact with a computer. Due to large market share, this industry branch has already shown how to change the graphics board market, and, due to competition, it may well further advance the interaction area too.

Relating to both the previously mentioned media integration issue and the application of entertainment is the area of mobile and ubiquitous computing. Whereas we already see the integration of multiple devices into a single one (mobile phone), more and more work is being performed to further integrate and hide information technology. This may take the form of further experiments to veil electronics in everything we use (embedded computing, section 3.1), from clothes up to the bedside lamp. In order to interact with these computational units, the devices may rely on involuntary actions to provoke actions or completely new forms of voluntary actions like the usage of isometric muscular activity (Costanza, Perdoma et al. '05). The field is still in its definition phase, with new directions like ambient computing rising, but not fully grown. Hence, it is still unknown what kinds of interaction truly will need to be developed.

Finally, a field that is currently growing is that of man-robotics interfaces. There have been multiple approaches in which robotics have been used in the haptics area. However, research efforts are heading towards the development of appropriate interface methods to communicate with or control robots. These interface methods may take the form of direct communication (speech methods), remote control via mapping of body movements to robotic control, or even cooperative interaction techniques where a user could be performing an action together with the robot, for example using AR techniques.

Social and ethical issues

Due to the experimental nature of both unconventional interfaces and the field of spatial environments, there are many additional factors that influence the level of usage of these interfaces. This section takes a short look at some of the issues that have a social or ethical nature. However, it does not provide a rating (reflection) in an analytical sense. It is clearly neither the aim to label unconventional techniques as good or bad or to analyze possible social changes affected by its usage. Nevertheless, some issues or guidelines need to be stated that reflect a basic level of end-user centered design.

Many techniques presented in chapter 2 may be regarded as unethical. For example, the PainStation (PainStation '05) inflicts pain to the user, thereby clearly endangering the health of the user by damaging the skin. Another example is the urinary control installation (Figure 3.8,(Maynes-Aminzade and Raffle '03)). These installations are used of free will and should be regarded as such. As long as these installations do not willingly penetrate the intimacy of a user, who does not want to be confronted with it, these experimental installations can quasi co-exist. Nonetheless, a border will be crossed, passing by social, cultural and ethical values when these installations hit the public space accessible by every kind of user. The same kind of border can be passed by the wider application of bioware (bioengineering / assistive technology) on non-impaired users. How far is someone allowed to tune her body? Some people see the integration of bioengineering technology and robotics in both the social environment and the human body in a very positive way, like Kurzweil (Kurzweil '06). However, it remains to be seen how far the medical usage of bioengineering technology will affect general usage of both invasive and non-invasive technology. In general, people still are reluctant to use invasive technology, especially when there are no medical reasons.

Figure 3.8: *Socially and ethically questionable: You're in Control. Courtesy of D. Maynes-Aminzade and H. Raffle / MIT Media Lab*

Many of the social and ethical issues that are valid when observing unconventional issues are those that have been identified over the past decades when dealing with computational technology. These include: individual and professional responsibility, access and equity, integrity, risk, and privacy, and they should be regarded from a social-cultural perspective (Jacobs '88; Granger, Little et al. '97). The major consideration with most unconventional techniques is, whether they are used privately or in public space. Techniques that are used privately are mostly used of free will – as stated before, people make use of this techniques without being forced. As soon as such techniques are used in public space, in which people are confronted with them even if they are not willing too, borders probably need to be found that fence off intimacy problems, including those problems related to age. Obviously, the usage of the PainStation is not intended for kids.

The question of health is a different one. The usage of the PainStation is definitely not conform to international health regulations (Figure 3.9). In its current stage it would also not hit the market. But, borders are rather thin, which can be seen with the Bioforce from Mad Catz (MadCatz '06), a device that was almost put on the market. The BioForce makes use of small electroshocks to stimulate the user with haptic-like effects. What may seem to be a cruel device is generally allowed in the medical or sports area, in which the same principle is used for muscular training. Hence, it may not always be easy to draw the line between allowing and not allowing devices that have health limitations.

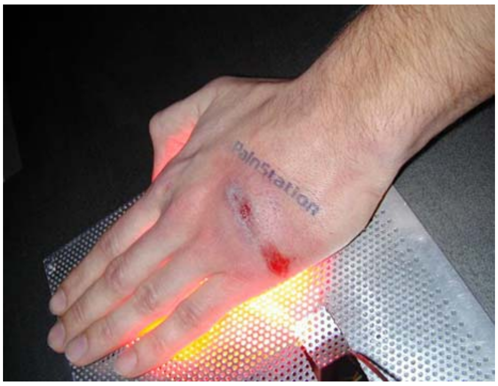


Figure 3.9: *Health infliction*: *hand of user after using PainStation. Courtesy of www.painstation.de*

Directly related to health issues are the reflection of ergonomic issues – of course, force and movement of devices need to be kept within range of the user, adapted to the user's body parameters. A large haptic device may work well for an adult, but may inflict serious problems on a child. These ergonomic regulations are well known and can be found in a human factors book like (Salvendy '97), which also holds references to the related ISO standards.

3.6 Developing devices using garage interface design methods

This section may seem an "odd duck" in this chapter, but actually presents one of the major directions in developing unconventional techniques from a device oriented perspective. Garage interface design basically is grounded in experimenting with different kinds of devices, taking them apart, using parts of them to built new devices. A similar approach is taken by looking at single sensors or actuators, investigating what is possible when they are used in different situations, or in different combinations (Greenburg '02; Forman '03; LaViola '04). This latter approach goes much further than just taking a mouse apart and using its electronics within a completely other device: there are multiple examples, in which basically every kind of object has been considered as interface device. Just one example is the usage of a coffee machine to provide olfactory output (Dinh, Walker et al. '99).

As can be guessed, the garage interface design approach is rather experimental. It should be seen as the basis for finding new techniques, to create a prototype, and not as a way of creating industrial devices at once. Nonetheless, some of the experimental devices have made it into the industrial league; The Cubic Mouse (Froehlich '00) is just one example. Basically said, garage interface design is an excellent way to test new ideas or to improve current interfaces.

In order to create a new device, only a few parts are needed: sensors and / or actuators, some electronics to connect these elements, and a suitable housing. Probably the best way to get any sensors and the electronics to connect to them is to dismantle existing devices. Mouse, joystick, it does not matter. Probably any existing input device has been taken out of the housing already to see how it works and what can be done with it. The great advantage of using such devices is that one gets cheap and possibly robust sensors, which most of the times can rather easily be built into another housing.

Furthermore, using the standard device electronics to connect the device to a computer saves much time, since one can directly use it with the original drivers. So, the order to go ahead is: analyze which buttons and actuators are needed, and, then, take a look which readymade devices they can be taken from. If it is impossible to take the elements from existing devices or if they are difficult to port to a new housing, one should make use of the sensors or actuators that one can get in the electronic store around the corner, or to go for industrial quality elements. Such elements have the disadvantage that they are sometimes more difficult to connect to, but due to the availability of a huge number of different sizes and sorts of sensors and actuators, it is likely that a good one for the intended housing can be found. Some providers of sensors and actuators make use of MIDI interfaces and offer a good amount of elements that can be directly connected to a computer via a MIDI controller. This provides a rather friendly way of building interfaces. However, sometimes the range of elements to choose from might not provide the right one for the intended housing: some of the sensors are rather large.

Yet another and potentially simpler way is to make use of “suites” of sensors and actuators that already come in some kind of housing. LEGO Mindstorms (LEGO '05) is just one example, but may pose difficulties when one needs to put the blocks inside a smaller housing. The same goes for other toolkits like ActiveCube (Kitamura, Itoh et al. '01), Phidgets (Greenburg '01), or any of the biocontrol systems handled in section 2.2.6.

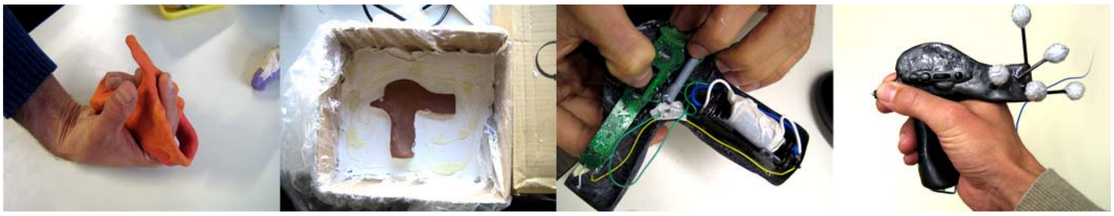


Figure 3.10: *From clay model to final device (Eye of Ra).*

The second step is to create a housing for the device, for which there are several approaches. First of all, the basic form of the device needs to be defined, considering the task characteristics (amount of buttons, needed degrees of freedom) and the needed ergonomic considerations in order to perform the tasks in an orderly way. To define the basic form, one can make use of foam or clay to experiment with different kinds of grips and button placements. Another way is to make use of LEGO bricks, but for many hand-held devices, the square forms of the blocks will be unsuitable to come to an ergonomic form. During experimentation with different forms, one should consider where to put the electronics, since, sometimes, this is rather difficult – often, there is not much space inside the housing. In addition, when one can do without many cables, this is highly preferable.

Once the final form has been found, one should consider how to prototype the real housing. To produce a highly precise housing, making use of rapid prototyping methods like stereolithography is a good, but time-consuming way, since it involves modeling the device in a modeling program. When this possibility is too costly or simply not available, the only way is to do it yourself. A good way is shown in the Figure 3.10, in which a negative model was made from the final clay model, by using plaster. In the plaster, glass fiber and carbon mats were laid together with epoxy glue to create a lightweight though sturdy housing. Afterwards, the electronics where put in to create the final prototype of the design (see section 4.8). Once the device has been created, a software interface needs to be found to connect to the buttons. In case where the new device is based on parts of an existing device, this will not be too difficult: most if not all of the sensor and actuators can be accessed using the driver of the original device. A similar case occurs, when the before mentioned toolkits like LEGO Mindstorms, or a MIDI-based system have been used. The devices can be used directly. When separate sensors or actuators are used, the case gets slightly more difficult. Sometimes these elements can be accessed using standard print boards of another device. However, sometimes, one needs to make use of a microcontroller which needs to be programmed for that purpose (Forman '03). Finally, when multiple different devices are being combined, device infrastructures like VRPN (Taylor, Hudson et al. '01) or OpenTracker (Schmalstieg and Reitmayr '01) come in handy, taking over the device control layer of many currently available devices.

Concluding, the following issues are important when designing an interface yourself:

- *Every object is an interface*: basically every object can be seen as an interface, as has often been probed in the large multitude of tangible interfaces (Ishii and Ullmer '97). For a spatial interface, this often means attaching a tracking sensor and a button to an object.

- *Correct grip and coupling*: the form of a new device needs to be carefully analyzed, taking the task domain and the different kinds of end-users (with small and large hands, left and right-handed) into account.

- *Accessible buttons*: adding an additional button to the housing does not automatically lead to improved and accessible functionality. The placement of the button is of utmost importance. Thereby, re-grasping to reach the button should be avoided.

- *Appropriate weight*: most of the time, a light, but strong, device is needed to perform spatial interaction, especially when it is held in free air. Nonetheless, there are multiple occasions when the device needs to be or will automatically become rather heavy. An example is public space devices used by many people. Here, devices need to be extremely strong, often resulting in the use of heavier materials.

- *Hygienic material*: both public space and specific “hygiene critical environments” like hospitals require a careful choice of material, in order to avoid hygienic problems. Even with prototypes it is important to consider these facts (Pausch, Snoddy et al. '96).

- *Easy connection*: it is important to consider and test how to connect the device before actually building it. For example, the balance of the device needs to be checked with the cables possibly coming out of the back of the device. Wireless connections are often preferable to reduce the off-balancing by cables. However, these connections are often disturbed. Furthermore, in case of public space installations, one is often bound to cables, in order to assure the devices are not taken.

3.7 Evaluation

Though often ignored when designing spatial interfaces, evaluation is an integral part of the development process. Both in more experimental design processes, and those that focus on truly integrating unconventional interfaces in more conventional environments, evaluation delivers key understanding of facts required in order to state the success of the technique in relation to the identified goals (whenever these goals can be identified). Most of the times, evaluation is rather experimental, getting a first impression of the technique or device reacts, and it is not necessary to carry out a formal performance test.

The evaluation of unconventional interfaces largely overlaps with informal and formal evaluation methods applied in the field of 3DUIs. Such methods include heuristic evaluation (Nielsen and Molich '92), cognitive walkthroughs (Polson, Lewis et al. '92), formative and summative evaluations, and the usage of questionnaires and interviews (Hix and Hartson '93). A good overview of factors involved in the evaluation of spatial interfaces can be found in (Bowman, Kruijff et al. '05), including user and task performance metrics to set up the boundaries for an evaluation or general issues that need to be regarded when evaluating, due to their difference with desktop interface evaluation. Hence, this section will not provide a general overview of evaluation principles. Some factors that are mentioned in the before mentioned source, and factors which can be derived from the previous statements in this chapters will be handled separately, due to their importance for unconventional interfaces.

- *Human potential limits*: a key issue when evaluating the performance of an unconventional interface, is the limiting factors of an unconventional interface. Evaluators need to regard to which extent the technique can be applied by a wider public without crossing the boundaries of the abilities of a specific input or output channel. These boundaries are highly user-specific, for example heavily dominated by age. Creation of a user group with highly varying psycho-physiological characteristics may be required to get an actual, non-biased result of an evaluation. In some cases, the particular focus on cognitive or motor overload should be guaranteed. To define the performance metrics, a good place to start are the factors identified in section 3.5 and 3.6.

- *Experimental status versus integration*: evaluation generally goes through multiple stages. The characteristics of an initial evaluation, in the experimental stage of development generally focuses on different issues than an evaluation that tests the ability to integrate the technique for “daily” usage. Initial evaluations might focus on the applicability of a technique, finding out for what purpose the technique might actually be used and what general problems occur. Looking at integration, issues like meeting specific goals or applicability to standardizations come into the foreground.

- *Generalization of results:* especially with initial experiments, it may be hard to generalize evaluation results. Just like with general 3DUIs, there are many unconventional interface solutions “looking for a problem”.

- *Social issues matter*: as identified in section 3.5, social matters truly matter. Due to the experimental characteristics of many unconventional interfaces, social acceptance might be very low, possibly leading to complete rejection.

> Due to the generalization problem mentioned before, it may be hard to identify the actual user group in the initial phase, and, therefore, to know if there is a chance for interface reject – hence, the development and evaluation process always includes an element of insecurity.

Even with the similarities to general (spatial interface) evaluation methodology, evaluating unconventional interfaces can be difficult. Most of the time, unconventional interfaces are being developed under the limited conditions of a research project, lacking the possibilities of a large-scale evaluation as can be paid in the industry. This is often reflected in the limited variety and number of users or the lower level of formality which can be guaranteed in comparison to professional evaluation laboratories. Regularly, it makes more sense to start with an informal evaluation to explore the factors affecting the experimental technology or interface than to dive into a large formal experiment. This approach is strengthened by the regularly failing strict hypotheses. Frequently, hypotheses can first be defined after having performed the initial test, since the technology or interface may not be comparable to conventional interfaces. Most of the time, unconventional interfaces are developed based on a hunch that the new technique could be useful for certain situations, but what these situations are and how it would react is mostly unknown. At this point, it is important to look again at the main goal of this dissertation (see preface):

> *The ultimate goal is to find out how the potential of the human body can be used to design, develop and analyze new spatial interaction methods that surpass performance or application possibilities of currently available techniques*

Once more, the human potential itself is the starting point of development approach that is promoted in this work and has found its way into several of the case studies in the next chapter too (especially Shockwaves, BioHaptics and Tactylus). The human potential is explored, after which techniques are being designed which are experimented with to see how they react. The outcome is knowledge on how things work coupled to a possible practical application. Hence, phenomena are explored that, hopefully, result in the definition of performance characteristics of an interaction technique. The explored phenomena mostly deal with top-level issues: the design of unconventional interaction techniques is rather often the “science of fuzzy problem solving.” For example, the aim of the Shockwaves study (section 4.2) was to get some kind of haptic feedback for larger groups of people (problem: no available group-based haptic feedback methods). However, there was no clear idea how the new method could be realized.

As a result of these explorative studies, techniques can be refined to focus on a specific problem. The refined technique can be regarded as any other spatial interaction technique and, based on its performance characteristics, easily tested in a standard performance test.

To conclude, it is important to note that “explorative” evaluations are very important. Without these, emerging theories and techniques are hard to develop.

3.8 Summary

In this chapter, theoretical and practical issues have been treated that affect especially the design and development of unconventional interfaces. Some issues in this chapter have been put in the foreground, of which sensory and control substitution are the most dominant. It is not hard to imagine that many of techniques presented there were based on the process of sensory or control substitution, addition, or integration. Hence, it should also be seen as one of the main directions or aids in creating unconventional interfaces. What can also be seen is that many interfaces tend to be integrated with everyday objects. There are a large number of interfaces that are based on the “tangibility thought” as introduced by Ishii and Ullmer (Ishii and Ullmer '97), which is currently moving into the direction of packing everything with sensors or actuators to make use of them as interfaces. A short look at the how-to’s has been provided in section 3.6. However, it is just the basis for the general slogan: see what is possible and try everything, looking at both the user’s capabilities and the affordances of everyday objects. That this process is rather creative seems to be natural. In 2000, Shneiderman recognized the strength of creativity for innovation, its basis being an inspirational moment, a structural approach or the influence of situational factors (Shneiderman '00). After all, the designing of unconventional interfaces stays experimental, the best approach depending on the “inventor” and task at hand.

When designing unconventional interfaces, one should look to related fields of research. Caused by the highly interdisciplinary nature of spatial interfaces, research fields such as tangible interfaces, ubiquitous computing, and ambient intelligence can feed the development of new and possibly unconventional spatial interfaces. Results of this research should be critically reviewed, before starting own development, since some pitfalls can be avoided when a specific human I/O channel has already been tried for a similar purpose in another field of application. This is especially the case when technology is re-applied in a new field of application, as is, for example, the case when dealing with biopotential interfaces.

Design of unconventional interfaces often deals with fuzzy problems. Frequently, the design of an unconventional interface does not address a specific problem directly, but rather meant to create an understanding of the phenomena that may be related to top-level problems such as designing interfaces for a specific user group (like the disabled).

Human potential analysis can be the key to identifying new interaction possibilities. Being the main method presented in this dissertation, the analysis of human potential (*human-oriented design approaches*) can be a highly human-centered approach in the development of new kinds of (unconventional) interfaces.

Technology can be a starting point of design, but should best be exhibiting human potential. Technology (*device oriented design approaches*) can take an influence on the analysis of human potential, by evaluating how this potential can be “unlocked”. For example, new technology can be re-applied to another field of application or gadgets can be used to make something artistic. Nonetheless, within the approach taken in this dissertation, it is always best to reflect technology to human potential.

Multisensory is not the same as multimodal. Recently, theories have appeared, proving that sensory modalities should not be seen as separate entities (as has been viewed upon in multimodal approaches), but that entities may affect each other. Most of the time, vision alters other sensory modalities, but experiments have shown for instance that visual perception can also be altered by other sensory modalities.

Both limits and possibilities of human potential should be regarded. When defining a new technique, it should be adapted to the psycho-physiological limits of the user or user group at hand. Thereby, these limits can also be seen as a possibility, to truly use the potential of user abilities.

The four stage analysis model can aid in designing techniques. When observing human potential, four different levels of analysis need to be taken into account. Following the model from Gopher and Sanders (Luczak '97), task variables, processing stages, energetical mechanisms (mechanisms that foremost focus on the effort to plan and perform a task) and cognitive resources can be identified that have an high impact at the design of new techniques.

Consider sensory and control substitution, addition, or integration. Based on the “sensory plasticity” of the brain and the ability of body parts and sensory systems to take over (or resemble) functionality of each other, techniques can be designed that make use of alternative output and input methods. Depending at the user and task at hand, using alternative techniques can make sense, for example to reduce cognitive load, or to allow performance of a task when another sensory or motor system is blocked or even unavailable (due to disability).

Supporting flow of action is important for interaction with complex applications. Interaction flow, which foremost refers to the output stream of a user to an application, should be carefully regarded, especially when multiple control devices are being used. Clear feedback mechanisms and the careful identification of possible “chunking” of actions can lead to improved flow, whereas a disturbed flow regularly leads to considerable performance decrease, potentially caused by cognitive overload.

Regard social and ethical issues. The acceptance of a technique is affected by social and ethical issues, such as health risks or privacy. Unconventional techniques are often highly experimental and may therefore regularly touch social and ethical boundaries.

Garage interface design is a powerful approach to create new and possibly unconventional interaction devices. Creative trial and error approaches are powerful ways to create new techniques. Approaches include the usage of media toolkits to try out different combinations of sensors and actuators, or the usage of parts of existing devices in a new housing.

Evaluation is very important, but can be difficult. Evaluation is an integral part of the development process, and should especially be regarded when the limits of user capabilities are reached closely. However, within most non-industrial projects, evaluation is performed under limited conditions, which mostly leads to less formal results. Nonetheless, these results are extremely valuable and especially deliver good insights when the design process is highly experimental.

CHAPTER 4
Case studies

4.1 Introduction

In this chapter, several case studies will be presented that illustrate and strengthen the information handled in the previous chapters. The case studies tackle two main issues for unconventional interfaces: using sensory or control substitution, addition or integration to create new techniques and the using unconventional techniques in conventional environments, leading to so-called hybrid interfaces. Additionally, one study focuses specifically on the performance analysis of an unconventional device in comparison to several generally accepted 3D interface methods.

The first three case studies focus on developing purely unconventional interaction techniques, by applying substitution methods for haptic stimulation through alternative feedback methods. *Shockwaves* explores the usage of audio and air-based shockwaves to provide pseudo-haptic feedback in large projection systems, by using large subwoofers, vibration elements, and an air-propulsion device. *BioHaptics* explores the usage of small electroshocks that stimulate the user's muscles to simulate force-like feedback, by using a transcutaneous neuroelectrical stimulation device. The *Tactylus* makes use of closely-coupled audio and vibrotactile feedback to provide predominantly tactile information, strengthened by auditory cues for collision detection and texture recognition purposes. This study also looks closely at the interrelationships between visual, auditory and vibrotactile events, by investigating multisensory integration.

This section is followed by an analysis of a prop called the *Cubic Mouse*, introducing a new method for studying the performance of three dimensional placement tasks, resembling Fitts' law tests.

The latter three case studies are focused on combining traditional 2D interaction methods, enabled by a touch screen or Tablet PC, with spatial and potentially unconventional methods. This combination of 2D and 3D interaction methods is regularly referred to as hybrid interfaces. In *ProViT*, macro-level factors affecting the combination of different techniques were studied and partially applied in the *Capsa Arcana* display interaction console. This console combined a touch screen and AV-streaming equipment with unconventional commercial and newly designed sensors to control actions in a virtual heritage application. Finally, the *Eye of Ra* is a pen-like input device that can be used control both 2D and 3D (medical) applications. The highly unconventional form of the device represents a study in human factors and ergonomics, the results of which will be presented and analyzed.

Summarizing, following issues from chapter 2 and 3 are specifically focused on:

- Somatic and kinesthetic feedback (section 2.2.3) in Shockwaves, BioHaptics, Tactylus
- Biopotential systems (section 2.3.6) in BioHaptics
- Multisensory processing (3.2.1) in Tactylus

- Sensory and control substitution, addition and integration (sections 3.3.1 and 3.3.2) in Shockwaves, Tactylus
- Flow of action (section 3.4.1) and feedback mechanisms (3.4.2) in ProViT, Capsa Arcana
- General application and transfer issues (section 3.5) in ProViT, Capsa Arcana, Eye of Ra, and social and ethical issues (section 3.5) in BioHaptics
- Garage Interface design methods (section 3.6) in Shockwaves, Tactylus, Capsa Arcana, Eye of Ra
- Evaluation (section 3.7) in Cubic Mouse

4.2 Shockwaves

This case study focuses on the generation of haptic sensations by using sound and air-based shockwaves (Kruijff and Pander '05). Normally, actuators used with ground- or body-referenced devices are focused on providing one user a realistic haptic sensation by putting an actual force onto the person’s body. During the setup of a new multi-speaker audio display inside a large projection system (the iCone at Fraunhofer IMK), it was envisioned to enrich the capabilities of this display system by generating haptic sensations for groups of people. Hereby, the devices to generate these sensations could not interfere with the setup of the visual display system or burden the users in any possible way. Therefore, the usage of body or ground-coupled devices was not possible. Using what is considered the “third kind of haptic device” (see (Burdea '96)), the tactile display, produces vibrotactile feedback which is in general less accurate than body and ground-coupled devices, but which could potentially be used to generate sensations for more than a single user. Unfortunately, general vibrotactile approaches normally stimulate only small surfaces. Hence, new methods needed to be found. Inspired by rock concerts and discotheques, an initial idea was to somehow “blast” the user away, by using a harmless shockwave propelled at a user. The shockwaves that are examined in this case study are sound and air-based.

4.2.1 Background

From several sources, background information can be obtained on how sound waves and air can affect the human body. These sources are:

- General literature on the mechanisms of acoustics, specifically on *vibroacoustics*. Vibroacoustics is the field of research concerned with the vibrations caused by sound waves in the human body. Much work is performed in the field of vibroacoustics on music, for therapeutical reasons, both at universities and through more “non-scientific” experiments.

- References to work performed on *acoustic weapons*. These deliberately harmful technologies (non-lethal weapons) investigate the effects of, among others, the usage of infrasound, and the creation of so called “sonic bullets”, being sonic energy propelled towards a target. Research performed in this area provides a good overview of ergonomic considerations useful for the experiments carried out in this case study (Cook, Fiely et al. '95) (Altmann '01)[1].

- Work performed on *air propulsion devices*. This work is predominantly practically oriented and can be found in several areas, including machines used to propel small or light objects like confetti or smoke and mechanisms used in air guns and cannons. The latter involves several techniques that are, just like the previously mentioned acoustic weapons, not suitable since they can harm a user. Nevertheless, potentially useful methods can be obtained, if adapted appropriately.

In order to understand the different effects of sound and air shockwaves, more detailed on these issues will be given.

Sound-based shockwaves

As stated before, sound based shockwaves are predominantly researched in the fields of vibroacoustics and acoustic weapons. For this case study, it is important to differentiate between *acoustic* and *tactile* components of sound waves. The acoustic component, mainly sensed by the auditory system, is the sound we “hear”, whereas the tactile component, which can be sensed by multiple body parts, is concerned with how we “feel” the sound.

Sound waves can be sensed via transcutaneous sensing, bone structures (bone conduction), and via the cavities of the human body that pick up the sound frequencies (Cook, Fiely et al. '95; Altmann '01). Sound waves can generate anywhere from vibrations up to shocks in the human body (Figure 4.1). These effects can be harmful, but can also have positive effects. Vibroacoustic effects are believed to have positive therapeutical purposes. Probably pioneered by Olav Skille in the 1980ies, physiological vibration is sometimes used to control psychological events, lessen stress, and provide muscle relaxation and pain relief. Multiple sources on this work can be found in the Journal of Music Therapy (AMTA '06). Work at universities includes efforts such as the Music Vibration Table (Chesky and Michel '91). A multitude of commercial products like music chairs and beds are available, the effects of which cannot always be proven in a strictly scientific way.

Most of the therapeutical work is concerned with lower-intensity sound waves. In order to generate haptic sensations, it is important to understand the effects of different frequencies and higher intensities on the human body. Figure 4.1 provides a basic overview, based on (Altmann 2001), describing some advances in the development of devices that make use of acoustic waves affecting the human body.

The table provides some insights – it can be concluded that some frequencies cause resonances in the human body that can be felt, probably as vibration-like sensation. The actual frequencies (ranging from about 2Hz to 2.5KHz) seem quite wide, even though indications can be found that frequencies below 100Hz can produce useable effects. The intensity of the sound wave needs to be tuned exactly – when the intensity will be too high, important side effects (health issues) will become evident. The auditory “pain level” lies around 130dB – it is clear that any usable usage of sound for producing haptic sensations need to be far enough below this sound level.

---

[1] The author hereby clearly states that the interest in acoustic weapons is clearly based on its results showing ergonomic effects on the human body. In no way it is intended to reproduce any of the lethal effects produced by such weapons in a VR projection system.

Inferring from the information in the table to a general-purpose projection system setup, we need to be looking at intensity of between 85 and 100dB, since increasing the intensity will probably result in a system that is too loud. Keep in mind that a front-row in a rock concert will be around 110dB, about ten times as loud as an 100dB sound source.

Sound source	Frequency / intensity	Reported effects
Infrasound	Below 20Hz, high intensity (up to about 150dB)	Resonances in inner organs, vertigo, imbalance
Sonic Boom	2-20 Hz, extremely high intensity (up to 170dB peaks)	Resonances in body (air) cavities, resonances in inner organs, extreme health issues
Acoustic pulse	Around 10Hz, focused beam, extremely high intensity	Possible knock out
General low frequency sound	Mainly 20Hz – 100Hz, up to 2.5 KHz, higher intensities (up to peaks of +/- 135dB)	Resonances in different body (air) cavities, resonances in inner organs

Figure 4.1: *Effects of different sound sources.*
After (Altmann 2001)

Air-based shockwaves

The basic idea behind air-based shock waves is to blow balls of air at a user. Most techniques to produce these bursts of air are based on pressing some air to a small(er) opening. The devices are generally known under the name *vortex generator*. Using a box with flexible backside and a hole, when the volume decreases by pulling the flexible backside, the pressure increases. This action forces some of the air out of the hole. The velocity at which the air leaves the box is inversely proportional to the diameter of the hole; the smaller the hole, the greater the velocity of the air. A good overview of vortex generators can be found in (Beaty '04).

AirZooka™ is the "fun gun" that blows a harmless ball of air towards any object or person.
Laugh with amusement as, seemingly from nowhere, you are able to mess up a person's hair, ruffle their shirt, dress or office papers from a distance, or just plain BLOW 'em away !!!

Requiring no batteries or electricity, AirZooka operates simply by pulling and releasing a built-in elastic air launcher. And here's the best part: because it shoots air, you'll never run out of ammo! AirZooka also features a pop-up site for improved aiming. Try your luck at indoor target practice, or take aim at your friends with surprising accuracy !!

To Operate AirZooka:

1) Flip up site (A) to take aim.
2) Firmly grip the handle (B) with one hand while gripping the elastic air Launcher (C) with your other hand.
3) Fully extend your arm and pull straight back on the elastic air launcher. Keep Launcher centerd, and pull back.
4) Release the elastic air launcher to launch a powerful yet harmless ball of air!

Figure 4.2: *The AirZooka air cannon.*
Courtesy of ZeroToys

The more harmless versions make use of a rubber-like surface that is pulled back and released in order to change the size of a volume (children’s air cannons, Figure 4.2, (ABC '06)). The same effect can also be achieved by mounting a loudspeaker on the back of a box or cylinder, making use of the movement of the membrane. The higher intensity air propulsion devices almost all make use of some kind of pneumatics, either to propel small objects like confetti or smoke or to shoot projectiles (airgun).
Just like wind, the air is sensed by the movement of hairs on our skin, thereby being a cutaneous sensation. Extreme blows of air can result in slight deformation of the skin, thereby activating the transcutaneous sensing mechanisms.

Hypotheses
Resulting from the searches through background literature and personal experiences, it was hypothesized that:

- A combination of specific sound frequencies and intensities might generate a haptic sensation inside the user’s body
- Sound effects could be sensed by the body and interpreted as pseudo-haptic sensations
- The usage of a large-area vibrotactile display could easily provide some kind of haptic feedback based on vibration focused on the user’s feet
- The usage of air-based shockwaves might provide an additional kind of haptic sensation on top of using sound waves.
- A combination of different methods could lead to a way of providing both a somehow effective and entertaining (exciting) kind of haptic sensation for groups of users

4.2.2 Related work

The work performed in this case study is directly related to the body of research performed on haptic feedback. A general overview of haptic feedback is provided in section 2.1.3 or can be found in (Burdea '96). More specifically, research related to vibrotactile displays was a source of inspiration. Some of these sources are (Tan and Pentland '97; Okamura, Dennerlein et al. '98). The developed system described in the next section was inspired by the usage of vibrotactile elements in immersive displays to simulate earthquake vibrations (Dombois '02). The concept of “blasting” the users body, by stimulating larger areas of the body has been tried in experimental systems. Examples include tactile vests using vibrotactile actuators or contractile shape-memory alloys (Jones, Nakamure et al. '04), and the Aura Interactor (Figure 4.3, not produced anymore). The latter device is basically a large loudspeaker mounted directly on the torso of the user, vibrating heavily when activated.
The usage of alternative ways of providing haptic sensations has been greatly inspired by Yanagida’s air cannon (Figure 2.17, (Yanagida, Kawato et al. '03)), which lead to the exploration of loudspeaker-based air cannons. This device makes use of a small loudspeaker placed in a cube with a narrow opening to generate airstreams that hold different smells, in order to provide localized smells to a user. A non-electrical version of the air cannon using the same principle (air flow) as Yanagida’s device is sold by Zero Toys, called the AirZooka. Another toy made in the 70ies, called the Wham-O Blaster, also makes use of the same principle, but is not sold anymore.

Figure 4.3: *The Aura Interactor shock vest*

4.2.3 Interface design and development

The particular case at hand was to install a new audio system in the iCone display environment at IMK.VE (Figure 4.4). The iCone is a spherical visual display of about 5,5 meters diameter, with a 240 degrees field of view and slightly tilted walls (7 degrees to the back), giving it a conical form. One of the advantages of the conical form is reduction of acoustic reflections, making it better for sound display than normal spherical display systems. The previously used 9-speaker system (using Genelec 1030 speakers) was to be replaced by a system that would support full surround sound. Therefore, new speakers where acquired (Genelec 1029a), that where to be placed in a spherical setup, in a top and bottom ring. With the new setup, subwoofers were also installed. Previous listening tests concluded that the lower end of the sounds spectrum needed some support, since the Genelec 1029a does not perform well in the lower frequency range.

Figure 4.4: *iCone display system.*
Courtesy of Fraunhofer IMK

In combination with the incentive to use subwoofers to generate haptic experiences, a first subwoofer was built to test the effects of low frequency up to subsonic sounds (below about 20hz) on the human body. As stated before, theories on the effects of audio sources show that specific frequencies between about 2hz and 100hz have an effect on the human body, for example by vibrating organs. Especially organs with volume like the lungs, will start to vibrate at specific frequencies, producing a haptic-like sensation. Hence, a system producing a good amount of low frequencies was needed.

The first, prototypical system was called the SoundMaster 2000 (Figure 4.5). It was built of three massive Visaton 17" (43 cm) low frequency speakers built in a triangular MDF casing and placed in the centre of the iCone. Initial experiments showed that the low frequencies produced an acceptable range of sounds that could be combined with the Genelec 1029a speakers for normal sound production. A wide range of sound frequencies were produced to test the effects of subsonic and lower frequencies on the human body. However, it only produced minimal haptics through vibration of the floor and only very little haptic sensations by having an effect on other body parts than the feet. Assuming that the volume of the speaker casing was too small, a tile of the double floor was taken out, whereby the speaker was placed over the opening. Through this, the resonance volume was enlarged considerably, taking effect of the large double floor space. Nevertheless, the loudspeaker did not produce the wanted haptic effects. Furthermore, the construction was occupying too much space in the display system.

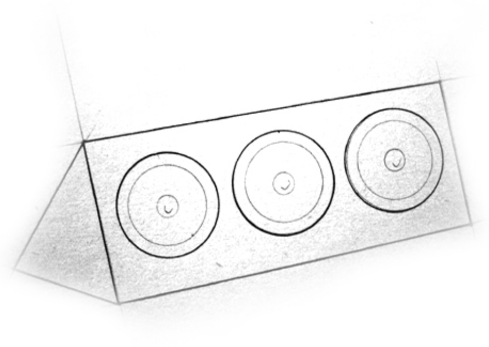


Figure 4.5: *Artist impression of the SoundMaster 2000 subwoofer installation.*

The next idea was to take the large subwoofer apart, making multiple subwoofers to be placed outside of the iCone display. Hereby, 4 loudspeakers (Figure 4.6) where built with a size of about 18" x 18" x 18" (50x50x50 cm). The volume of the box is about 125 liters which for a closed box results in a resonant frequency of approximately 56 Hz. Holes where made in floor tiles so that two of the subwoofers, placed just outside the iCone in the left and right corner, could exactly fit with the speaker into the floor. The floor in the iCone is a double-layered floor made of tiles on a concrete floor. The space between the tiles and the bottom floor is about 15 cm. The total volume is about 1850 liters between tiles and concrete floor, resulting in a resonant frequency that would be around 31 Hz. A second pair of speakers was placed behind the iCone, taking advantage of the reflection of the back corner for producing low frequency sounds. When placing loudspeakers in a corner the lower frequencies radiating in a circular pattern are reflected by the walls and sum up with the non-reflected sound waves. This results in directed sound waves with higher intensity. The distance of the speaker to the wall must be less than 1/4 of the desired wavelength to make optimal use of this effect.

Figure 4.6: *Single and stacked subwoofers.*

Using sound samples and music files, the loudspeakers where calibrated first to produce a balanced audio spectrum together with the Genelec 1029a speakers. In a second step, it was tried to generate sound-based shockwaves. Hereby, the loudspeakers proved highly effective to produce the lower basses – basically, using higher volumes, a large part of the building was slightly vibrating. Due to the large double floor space used as extra resonance volume, all rooms that are at the same level, and below (cellar) the double floor noticed considerable lower frequencies (vibrations). Nevertheless, the shocks could not be directed towards the user's body, and where mainly floor vibrations. Additionally, the audio waves did not have a maximum in the centre of the iCone (the hotspot where most user's will be immersed), since the vibrations where felt most strongly at the border and outside the display system. Hence, the sound waves had to be adapted in order to produce the necessary results in the middle of the iCone. Hereby, the subwoofers standing in the left and right corner where delayed for about 7ms, to create a maximum for frequencies around 30 Hz in the centre of the iCone. This resulted in vibrations and a clear experience of lower frequency sounds (related to the sound samples) in the hotspot of the display system. Tests where run, whereby low frequencies where generated at around 95-100Db. In the software the sound spectrum was adapted to produce more lower frequencies by applying separate filtering on the subwoofer channels. Nevertheless, this adaptation would not disrupt the generation of sound for other applications using the display systems.

Even though the lower frequencies could be sensed in the body, mainly through floor vibration, its effects where unsatisfactory. It generally produced a deep “bass sound”. Since higher intensities (above 100dB) would generate too much vibration inside the building and could de-calibrate the display hardware (projectors), the idea of producing audio shockwaves by using solely subwoofers at considerably high intensities was abandoned.

Following up the previous success by colleagues with vibration elements for simulating earthquake vibrations, a next step was to install a large area vibrotactile area in the iCone. Because the height of the floor is limited it was not possible to construct a vibration floor without keeping the centre of the iCone leveled with the rest of the room. Hence, in the centre of the iCone, five floor tiles where mounted with Paraseats (Figure 4.7). Paraseats are 25W loudspeakers (“tactile transducers”) exclusively producing low frequencies. Paraseats are extremely effective for producing vibrations: Mounted under the floor tiles, the surface produces clear vibrations, up to small shocks.

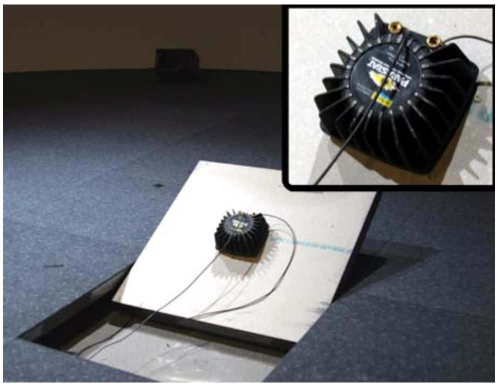


Figure 4.7: *Emphaser Paraseat 25W vibration element mounted under a tile of a double floor.*

The vibrations were in sync with lower frequencies generated by the subwoofers, thereby producing a nice experience of “bass sounds.” The Paraseats are driven through a separate audio channel in the software, so special effects like rumble or other vibrations can be directed by the application. A similar setup was in use in the CAVE at the IMK-VE group. The original AVANGO Cyberstage Soundserver has the capability of using a separate channel for driving vibrations elements (Eckel '99).

Figure 4.8: *Air cannon prototype.*

Leaving the trail of sound-based shockwaves, experiments were started using air-based methods. Inspired by different vortex generators, several designs were analyzed. Encouraged by designs using loudspeakers mounted in a cylindrical volume, a small air cannon was produced (Figure 4.8). The air cannon was made of a 10” mid and low frequency speaker from Visaton, mounted in a recycling bin. On top, a round wooden plate was glued, with an opening through which a small plastic pipe (about 2” / 5cm diameter, and 11” / 28cm) was placed. For testing, the loudspeaker was connected to a 9V battery. Testing the device, it produced only slightly noticeable air balls, up to a distance of about 2 feet.

4.2.4 Evaluation and results

The setup and usage of the different developed components has been probed over a period of over two years. Through multiple sources, the following sections report our experiences with the different setups, and discussion of the effects. These sources are:

- Personal observation and experience of audio experts
- Direct and indirect feedback (observations) from several hundred users experiencing demonstrations inside the display system, including direct feedback from VR experts
- Acoustic measurements (spectral analysis)

From its first setup, the acoustic quality of the system was never an issue. Most of the direct feedback confirmed the richness of the sound in the system. It produces crisp higher frequencies and deep basses. Users comment especially on the effects of the sound floor; Acoustic effects like a heartbeat in a medical demo can be clearly experienced via vibration. Most users do not directly realize the haptic effects of the system, but just seem to enjoy the tactile characteristics of the sound display. The user's body is clearly shocked by the vibrations of the floor, generated by both the subwoofers and the Paraseats. Thus, most tactile/haptic effects seem to be generated by conduction of vibration from the feet, through the bone structure of the user. By using separate channels for low frequency reproduction, it is possible to tune the maximum for certain locations in the room. However, these locations are only valid for certain frequencies. The central area of the iCone is tuned for frequencies between 30 and 60 Hz. Using a single subwoofer, this kind of effect would be difficult to realize.

The “blast effects” that are reached by high intensity sound environments (discotheques or concerts) could not be fully repeated inside the display system, due to several reasons as stated in the next section. We could increase the sound level to such a point that the full building (including users) was vibrating, but under general usage, this is not acceptable. Effects as stated in the table on effects of low frequency sounds could not be fully tested or reported on; Body cavities like the chest of users where vibrating, but only at a lower noticeable level.

Measurements have been done for calibrating the loudspeaker setup. This is part of the normal setup procedure for lining up loudspeaker arrays. Measurements showed that there are a few resonant frequencies due to the nature of the room and the display setup. These were around 400 and 800 Hz. Filtering has been applied to all audio channels to attenuate these areas in the spectrum. Furthermore, the Genelec 1029A have the tendency to sound a bit “sharp,” which results in the undesired effect that the listener can hear where the separate loudspeakers are located, therefore a bit of attenuation of the higher frequencies is also applied.

The air cannon was evaluated negatively. Since it only produces slightly noticeable pops of air, within a distance of up to maximum 2 feet, it was unusable in the display system. A newer version of the system is under design, but has not yet been tested.

4.2.5 Reflection

The setup of the different devices has been a trial-and-error process. Starting with the subwoofers, we were able to produce an incredible amount of lower frequency sound for the size of space of the display system. Using the sub floor space as resonance volume is a good way of generating deep basses and vibrations, but its effect is largely non-directional. Through the additional usage of the Paraseats, users clearly notice shocks coming through the feet, being conducted by the bone structure throughout the body, into the stomach and the chest. To these vibrations, the user does not connect any specific direction (directional cue). Therefore, in interfaces, the vibrations could only be used for general purposes. The validity of the vibrations, therefore, mainly seems to be to enhance a sound system by making sound effects tactile. The tactility can be applied for simulating the effects of collisions, like running into a wall. Additionally, the vibration seems to provide motion cues. The most plausible explanation for this effect is that a part of the inner ear is sensitive to vibrations above a certain level (90dB) and generates motion sensations.

For reasons stated before, we were unable to “blast” groups of people using low frequency sounds. Even though the vibrations are impressive, intensities that are high enough to really shock the complete the body of a user (in case of group experiences) is not possible within smaller or mid-size presentation theatres, using the taken approach.

In addition, we can report on a lack of any negative side effects of using the system, to date. ISO 2631 (ISO '03) reports that vibrations from 1 up to 80Hz can involve health issues like drowsiness or stomach problems. No user of the system reported such issues. In addition, the demonstrated methods surpass previous efforts solely using Paraseats or similar bass shakers, by providing a wider range of frequencies, thereby stimulating more organs or bone structures than normally achieved with the bass shakers solely. Hence, the sensation is richer and can be better controlled.

The first attempts of coupling audio and air-based shockwaves were futile. Due to the problems with the first prototype of the air cannon, no air balls could be propelled to the user. Nevertheless, the concept seems realizable, if enough devices can be found that effectively propel air to user. The problem with most devices is their size. A pneumatic air cannon using multiple nozzles, mounted under floor tiles was conceptualized, but has not yet been built.

Reflecting chapter 3 issues, the test successfully showed how haptic feedback can be provided using alternative way by using sensory substitution methods. Bone conduction and the vibration of human organs provide a useful way of delivering haptic-like feedback without having to use grounded haptic devices. The actual quality of the feedback is slightly blurred by the informality of the evaluation: a formal user test with a large user group, which was impossible to perform to date, needs to be carried out to understand the effects of the feedback in a more precise way.

4.3 BioHaptics

This case study (Kruijff, Schmalstieg et al. '06) focuses on the usage of electro stimulation methods to create voluntary, but non self-induced, muscle contractions. These contractions can be perceived as pseudo-haptic feedback. Ultimately, muscular stimulation could be used to create new poses through focused triggering of specific muscle endings to certain extent. Within the current case study, the user attitude was tested. In addition, observations of users’ physical reactions to this rather unconventional form of feedback were made, leading to a discussion and roadmap for further development.

4.3.1 Background

Within the medical and sports areas, the usage of electrical stimulation devices has been widely used for pain relief and muscular training. Transcutaneous electrical nerve stimulation (TENS) is generally applied to stimulate nerve endings in order to block pain (Johnson '01), whereas neuromuscular electrical stimulation (NMES) is widely used for training muscles, both in the sports area and for rehabilitation purposes (Alon and Smith '05; Porcari, Miller et al. '05), possibly aided by virtual reality aided methods (Steffin '98). Both methods are based on the electrical stimulation of nerves or receptors, using impulses at different frequencies and intensities.

Only recently have researchers started to explore the usage of electronic stimuli for triggering somatic and kinesthetic events for human-computer interfaces. The somatic and kinesthetic systems handle the sensations that relate to force and touch. The somatic system perceives cutaneous (skin) and subcutaneous (below skin) sensations, whereas the kinesthetic system senses mechanical sensations in the joints and muscles. These sensations are generally known as haptic feedback and relate to the communication of information on geometry, roughness, slippage, and temperature (touch), weight and inertia (force). Skin and muscle sensations are received by several receptors: thermoreceptors, nocireceptors, and mechanoreceptors, including proprioceptors and chemical receptors. Through electrical stimulation, theoretically every kind of nerve, nerve ending, or receptor can be triggered, depending on the kind of stimulus provided to the user. These stimuli differ in pulse length, frequency, amplitude and triggering mode.

In order to provide neuroelectrical stimulation, electrodes are placed on the skin’s surface. Implantable solutions or methods using needles also exist, but are not generally used yet. An electrical stimulus is able to reach a nerve or receptor due to the permeable properties of the tissues below the skin (Sörnmo and Laguna '05). Under effect of specific ionic substances in cells, membrane potentials can be generated that flow through the surrounding tissues, eventually resulting into a pseudo-haptic event through stimulation of the motor nerves. This permeable characteristic of the skin tissues is also used for biopotential interfaces, by reading so called action potentials. An example of an interface using action potentials is the electromyography (EMG)-based joystick by Jorgensen et al (Jorgensen, Wheeler et al. '00).

Within this study, a closer look is taken on force-related events that can be triggered using electrical stimuli. These stimuli are envisioned to cause pseudo-haptic events by changing the users pose through voluntary, but not self-induced, muscular movements. Using NMES-based methods, wearable interfaces can be built that surpass limitations with body or ground-referenced devices, such as cost and immobility. By means of an initial user test, first experiences have been made with electronically triggered muscle events. These events will be discussed by illuminating the physiological background, deducing useful factors in building up NMES-based haptic interfaces.

Hypotheses

- Using electric stimuli can be used to control muscle activity in order to provide haptic feedback
- Exact control of the stimuli might be difficult
- Users might react to electric stimuli as a rather experimental way of feedback and disapprove its usage

4.3.2 Related work

Using electric stimuli to generate haptic or pseudo-haptic feedback is a largely unexplored area. The usage of TENS devices and methods has found wide application in medical scenarios, for pain relief (Johnson '01), but has not found particular usage in haptic-related applications. Similarly, electro muscular stimulation is widely used for training muscles, both in the sports area, and for rehabilitation purposes (Alon and Smith '05) (Porcari, Miller et al. '05). A virtual reality system developed by Steffin (Steffin '98) made use of muscular stimulation methods for helping people with multiple sclerosis and spinal cord injury to increase accuracy of movements. Folgheraiter et al experimented with a glove interface, applying basic muscular stimulation methods for providing haptic feedback (Folgheraiter, Gini et al. '05). Using electrical stimulation for tactile purposes has been done on several occasions, including the finger-mounted electrotactile system by Kajimoto et al. (Kajimoto, Kawakami et al. '99) and the tongue-based electrical stimulation by Kaczmarek et al (Kaczmarek, Weber et al. '91), which was used for sensory substitution purposes. A rather strange art installation was made by Elsenaar (Elsenaar and Scha '02), in which face muscles could be controlled remotely to generate different kinds of facial expressions, by usage of small electrodes placed around key muscle endings. Finally, some game environments have experimented with electrical input to the skin.

Figure 4.9: *Bioforce controller from Mad Catz.*
Courtesy of Mad Catz

In 2001, Mad Catz (MadCatz '06) announced a controller called Bioforce (Figure 4.9), which delivered electric impulses to the user’s forearms. Stimulation would result in slight spasms in the forearm, up to larger shocks possibly letting the player loosen grip on the controller. Till now, the product has not gone into production. Finally, the PainStation (PainStation '05) induces small electroshocks to users playing pong, focusing on pain oriented feedback as a penalty during game play (also see section 2.2.3 and Figure 2.15).

4.3.3 Experiment setup

Due to the novelty of the method, an initial user test was prepared to create a basic understanding of the effects of neuroelectrical muscular stimulation. Within this test, muscular behavior (contractions) and possible side effects such as pain were observed, along with the user's attitude to this rather unconventional kind of feedback. The muscular contractions were caused by surface electrodes attached to either the biceps (forearm) or the brachioradialis (lower arm). The experiment made use of a non-immersive setup (Figure 4.10), resembling the setup proposed by MadCatz as a setup for games involving force feedback. The MadCatz setup, though, solely triggered the brachioradialis, whereas (as stated before) we also triggered the biceps.

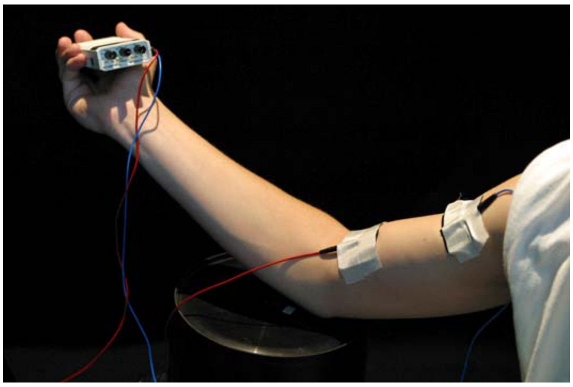


Figure 4.10: *Connected TENS with electrodes.*

A basic 3D environment was used (Quake3), running on a laptop with a 14" screen (Figure 4.11). Users interacted via standard input devices, a mouse and keyboard, to control the game. The used electrodes were connected via cable with a TENS device (9V), a Schwa-medico SM2. Seven users, of which six were male and one female, participated free of will in this test, being informed of the possible health issues of using the system. The users had widely varying anthropometric characteristics.

Figure 4.11: *Test setup showing electro pads and display with Quake3.*

The evaluation consisted of three stages. In the first stage of the experiments, the TENS was used to identify at which stimulation level the user would obtain feedback (muscle contraction) without causing pain, thus calibrating the device for each user. Using contact fluid, the electrodes were placed on either the biceps or brachioradialis muscle endings and were not moved until the end of the experiment. The device was put on a low pulse rate (3 - 5 Hz, biphasic), after which a range of short and longer (up to 3 seconds) pulses were given to the muscles, to come to an appropriate stimulation level. The resulting maximum intensity in continuous mode turned out to be between 10 - 15 mA, up to short shocks of up to 25 mA. The maximum stimulation level differed between users and was clearly dependant on the muscle fat level and thickness of the arm: thicker skin and muscle tissues resulted in higher Ampere stimulation levels.

4.3.4 Evaluation and results

Biceps stimulation (four users) could be clearly noticed on two users. For the other two users a calibrated stimulation level was used at which contraction was minimal. The reaction of the muscles when placed at the brachioradialis (three users) could be better observed. With one user, a clear spasm in one of the fingers could be seen, probably caused by the triggering of a different muscle than intended. As a result, some changes in the biomechanical configuration (pose) of the arm and fingers were triggered and could be observed. The levels of input (up to 25 mA) were not expressed as being painful. Users expressed slight discomfort or some (mental) excitement and never seemed to loose grip on the input device, as was previously stated in informal statements on the usage of the MadCatz Bioforce.

The second phase of the test focused on establishing small muscle contractions during game play. Whenever the user would get shot, a short stimulus (when hurt by explosive weapons) or continuous stimulus (when hit by a gun) would be given to the user. Users played a single round, lasting approximately 10 to 15 minutes. The TENS device was triggered manually, which resulted in a small delay to the feedback. Only one subject reported negatively on this delay. Due to the observation angle, the observer could observe both the game play and the muscle contractions without having to switch focal direction and, therefore, attention, between the two. Hence, the observer could get a clear impression of both.

As within the first phase, muscle contractions could be clearly noticed with most users, especially when stimulation would be provided in continuous mode. When the biceps was triggered, the change of pose (noticeable in the change of the elbow arc) was, at a maximum, around 10 degrees, but regularly just a couple of degrees. During stimulation of the brachioradialis, contractions mostly lead to a change in the pose of the lower arm and hand. Contraction was not always completely continuous; A higher pulse rate could improve contraction, since it would not allow the muscle to relax.

After the second phase, the users were questioned about their experience, using a questionnaire with a 5 point Likert scale.

The results are visualized in Figure 4.12 (higher scores are better). The stimulation did not result in painful reactions; All users expressed that the feedback was at most uncomfortable (avg. 3.14, stdev 0.37). This result confirmed the calibration results and the observations during game play. The users stated they had no to limited reaction loss (avg. 4.29, stdev 0.95). One user said that the stimulation was “irritating”, when stimulation was provided in continuous mode for a considerable amount of time (5-10 seconds). In all cases, the electro stimulation was noticed as “somehow noticeable feedback” (avg. 3.00, stdev 0). 5 of 7 users found the stimulation to be funny or interesting; one user was even pretty excited. Two users did not find the stimulation exciting at all (note 1), which resulted in a diverse reflection (avg. 2.57, stdev 1.13).

Finally, the reactions on usefulness varied widely (avg. 3.42, stdev 1.86). Three users reacted extremely positive and stated a “definitive” desire to use the system further, also outside the games area. Two users reacted negatively (the same users as reacted negatively on excitement) on the feedback and certainly did not recommend usage outside entertainment purposes. These users also had most problems with user comfort. Finally, as expressed in direct discussion, none of the users had any problem relating the feedback to the game play. The “shock”- like feedback could clearly be connected by the actual event of getting hit.

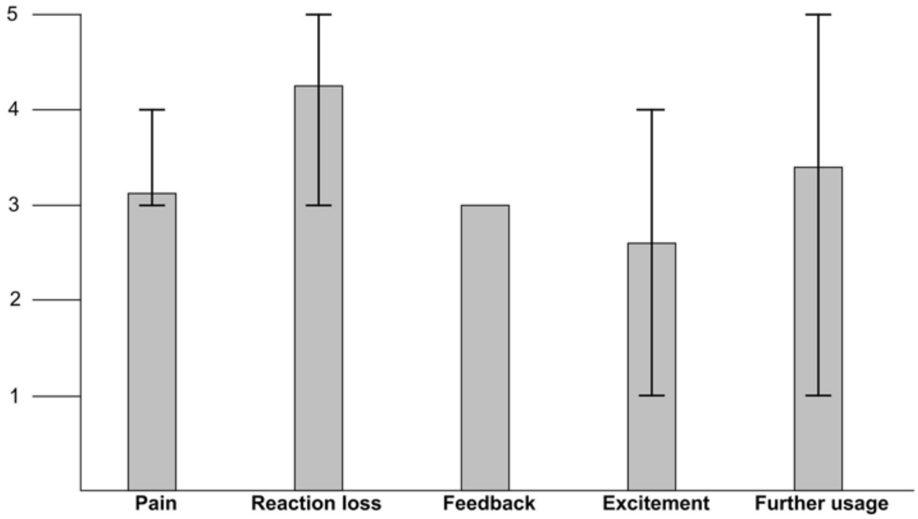


Figure 4.12: *Evaluation results.*

4.3.5 Discussion and roadmap

The evaluation showed various issues that relate to potential, but also the problems of using neuromuscular electrical stimulation for pseudo-haptic feedback. Theoretically, neuroelectrical muscle stimulation can produce muscle contractions that can lead to the same kinds of movements as performed voluntarily using self-induced muscular activities. It is to be expected that stimulating solely the arm will not provide the full spectrum of movements such as afforded by methods like an exoskeleton, simply because some movements of the arm are also triggered by the shoulder, hence also using muscles at the back of a user.

The contraction is hard to control; One needs to trigger specific muscles to the right extent in order to create the wished change in the biomechanical configuration (pose) of the user. Furthermore, there may be a conflict between movements caused by electro stimulation and the body-internal voluntary signals that may level out muscle activity, depending on the thresholds of the potentials triggering the muscles. Thus, it may well be that the voluntary control rules out some kinds of feedback or results in distinct parameter changes in the NMES-based control, like the increase of intensity to overcome a specific potential threshold. However, there are no current experiments known to the authors that deal with these problems.

Model of muscular behavior

The usage of NMES-based feedback is centered on the creation of a precise model of muscular behavior, under effect of electrical stimulation, showing how the biomechanical configuration in the arm reacts to different electrical impulses in order to create directional feedback. This model needs to integrate effects of conditioning of the muscles, since stimulation effects will change over time (Alon and Smith '05). Electrical muscle stimulation applies low frequency, higher intensity pulses. These pulses, which are mostly biphasic, trigger the alpha motor nerves, which excite the muscles. This stimulation leads to a non self-induced contraction of a muscle. The higher the intensity of the stimulus, the more muscle fibers will be excited, leading to a stronger contraction (twitch). A twitch can have different contraction speeds (explosivity). The duration of the contraction depends on the frequency of the impulse; When the frequency is high enough, the muscle will not have time to relax, thereby staying continuously contracted.

The model depends on the effects of stimulation resulting in both isometric and isotonic muscle contraction. Isometric muscle contraction leads to a tension in a muscle, without changing the length of the muscle, whereas during isotonic muscle contraction, the muscle does shorten. Thereby, muscles can be classified in four different functional groups (Goldstein '02): prime movers, antagonists, synergists, and fixators. The different muscles play an important role in the lever system that characterizes the bone-muscle relationship. This mechanical system defines the force or effort to balance a load, moving on a fixed point, the fulcrum. Most movements in the human body function according to a lever system. Excitation of a motor neuron by the nervous system produces exactly the same result as when provided through electrical stimulation; However, now, the brain and spinal chord are not involved in the muscle activity.

Calibration methods

In order to stimulate the muscles to the right extent, and to ensure a high level of user comfort, there is a strong need for exact calibration methods. One of the problems noticed during the experiment is the trade-off between high intensity stimulation resulting in noticeable pose changes and user comfort. Different effects of similar stimulation, caused by the different anthropometric variables such as arm tissue thickness or the level of skin hydration could be noticed.

Calibration will most likely require creating an exact stimulation – biomechanical change model by tracking the user’s arm, via an exoskeleton or bend sensor(s). Information from these sensors can also be useful for real-time controlling and adaptation of stimulation, thereby also take care of conditioning of the muscles.

Triggering of skin receptors

Another issue which should be dealt with is the triggering of skin receptors. During the experiment we had the impression that not only the muscle endings were stimulated, but also specific cutaneous or subcutaneous receptors. The triggering of these receptors might also have caused the feelings of pain or “buzzing,” as was sometimes noticed by the subjects. The triggering of skin receptors can go into two directions: either avoiding the receptors to be triggered to prevent unwanted side effects or to deliberately trigger the receptors to create specific tactile sensations.

Electrotactile stimulation can focus on one or multiple of the six available receptors that can be found in either glabrous or hairy skin. The receptors have different receptive fields (1-1000 mm2) and frequency ranges (0,4 – 800 Hz), producing diverse sensory correlations. The receptors can roughly be classified according to the speed of adaptation to a step change in applied pressure to the skin (Kaczmarek, Weber et al. '91). There are fast adapting broad receptive-field receptors, like the Pacinian corpuscle, producing vibration tickle sensations, up to slowly adapting, small field receptors such as the Merkel's cells, handling pressure sensations. Within the body, the fingertips are by far the most sensitive, having a high spatial resolution. Not surprisingly, most haptic interfaces focusing on electrotactile feedback stimulate the fingers. Electrotactile stimulation and perception is rather difficult and does not necessarily lead to unanimous results; Depending on the stimulus characteristics (intensity, waveform) electrode size and material, and skin characteristics like thickness and hydration, perception may range from tickling, buzzing, beating, pressure, up to pain. Thus, a model of electrotactile stimulation should be carefully coupled to the model of muscular stimulation, thereby taking care of anthropometric variables.

Wearable hardware setup
A final issue that should be regarded is the actual hardware setup of the system. As can be concluded from this article, neuroelectrical stimulation is well suited for lightweight installations, since there is no dependency on large body or ground-referenced devices. Hence, a wearable and thereby ergonomic and easily installable system could be developed. The problem is dealing with anthropometric variables: there is no one size fits all solution, since electrode placement most likely differs between users. Electrodes sewn in cloth-like constructions show good results, but more progress needs to be made (Jorgensen, Wheeler et al. '00).

4.3.6 Reflection

Concluding, in this case study, a novel way for providing pseudo-haptic feedback by using neuromuscular electrical stimulation methods was presented. An initial user study was performed and an extensive physiological discussion addressing specific problems, which will hopefully lead to further investigations.

Reflecting chapters 2 and 3, the BioHaptics technique presents a good way of using biopotential to obtain feedback from a computer. It applies a similar way as control using EMG, since muscular activity can be both used as input and output method. The evaluation showed the experimental and unconventional characteristics of the new method, but is still believed to have great potential for further research following the provided roadmap. Using a more advanced system, a more comprehensive user study needs to be performed, covering a wider amount of factors than the user attitude and initial analysis of physical activities of the muscles. Potentially, an EMG can be used to come to exact results. Thereby, a combination of both EMG for user output, and muscular stimulation as input method can be a powerful new way of interaction.

4.4 Tactylus

Recent studies have shown that through a careful combination of multiple sensory channels, so called multisensory binding effects can be achieved that can be beneficial for collision detection and texture recognition feedback. This case study (Kruijff, Wesche et al. '06) focuses on a new pen-input device called Tactylus, which explored multisensory effects of audiotactile cues to create a new, but effective way to interact in virtual environments, with the purpose of overcoming several of the problems noticed in current devices.

4.4.1 Background

Many virtual reality (VR) setups make use of some kind of pen-input device to control a VR application. When observing frequently used pen-input devices, some ergonomic problems can be noticed, as well as control and feedback issues. These issues become especially evident and problematic in more complex task environments. A pen-input device is generally regarded as a time-multiplexed device. Multiple tasks are performed in a serial manner, thereby resulting in a composite flow of action. However, the control possibilities do not always match the complexity of an application, regularly leading to the usage of a system control technique to change the mode of interaction, or the usage of a second device. Most pen-input devices (such as produced by Polhemus or InterSense, see Figure 4.13) we evaluated are not very ergonomic when analyzing grip and balance issues. Using these devices, manipulation tasks are not always easy to perform, and they might benefit from more advanced feedback mechanisms besides visual and self-maintained feedback, especially in visually challenging (occluded) tasks. Finally, we noticed an increased demand on pen-like devices to support integration in hybrid interaction setups.

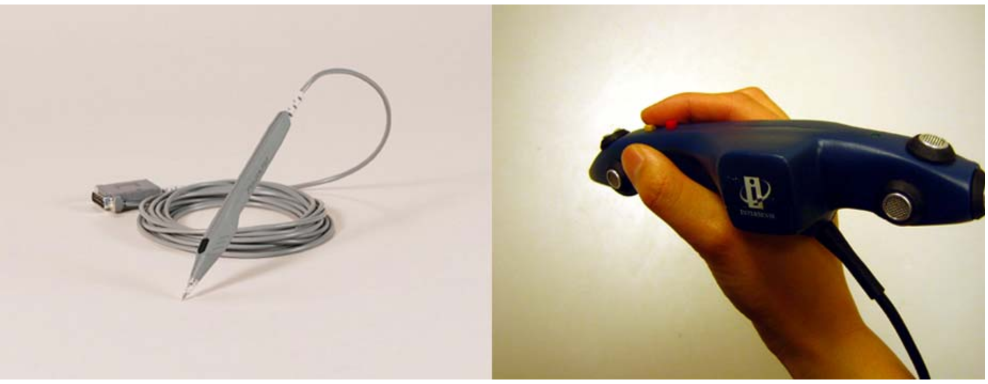


Figure 4.13: *Pen-input devices from Polhemus and InterSense. Courtesy of Polhemus and J. Chen*

In order to potentially solve the previously mentioned problems, several issues where addressed during the design of a new input device, called the Tactylus. We created an ergonomically improved form, limiting fatigue problems, and increased the number and diversity of buttons to improve the quality of system control, including the support for hybrid interaction. Finally, we searched for apt feedback solutions to increase performance, resulting in the application of sensory substitution methods (Kaczmarek, Weber et al. '91) by using audio and vibrotactile cues to simulate haptic feedback. A key issue was to analyze multisensory effects of closely coupled visual, auditory and tactile cues. Recently, some research (handled in detail in the next section) has been focusing on the effects of a process called multisensory binding. The usage of congruent signals can lead to a better perception of roughness of textures, up to perceptual modulation that may serve as a collision detector (Shimojo and Shams '01;

Weisenberger and Poling '04). The objective when designing the Tactylus was to take advantage of multi-sensory binding, finding out to which extend it could be used to increase performance and how it would affect perception of certain events. The research on the Tactylus continues work on hybrid interfaces by coupling desktop and immersive interaction (Szalavari and Gervautz '97; Watsen '99) and feedback methods, such as vibrotaction (Cheng, Kazman et al. '96; Lindeman, Sibert et al. '04). The specific integration of vibrotactile and pen-like interface devices is still rather rare. Some undocumented examples exist, but only a few have been published. Some directly related developments include the Haptic Stylus (MERL '04), which emulates the feel of buttons and the SenStylus (Fiorentino, Monno et al. '05), which integrates vibrotactile feedback in a pen-like device.

Hypotheses

- A combination of visual, auditory and tactile feedback can increase collision detection performance over purely visual feedback
- A combination of visual, auditory and tactile feedback can aid in more complex texture recognition tasks
- A suitable I/O device should be rather easy to create using existing parts, providing a better ergonomic form than current pen devices

4.4.2 Multisensory feedback binding

As discussed in section 3.3.1, the majority of approaches dealing with multimodal feedback regard output modalities as separate entities. More recently, this idea is shifting towards what is called multi-sensory binding (Spence and Squire '03) or cross-modal integration (Shimojo and Shams '01). This integration refers to activities in the brain that affect the perception of feedback obtained from multiple sensory channels. Based on the plasticity of the brain to associate sensory information, modalities interact among each other, through which perceptual events can be integrated (Pai '03). Three different situations can be identified: cross-modal bias, enrichment, and transfer. During cross-modal bias, stimuli from two or more sensory systems can differ and can affect each other leading to modified or even incorrect perception. Enrichment refers to the strengthening of perception by adding one or more sensory channels, whereas transfer deals with sensory potentials that trigger potential in another sensory channel, thereby not necessary biasing perception. One of the keys to integration is the weighting of sensor potentials, dealing with the domination of one sensory channel over another one. Generally, visual alters the other modalities, though it has been shown that sound alters the temporal aspects of vision and, based on the modality appropriateness theory, may also alter other aspects of vision (Kaczmarek, Weber et al. '91). This refers to task-specific characters of actions; The modality that is most appropriate for the task at hand will most likely dominate over any other sensory channel.

As stated earlier, the effects of multi-sensory binding can be well used to develop sensory substitution systems, such as those using vibrotaction (Kaczmarek, Weber et al. '91) (Lindeman, Sibert et al. '04), in which sensory potential and resulting perception lie closely together. An example which has greatly inspired the design of the Tactylus is the integration of visual, tactile, and sound stimuli to provide correct collision information by Shimojo and Shams (Shimojo and Shams '01), suitable for many engineering environments. Shimojo and Shams showed that in an ambiguous motion display, crossing objects could be perceived as bouncing objects, once an additional stimulus was added, in the form of a visual flash, sound, or vibration with high enough synchronicity with the visual event. Ultimately, vibrotactile stimulation provides directional cues, but this would require multiple vibrotactors, which is hardly possibly in small devices. By integrating spatial audio cues and vibrotaction, this problem can potentially be solved.

The second issue addressed by the binding of visual, vibrotactile, and auditory stimuli is the perception of surfaces. Lederman et al. report on the strength of vibrotaction for exploring textures, even though the resulting psycho-physiological function differs from using the bare finger (Lederman, Martin et al. '04). As observed by Weisenberger and Poling, audio can contribute to the perception of surfaces, like material properties and texture roughness (Weisenberger and Poling '04). In performed tests, observing both haptic and auditory cues, they found out that when perceiving textures, the weighting is predominantly haptic (62%), though auditory cues play an important role too (38%). They further noticed that auditory cues are not particularly salient upon initial experience – the cues need to be learned. This may be in line with Lederman et al’s study, which reported that sound may actually slow down texture recognition. An important factor when dealing with adding auditory cues is the level of correspondence – binding of different stimuli can lead to a different perception of textures as compared to when only visual information is provided. Hence, it seems that haptic and audio cues can either bias or disambiguate visual information. In any case, auditory cues need to make use of suitable sound models (Guest, Catmur et al. '02) to support apt interpretation.

4.4.3 Interface design and development

In order to create the new device, an analysis of the control-body linkage and control tasks was performed. This eventually led to a new kind of stylus that, in comparison to the traditional pen-input devices, has a different shape, size and balance, and an extended control and feedback spectrum.

The first step performed was a close analysis of different types of grips. Observing how users normally grasp a pen-like device, it was found that three grips are valid: two power grips and one precision grip. An important interplay between the thumb and the pointing finger can be noticed (Balakrishnan and MacKenzie '97). The traditional pen-input devices are mainly built to support a precision grip, even though multiple actions during common interaction sessions using a stylus are rather coarse. Holding the Stylus for longer durations in a precision grip, it becomes clear that the device is wrongly balanced, partly due to the pull of the multiple cables coming out of the device at the back. The balance is weakened due to the minimal surface one can place the fingers on, leading to a limited grip. This incorrect balance quickly leads to fatigue in the users arm, as they are unsupported, floating in mid-air. In order to counterbalance, the usage of the device leads to dynamic coupling between hand and device, i.e. the re-grasping of the device either in a similar grip or to a different kind of grip, to relieve the muscles in the hand (foremost the metacarpal muscles strengthening the pointing finger). With increased fatigue, the hands and finger muscles will cause slight vibrations in the hand that decrease the input performance in terms of precision. Of course this is an unwanted effect; Several studies reflect that input using a pen-like device with a precision grip, makes use of smaller muscle groups that work well for performing fine-grain tasks (Zhai, Milgram et al. '96). Loosing this effect makes the device less effective for fine manipulation scenarios with longer duration.

The second step focused on the variety of tasks performed with pen-like devices, which turned out to focus on a wide spectrum. Limited by the movements afforded by the hand (mostly gliding and angular movements like abduction or flexion) both speed and accuracy varies over the full axes. For example, the modeling task of creation a curved surface will be high speed / low accuracy (sweeping task), whereas the placement of a small object in a virtual engine will be low speed / high accuracy. Actions can be both discreet (step by step) and continuous. Some of these actions can be constrained to a limited number of dimensions, like using certain kinds of menus, others make use of all degrees of freedom.

Functionally, the traditional pen-input devices perform both within-reach and out-of-reach actions. Due to limitations in mapping (foremost small angular) hand motions to far object rotations or translations, out-of-reach actions are more coarse than within-reach actions (Mine '95). Techniques like the Go-Go interaction technique (Poupyrev, Billinghurst et al. '96a) have aimed at increasing the precision of such actions. Still, fine grain actions are better performed within-reach. A typical scenario of performing fine grain actions can be found in engineering applications at a responsive workbench (Kruger '94). Even with tricks like snapping, fine grain actions with a 1:1 control-display ratio may be hard to perform. Examples of generally occurring situations include visual occlusion problems or limited visual detail of the projection screen. For instance, think about a montage scenario in which a small object needs to be fit inside an engine. Hereby, (non-visual) aids like collision feedback can greatly enhance task performance or even enable completion of the task at all.

The control-body linkage and control task parameters led to specific design variables that predominantly affected shape and size. The grip of the new device should work for both power and precision grip and have an ambidextrous form: the device should be usable in both left and right hand. It was found that the thumb should have a good grip to counterbalance the weight of the device and, thereby, relieve the musculature of the pointing finger, when the device would be used in precision grip modus. Additionally, the main body of the stylus should be large enough to potentially hold the needed electronics of the device: buttons, 3D position and orientation sensor (magnetic or optic) and, to produce additional feedback, a vibrotactile element.

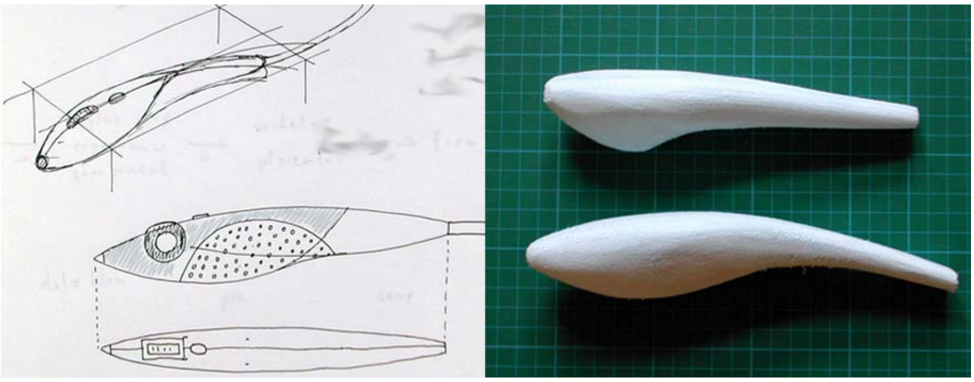


Figure 4.14: *First sketches and models of the Tactylus.*
Courtesy of M. Kunstman

In order to analyze the ergonomics of different forms of “styli”, several preliminary models were made of clay and foam (Figure 4.14). These models were tried out by a small user group (9 people) with different hand sizes in order to come to a potentially "one-fits-all" solution. As a result, a slightly curved form (over the length axis) was preferred, fitting well in the hand in both precision and power grips. An electronic components study was performed - a sum up of needed hardware components with related cables and the resulting cable thickness and control elements was made. This resulted in the decision to externalize the control boards (producing the output signals) in order to reduce weight inside the device. Also, most of the hardware needed to be off-the-shelve in order to reduce development costs and ease device communication issues (connectivity).

Increasing the number of buttons, it was decided to include a scroll wheel in order to support scrolling quickly through menu items. As a result, the performance of menu techniques like a ring menu (Liang '94) could be greatly increased. The wheel should include mouse click ability and, eventually, a second normal button would be needed for normal tasks like selecting an object. The first design included just a single button, oriented towards the centre of the device.

The design was modeled in a CAD program (Solidworks). The models were exported as stereo lithographic (STL) files and built using selective laser sintering. The resulting polyamide model served as first evaluation model for further design (Figure 4.15).

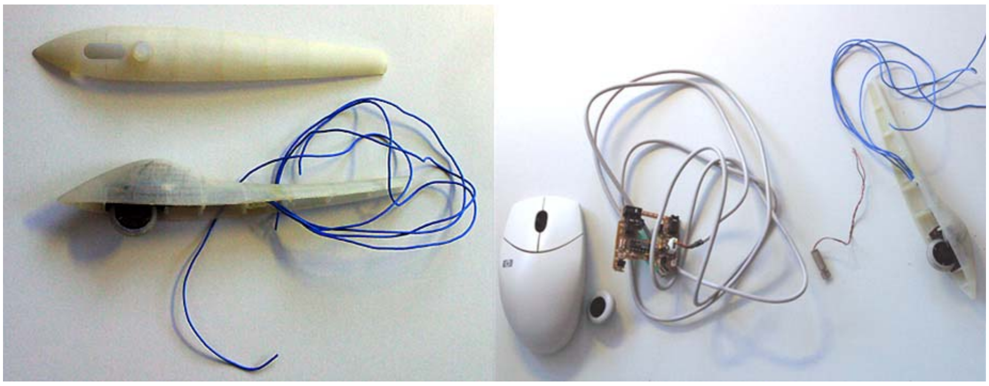


Figure 4.15: *First prototype of the Tactylus, with electronics.*

Carefully examining the grips of the user with the first prototype, it was observed that with different grips it would be ergonomic to have a button rather close to the tip of device (in precision grip) and one closer to the centre of the device (in power grip). As a follow-up, a design with two buttons was made. A button at the front belly of the device was considered, since with a power grip, the device is held similarly to a handle. Nevertheless, this option was abandoned due to construction problems. The button would hardly fit into the device or the device would be made larger than wanted, resulting in problems with fitting the device to people with smaller hands.

The second issue was the counterbalance support for the device by providing an anti-slip surface on the belly of the device, such that friction between thumb and input device could be generated to hold the device in precision grip. Anti-slip surfaces like a thin rubber coating were considered, but abandoned due to production difficulties and possible hygienic issues. The first solution was to include small rims at the side of the device. As such, the surface would be slightly rough, thereby providing a potentially good grip to the thumb or other fingers. Within a second prototype, these rims in the polyamide model proved useful, but not ideal, since the side was still too slippery.

Finally, to accommodate vibrotactile feedback, a simple tactor (vibration element) was chosen and placed in the belly of the device to achieve a good coupling with the housing, and thus with the hand of the user.

.

Figure 4.16: *Final prototype of the Tactylus.*

In the final prototype (Figure 4.16), a button was added on top to support button pressing in both precision and power grips. Additionally, the rims were made deeper on the sides to provide more grip for counterbalancing. A pen-tip, consisting of a tip of a Palm pen, was integrated in the "nose" of the device, to support controlling a PenPC or PDA. In order to track the device, a Polhemus sensor was mounted in the belly of the device. Furthermore, reflective markers for optical tracking were placed on the nose, to make the device usable in a wider number of device setups. Due to the lightweight construction, the markers did not disturb the balance of the device at all.

Of particular interest is the way tactile feedback is provided. The vibrotactor is connected to an amplifier, thereby rotating at the speed of sound frequencies put on a specific audio channel. Hence, it can vibrate at the same frequency as the audio heard through the speakers; The user is literally feeling the audio files. We took this approach, since the recorded audio files show particular roughness details when observing the waveform patterns. Audio recorded of collision (friction) between an object and rough textures shows patterns with quickly changing and highly differing frequency peaks, whereas collisions with smooth objects show a rather flat waveform. We believed that this is especially good for providing texture information, via sensory substitution. An additional advantage of using audio-based control of the vibrotactor is the potential high-synchronicity between audio and vibrotaction. Even though the adaptivity speed of a vibrotactor is slightly lower than of an audio speaker, the delay is negligible.

4.4.4 Experiment setup

Two experiments were prepared that focused on collision detection and texture recognition tasks, specifically dealing with analyzing the effects of multi-sensory binding, both in direction of either biased or disambiguated perception. 11 subjects (8 male and 3 female, novice (5) and more experienced users, aged 24 – 34) participated in the first experiment that focused on collision detection. The quality of multi-sensory feedback was tested on object placement ("key-lock") tasks with different levels of complexity. Every user performed 4 placement tasks, half in a clearly visual manner, half where the placement area was visually occluded. The users performed the same placement task under different feedback conditions, namely purely visual, visual and audio feedback, or visual and audiotactile feedback, thus resulting in a total of 12 placement tasks per subject. Cues included both collision and friction audiotactile cues. Performing the more complex key-lock placement under the visually occluded setting could be rated as moderately complex. Also, especially in the visually occluded placement task, visual information was highly ambiguous, since it was rather hard to see if an object would collide or not, due to absence of visual cues such as highlighting or shadows. We hypothesized that users would rate the audiotactile as most appropriate, especially in more complex task situations.

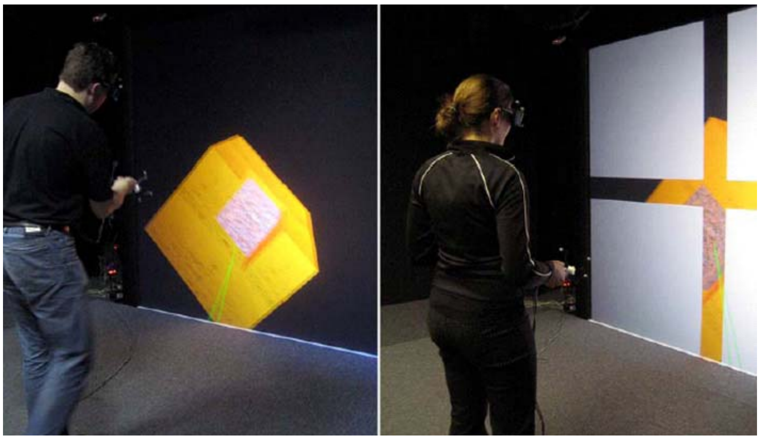


Figure 4.17: *Clearly visible and occluded placement.*

In the second experiment, dealing with texture recognition, 10 subjects (7 male, 3 female novice (4) and more experienced users, aged 24 – 49) took part, of which 8 users also took part in the first experiment. The texture recognition experiment focused primarily on biasing or disambiguating effects of visual information for recognizing different levels of roughness of textures. Subjects were presented with 15 different combinations of visual, tactile, and auditory cues, to test binding effects between different modalities. The combinations represented 5 different stages of roughness. For this purpose, 5 different textures (size 512 x 512 pixels) were selected with different visual roughness. Sounds (with a continuous roughness scaling) were generated by recording collision between real materials (metal) and synthesizing them. The vibrotactor would vibrate according to the wave pattern of the selected audio file. We hypothesized that by differing combinations of visuals, audio and vibrotactile feedback, audio and vibrotaction would alter the visual perception. Before the experiment, users were shown 2 reference textures in order for the user to correctly understand the continuous roughness scaling. Users were not informed about the methods behind the test (i.e., the differing of feedback combinations). After the experiments, users were interviewed and asked to fill out a questionnaire with 9 questions, using a 7 point Likert scale rating. The experiment was performed at a back projection display, the TwoView display, driven by a dual-pipe PC with 2,4Ghz, and a NVIDIA Quadro 4400 graphics board. Audio was provided by a pair of loudspeakers, and an ART tracking system was used to optically track the Tactylus.

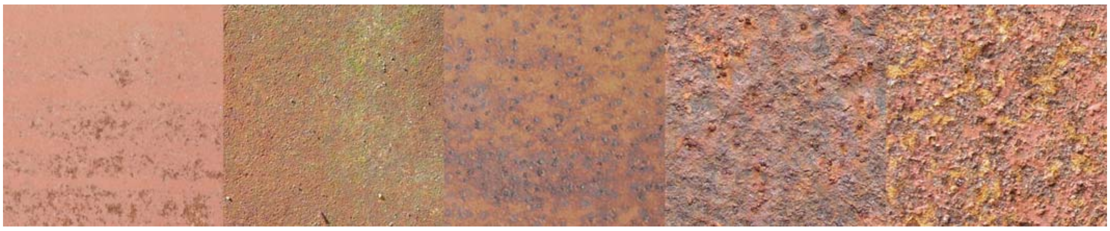


Figure 4.18: *Different levels of visual roughness.*

4.4.5 Evaluation and results

After the experiment, the results of the questionnaires and the observations were evaluated and summarized, as can be found in Figure 4.19 and 4.20. From the analysis of the results, the following statements can be made. Due to the rather non homogenous user group, these statements need to be further tested, but provide very interesting indications.

	Avg	***Stdev***	***Min***	***Max***	***Note 5-7***
1. Quality feedback in clear visual placement					
a) Audio only	*5*	*1.41*	*3*	*7*	*64%*
b) Audio and vibrotactile	*5.73*	*1.27*	*3*	*7*	*91%*
2. Ability to perform actions more precise	*5.45*	*1.29*	*4*	*7*	*64%*
3. Preference additional feedback in occluded task	*5.63*	*1.36*	*3*	*7*	*82%*
4. Accuracy of directional cues from audio	*3.45*	*2.02*	*1*	*7*	*36%*
5. Ability to map audio and vibrotactile cues	*5.18*	*1.25*	*3*	*7*	*64%*
6. Quality feedback in occluded placement					
a) Audio only	*4.64*	*1.29*	*3*	*7*	*45%*
b) Audio and vibrotactile	*5.45*	*1.13*	*4*	*7*	*73%*

Figure 4.19: *Test results collision detection.*

Audiotactile cues can enhance collision detection

The object placement experiment evaluating collision perception seemed to confirm the stated hypothesis: users preferred using audiotactile feedback to perform object placement to more correctly interpret collision between objects. In the less complex object placement tasks, users rated audio-only feedback as being good (average (avg.) 5.00, standard deviation (stdev) 1.41), but most users clearly preferred the combination of audio and tactile feedback (avg. 5.73, stdev 1.27). For the more complex tasks (occluded placement), users rated the feedback slightly lower, both in the case of audio (avg. 4.64, stdev 1.29) and audiotactile feedback (avg. 5.45, stdev 1.13). The lower rating can be explained by the higher complexity of the task. Even though most people (78%) found the audiotactile feedback more appropriate in the more complex than in the easier situations, the total amount of feedback is less, since visual feedback is highly reduced through occlusion in the complex scenarios. The majority of users stated they could perform their actions more precisely with audiotactile feedback (avg. 5.46, stdev 1.23), in which 63% claimed the precision increase to be very good (rating between 6 and 7).

Audiotactile cues can disambiguate visual information

Most users seemed to be able to disambiguate the visual information well in the visually occluded task by using the audiotactile feedback. The accuracy of spatial audio was rated less well and highly diverse: about half of the people felt they could make use of the directional information; The others had problems with it (avg. 3.45, stdev 2.02). None of the users reported problems with binding the audio and tactile information (avg. 5.18, stdev 1.25). Hence, it seems that audiotactile feedback can effectively be used to convey collision information to users, and can substantially aid in visually complex situations. It should be stated that this could be increased by adding visual cues, such as highlighting or shadows. Nonetheless, especially in visually occluded situations, as tested in this experiment, such cues become increasingly less effective.

Texture nr.	***Visual***	***Audio***	***Tactile***	***Avg***	***Stdev***
1	*3*	*3*	*3*	*2.40*	*0.70*
2	*4*	*4*	*4*	*4.10*	*0.74*
3	*1*	*1*	*1*	*1.50*	*0.71*
4	*2*	*2*	*2*	*2.60*	*1.07*
5	*5*	*5*	*5*	*5.00*	*0.00*
6	*2*	*3*	*2*	*2.40*	*0.97*
7	*4*	*4*	*5*	*4.20*	*0.42*
8	*2*	*3*	*4*	*2.30*	*0.67*
9	*5*	*5*	*3*	*4.60*	*0.70*
10	*4*	*4*	*1*	*3.80*	*0.92*
11	*5*	*3*	*5*	*3.90*	*0.57*
12	*2*	*4*	*4*	*2.80*	*0.92*
13	*4*	*5*	*5*	*4.30*	*0.67*
14	*2*	*4*	*4*	*2.80*	*0.92*
15	*4*	*2*	*2*	*4.20*	*1.03*

Figure 4.20 : *Simulation combinations and test results texture perception.*

Audiotactile cues can enhance texture perception

The texture roughness recognition experiment turned out to be a difficult and complex task. 67% of the users stated that they could correctly and well interpret the textures using visual, audio, and tactile information, thereby supporting the results reported in (Lederman, Martin et al. '04). Two users found it extremely hard, though. When observing the interpretation of the first five textures (represented by the “correct” combination of visual, auditory and tactile information), interpretation offsets were rather small, indicating that users were rather precise in texture roughness recognition. The average offset was extremely low for rough textures, whereas with light textures at most around half a scale in average. Strangely enough, some offset was seen with the interpretation of the smoothest texture, though shown as reference texture at the start of the test. The roughest texture (also shown as reference) was interpreted flawlessly afterwards. Some learning effects could be seen in users who also expressed they were more visually oriented; Biased on visual properties only, their recognition rate increased. As soon as the combination of visual, auditory and tactile information would get biased, some interpretation changes can be noticed. Based on the noted texture interpretations, a first important observation is that as long as the change in auditory and/or vibrotactile information is small, users seem to interpret the texture visually, or at least remember the original combination. Thus, changing one or two scales up or down with auditory or vibrotactile feedback does not necessary bias visual perception. This was especially true of biasing visually smooth textures with rough sound and vibration feedback. In a single case, similar visual and auditory texture of a scale 4 texture was biased by extremely light (scale 1) vibration: with around half the users, this lead to interpreting the texture as being smoother, but interestingly enough, the other half interpreted this texture as being rougher than before. Currently, we do not have any explanation for this.

Audio can alter vision
A clear offset was noticed when the (previously flawless interpreted) roughest texture was shown with level 5 vibration, but with level 3 audio roughness feedback. 89% now claimed that the texture was at least one level smoother (avg. 3.90), which is a rather clear proof that audio can alter visual perception. The interviews with the subjects made some things more clear that could not be clearly derived from the numerical analysis. First, most users reported they were sometimes annoyed by the feedback they got, supporting the hypothesis that people do not mainly base their perception on visual information, but tend to be biased by other sensorial information. Most users noted that they would first take audio into account (avg. 4.80 suitability / increase of interpretation over visual only, 1.75 stdev), whereas vibration only played a less important role (avg. 3.70 suitability for interpretation, with 2.01 stdev). These results were biased by a few users rating extremely low. Most users also reported on the lower usability of vibration for texture recognition: some noticed that the difference between vibrations was not good enough, some others found it good enough but just did not put much focus on it. Interestingly enough, about half of the subjects stated that audio did disambiguate the visual information: the textures were purposely represented in a flat way (put on a plain), such that users could clearly notice roughness differences, but not the height differences. Overall, our test results seem to imply that audio is more important than vibrotactile information for texture recognition, which contradicts the findings from Lederman (Lederman, Martin et al. '04).

4.4.6 Reflection

In this study, a new pen input device was presented. Though, at first sight, the device may seem rather conventional, its experimental characteristics can foremost be found in the new ways in which it provides feedback, surpassing recent research in vibrotaction. Reflecting chapter 2 and 3 issues, the study provided positive indications for both the usability and complexity of binding audiotactile cues. Audio seems to have a larger impact on perception of collisions and textures than vibration, even though this could be possibly levelled out when discrimination between levels of vibration is higher. The results support previous findings in the direction of multi-sensory processing and provide some ideas on how both its strength and weaknesses can be applied to increase interaction performance. Further experimentation is needed, especially by evaluation texture recognition using more combinations with different levels of feedback.

4.5 Cubic Mouse

This case study deals with the 6DOF Cubic Mouse, a 3D input device based on the idea of a coordinate system prop. The device is what can be labeled a task-specific device: it has been designed to support a specific breed of actions, for which it was hypothesized to be performing well. The design of the device has been rather experimental, and its final form can certainly be called unconventional.

The heart of the case study is a comparison of the Cubic Mouse with general-purpose 3D devices, in this case a set of gloves and a Polhemus Stylus. One of the particularities of the test lied in the observation of integrated versus separated control. The Cubic Mouse allows both integrated and separated control; The latter is specifically focused on constraining interaction by being able to control a single dimension, for example, the rotation along a specific axis. This separation has a large effect on interaction performance, as can be seen later on.

Finally, in order to study the different motion patterns of the devices, a new analysis tool has been created and used, based on 3D trajectory visualization.

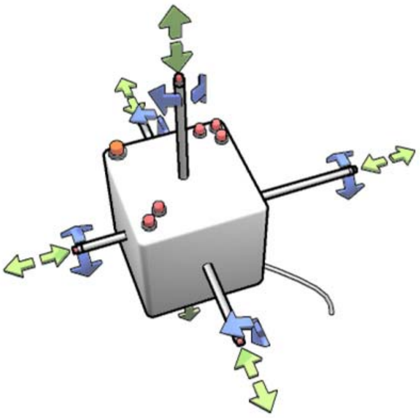


Figure 4.21: *The 6DOF Cubic Mouse.*
Courtesy of H. Seichter

Hypotheses

- Subjects would perform fine-grain manipulations easier and faster with the Cubic Mouse than with the Stylus or the PinchGloves
- The Cubic Mouse would not be preferred for coarse manipulations

4.5.1 Background

The Cubic Mouse (Figure 4.21) device consists of a cube-shaped box with three perpendicular rods passing through its center. There is a six degree of freedom (6DOF) tracker embedded in the Cubic Mouse, which is used to track the position and orientation of the Cubic Mouse. The rods typically represent the X, Y, and Z-axes of a coordinate system. They can be pushed and pulled, which allows the constrained input of three degrees of freedom. In addition, the 6DOF Cubic Mouse allows the rotation of the rods adding another three degrees of freedom. Altogether, there are a total of 12 "analog" degrees of freedom available with the 6DOF Cubic Mouse, which can be assigned to interaction tasks in various ways. Furthermore, there are six application programmable control buttons mounted on one face of the Cubic Mouse and a single button at both ends of each rod. Typically, users hold the device in their non-dominant hand and the dominant hand manipulates the rods and the control buttons.

The 6DOF Cubic Mouse is based on the previously presented Cubic Mouse (Froehlich '00), which allowed only pushing and pulling of the rods. For general 3D manipulations six degrees of freedom are often essential, which were not available with the original Cubic Mouse. As a result, the original model was mainly used for controlling three orthogonal cutting and slicing planes in 3D data sets.

4.5.2 Related work

Props are real-world objects, which represent a virtual object they more or less resemble. Hinckley et al. (Hinckley '94) use a head prop in their neurosurgical visualization application. In their system, also dealt with in section 2.2.3, users hold a doll's head or a small rubber sphere with an embedded tracker in one hand. This head prop is used to control the orientation of a head data set viewed on the desktop screen. The other hand holds a second prop that, for example is used to position a cutting plane relative to the head prop. This is, in contrast to our applications, where the dominant hand is used to handle controls integrated in the Cubic Mouse held in the non-dominant hand. Research within the area of props (Hinckley '94; Wloka and Greenfield '95) has shown that these devices are often very familiar to the user, obvious to use, supply strong feedback and utilize the two-handed frame of reference.

Two-handed interaction has been a popular research area since Buxton and Myers' study on two-handed input (Buxton and Myers '86) and Guiard's application of the Kinematic Chain theory (Guiard '87) (also see section 2.2.3). These initial studies have led to a series of experiments on two-dimensional and three-dimensional input, examining common factors with respect to bimanual interaction techniques. Examples are (Kabbash, Buxton et al. '94) (Hinckley, Pausch et al. '97b), as well as more specific studies on psychological and motor behavior factors by (Leganchuk, Zhai et al. '99; Balakrishnan and Kurtenbach '99b). Of specific interest to our research is the work by (Hinckley, Pausch et al. '97b), which emphasizes the importance of haptic feedback supplied by props for reducing the level of attention (cognitive load) for tool selection procedures.

Two-handed interaction techniques for virtual environments have become increasingly popular. Examples can be found in (Mapes and Moshell '95; Cutler, Froehlich et al. '97; Ullmer and Ishii '97; Mine, Brooks et al. '97a; Leibe, Starner et al. '00). The analysis of these two-handed techniques shows a clear tendency towards increased performance for most tasks in comparison to one-handed interaction techniques. However, as identified in (Balakrishnan and Kurtenbach '99b), precise guidelines for the development of two-handed techniques still requires more systematic experiments under a variety of conditions.

Several devices exist that share specific characteristics of the Cubic Mouse. First of all, dial boxes and slider devices also allow constrained input similar to the Cubic Mouse, but these devices are not tracked. The TouchCube (ITU Research, presented at the SIGGRAPH'98 exhibition) is also a cube-shaped input device, which is equipped with touch-sensitive panels mounted on five sides of the cube. The device is operated by applying gestures to the touch panels. Adding a 6 DOF sensor to this device could also result in an intuitive multi-DOF input device.

General issues with respect to 6DOF input are well explored in (Jacob '92). Zhai's influential study (Zhai '95) investigates a large number of factors with respect to the design and evaluation of 6DOF devices. Besides speed and accuracy Zhai also includes ease of learning, fatigue, coordination, and device persistence and acquisition factors.

4.5.3 Interface design and development

The Cubic Mouse has been developed for applications that deal with the manipulation of a single virtual model. Examples of this reference model are a geological model or a car. The Cubic Mouse implements the idea of virtually holding the reference model in one's hand. Navigation in this context entails rotating, translating, and zooming into the reference model. This is different from such techniques as flying or walking, which is often the preferred way of navigation in walkthrough scenarios. The 6DOF Cubic Mouse's rods are used to control virtual objects relative to the reference model. Each rod controls, simultaneously, up to two degrees of freedom.

All of the reference models being used have an up and down direction and assign the up direction to the Cubic Mouse's face with the buttons. The cabling comes off the opposite face that is, thereby, naturally assigned the down direction. The other directions are aligned with the natural coordinate system that comes with the reference model, e.g. the car model has a front, rear, left and right side.

The 6DOF tracker inside the Cubic Mouse is used to position and orient the reference model. Thereby, the reference model's orientation is always aligned with the orientation of the Cubic Mouse. For positioning tasks we use a 1:1 control-display ratio on the Responsive Workbench, 1:3 to 1:5 ratios for larger display devices like CAVEs or Reality Centers. The rods represent the X, Y, and Z axes of the reference model's coordinate system. Tracking the Cubic Mouse ensures that the rods stay aligned with the coordinate system axes. Pushing and pulling the rods is typically used for translating a virtual object, like a cutting plane, relative to the reference model in the appropriate direction. Rotating the rods is mostly used for rotation of a virtual object around the corresponding axes of the reference model. The buttons are used for clutching, zooming in and out, and various other application dependent tasks.

Adding another three degrees of freedom to the originally developed 3DOF Cubic Mouse makes the 6DOF Cubic Mouse a much more generally applicable device. The 6DOF Cubic Mouse puts the reference object's coordinate system into the user's hand and allows constrained translation and rotation with respect to the primary axis of this coordinate system. These are the most common operations for positioning and orienting objects in three-dimensional space and give users full six degree of freedom control. More specifically, (Froehlich, Plate et al. '00), differentiate between 1DOF control (separately controlling a rotation or translation along a single axis), 3DOF control (either rotating or translating an object around all axes), or 6DOF control (freely controlling rotation and translation along all axes).

4.5.4 Experiment setup

An experiment was set up to assess users' reactions to the Cubic Mouse in comparison to other commonly used input devices, such as the Polhemus Stylus and the FakeSpace

PinchGloves. The test application was displayed on a two-sided Responsive Workbench driven by a SGI Onyx2 InfiniteReality graphics system, using a Polhemus Fastrak tracking device and Stereographics CrystalEyes2 shutter glasses. The test was implemented using the AVANGO virtual reality framework (Tramberend '01). The main aim was to empirically characterize the Cubic Mouse in comparison to common input devices. That is, the experiment focused on the task-specific actions supported by the Cubic Mouse, in comparison to general 6DOF devices. This task-specificity was mostly defined by the inherent constrain methods afforded by its mechanical construction. Furthermore, it was aimed at getting a first analysis of trajectory behavior of the three devices using a new graphical trajectory analysis technique and to check user preferences for the three devices. The exact comparison of the completion times (as being a pure speed test) was not a main aim and would require an even more detailed study.

The test involved four basic docking tasks, using both the translational and rotational functionality of the 6DOF Cubic Mouse. These docking tasks represent tasks that are normally not done with the 6DOF Cubic Mouse, but due to the variety of the tasks, it is valuable in order to check device characteristics. Three of the tasks where basic „key and lock“ tasks in which a „key“ object should be placed into a „lock“ object. The fourth test involved the matching of two objects. The four tests included both coarse (movement of key to lock object) and rather fine-grain manipulations of objects (correct placement of key in lock object). The configuration of objects was unambiguous. Subjects were asked to perform the test as fast and accurately as possible. The test would end if both subject and observer agreed that the object was within a threshold range to the perfect target. No specific timestamps were set at the start of each test in the used log files during the test; Timestamps for the end of the test were automatically included when stopping the test by the observer. Start timestamps were added later in the analysis phase, to exclude the startup time of the subjects. Only an approximate indication of times would be needed for our analysis, since, as we will see in the results, time differences where extensive between different tasks.

13 subjects participated in the within-subjects test, resulting in 98 trials. The subjects’ experience with computers differed from novice to advanced level; Four users had used Virtual Environments before. Their age differed between 20 and 30 years, 11 subjects were right handed, one left-handed and one ambidextrous. Users had to fill out a questionnaire that had questions with respect to ergonomics, performance, and preference within the test scenarios. The questionnaire made use of a 7-point Likert scale. All subjects used all input devices in a randomized order.

4.5.5 Evaluation and results

The analysis of the mean times and standard deviations of the task completion times showed that the 6DOF Cubic Mouse was outperformed in the coarse movement actions of placing the key object within reach of the lock object, but was very fast and effective within the fine-grain actions when the key object had to be placed precisely in the lock object. A comparison of the overall times completion times showed differences of up to 255% time increase between completion times of the Cubic Mouse in relation to the other two devices. The actual times were affected by the required object selections for gloves and stylus, but the overall times were still considerably shorter than those for the Cubic Mouse. A comparison of the absolute distance between the key and lock object within the target area showed, without exception, the precision of the Cubic Mouse with fine grain actions. In comparison to the stylus and the gloves, the positional accuracy for the Cubic Mouse was generally much better for the alignment of the key with the lock.

Trajectory analysis

The trajectories for the docking tasks were visualized as curves in 3D space (Figure 4.22). Examining the representative trajectories provides excellent insight into the differences in performance during coarse and fine movement phases and clearly shows the effect of constrained interaction.

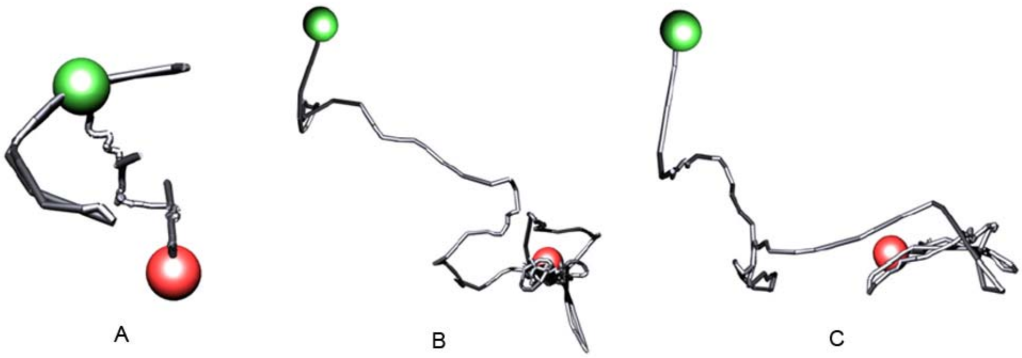


Figure 4.22: *Trajectory analyses of the Cubic Mouse (A), gloves (B) and Polhemus Stylus (C). Courtesy of H. Seichter*

Three dimensional trajectory analyses have been performed before. A good example can be found in (Jacob, Sibert et al. '94), in which the authors show a technique that visualizes the movement path and speed in a 2,5D representation. This technique is useful in many occasions, but lacks the rotational analysis that would be especially interesting for our examination of separable actions. As also described by Masliah (Masliah and Milgram '00), coordination of control is often disregarded. Our trajectory analysis for the three input devices also only compares the positional trajectory in 3D. The visual inspection of the trajectories and an analysis of our log files reveal the major differences of the devices during the initial coarse movement phase, when the key is moved near the lock and the fine movement phase, when the key is actually placed in the lock. Hence, the following seems evident:

- *Clutching*: clutching leads to obvious inefficiency when a shortest path comparison (Fitts '54) is made. Due to the clutching, several extra movements are made by users which lie away from the ideal path (as being an ideal path from point a to b in 3D Cartesian space).

- *Speed/accuracy tradeoff*: there is a noticeable speed/accuracy tradeoff between the three devices. The stylus and gloves perform fast overall, but relatively slow, inefficient, and non-precise during fine grain manipulation. This inefficiency can clearly be seen at the large nodes of movement of both devices around the target area. The Cubic Mouse performs slow during coarse movement, but is very efficient and fast during fine grain manipulation.

- *Coarse movement difference*: the integral devices (like the gloves and the stylus) show accurate coarse movement. When looked at the Cubic Mouse, a separable device, stair case patterns can be seen within the coarse grain movement part of all tests. This can be explained by simply focusing on the separation of control; With the Cubic Mouse, movement along a diagonal path can not be made, resulting in a continuous change of allocation of different controls (rods) to move approximately in a diagonal manner.

- *Fine grain manipulation differences*: the mentioned nodes around the target area, in both the paths of the stylus and the gloves, showed several clear differences in performing fine grain actions. The nodes are presumably caused by continuous compensation of certain dimensions, since movement (especially rotation) with integral devices along one axis often leads to unwanted movement along another axis. The path of the Cubic Mouse shows clearly the favorability of a constrained (separable action) technique of moving an object, since the movement is very efficient. Stylus users positioned the key object above the target area; the glove users rather positioned it within the target area. This difference remains unexplained.

Hence, the advantages of the Cubic Mouse during the fine manipulation phase seem to have their origin in being able to independently manipulate the different degrees of freedom, while exactly this turns out to be a big disadvantage for coarse movements.

Questionnaire

Subjects found the Cubic Mouse considerably more comfortable to interact with than the gloves or the stylus. Subjects reported that they had to look more at the Cubic Mouse than at a glove or stylus. This might be caused by the fact that during interaction with gloves and stylus, these devices are always in the view frustum of the user, whereas the Cubic Mouse is usually used eyes-off.

Overall, the users felt like they could control their actions about equally well with all three devices, although surprisingly, the stylus got the best ratings. The same tendency was found, when asking if the subject could control actions in a straightforward manner.

In the device comparison, the highest score was given for fine grain positioning and orienting. Here, 83% of the subjects preferred the Cubic Mouse over gloves and stylus for the “key” object. Notable is also the 83% preference for coarse orienting of the “lock” object with gloves, which was even higher than the coarse orienting of “key” objects choice of gloves (53%). In addition, the stylus scored well in the ranks for fine orienting “key” objects (50% voted for stylus usage, 50% for the Cubic Mouse). For coarse orienting and positioning of both scenes and objects, the Cubic Mouse was either not chosen not at all or by a very small amount of people (14% in case of coarse orienting of objects).

4.5.6 Reflection

Hinckley et al. (Hinckley '94) used a head prop and a separate slicing plane prop to position a slice through a computed tomography data set of a human head. This scenario is very similar to scenarios, in which the Cubic Mouse moves a single slicing plane.

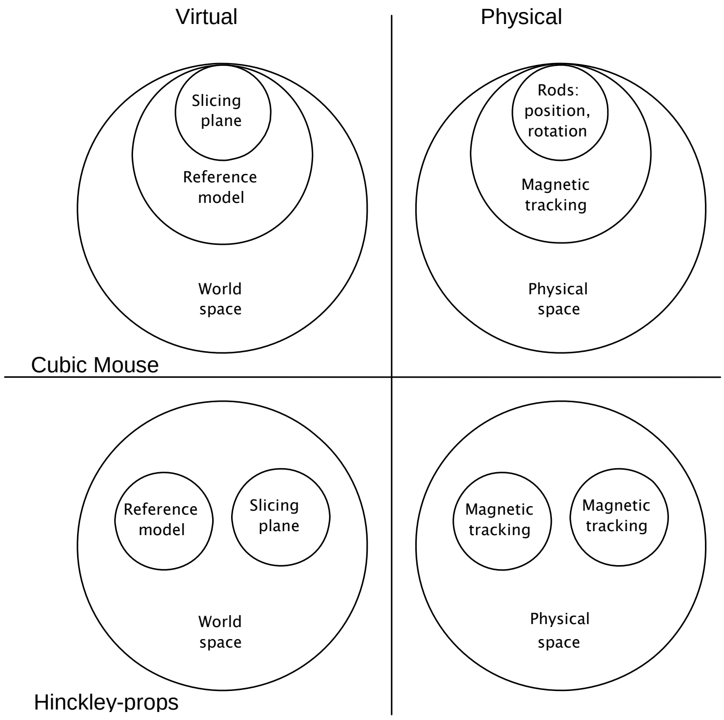


Figure 4.23: *Hinckley's props versus Cubic Mouse. Courtesy of B. Froehlich*

Figure 4.23 shows a schematic that compares both interaction methods. With the Cubic Mouse, users have persistence of the slicing plane, which is not possible in Hinckley's approach. Persistence allows the user to pass the Cubic Mouse on to another user without changing the position of the slicing plane. The tradeoff is that Hinckley's approach allows a user to move the cutting plane relative to the head model and also the head model relative to the slicing plane. With the Cubic Mouse, a user can only position the slicing plane relative to the reference model.

Another issue which is of particular interest is how control is applied with the 6DOF Cubic Mouse. In principle, the device allows both integrated and separated control of degrees of freedom. Classically, separation of control has been performed by using software-based constraints, integrated in interaction techniques. In the case of the 6DOF Cubic Mouse, constrained interaction is directly coupled to the hardware of the device and not only to the interaction technique. By allocating the integral and separated control actions over the different hands, a user can make optimal use of the motor behavior of both hands. The dominant hand can control the separated controls, thereby, allowing fine-grain manipulations on objects, whereas the non-dominant hand is in command of the integrated control, used for viewpoint manipulation. In viewpoint manipulation, fine grain actions do not normally need to be performed. Nevertheless, simultaneous control of all degrees of freedom (integrated control) when manipulating an object is not directly possible with the Cubic Mouse interaction model, even though this would sometimes be appropriate, as can be seen in the user test.

So, what can be concluded from this study? Obviously, the unconventional, task-specific design of the Cubic Mouse in partly separable degrees makes sense, at least, for a specific task domain in which fine-grain actions need to be performed. Furthermore, especially when a direct affordance between the (rather generic) form of the Cubic Mouse and the reference model can be established, the viewpoint control is rather easy. Nonetheless, as soon as coarse actions need to be performed with the device, there is a massive speed-accuracy trade-off, which does not always make sense. Switching devices would be a possible solution, but disturbs the interaction considerably. Hence, some interaction techniques that make use of the integral characteristic of the device should be developed, so that some of the disadvantages can be taken care of. However, it can be expected that for coarse actions, the Cubic Mouse will never be the ultimate choice.

The trajectory analysis method is highly interesting, even without rotational information. Adding this information, via a suitable visualization method, would clearly advance the interpretation. One possible method is to move a pointer (or lens (Viega, Conway et al. '96)) along the path to get detailed directional information at specific points. Adding this information along the full path would most likely result in an incredibly complex visualization.

4.6 ProViT

The following studies exemplify one of the main approaches to integrate unconventional interface techniques in traditional applications. This approach is generally referred to as hybrid interfaces, which combines multiple display and interaction techniques, in order to match the work processes at hand. Even though most hybrid interfaces combine 3D perception and interaction with 2D techniques, the basic issues being tackled are highly similar to combining unconventional and traditional techniques. As such, this section deals with the issue of porting unconventional interfaces to general work environments, handling some, but not all of its issues. For example, dealing with sensory substitution systems often implies other application factors, as dealt with in section 3.5.

This section handles the interaction concept of ProViT, dealing with general design issues for integrating multiple interaction techniques. These design issues were further applied in the second case study, Capsa Arcana, a console device integrating multiple unconventional sensors. Finally, the Eye of Ra deals with integrating 2D and 3D functionality in a quite revolutionary shaped new input device.

4.6.1 Background and related work

The ProViT project (BMBF, 2001-2004) focused on the usage of distributed virtual environments for design review in the ship, machinery, and airspace industries (Cao, Gaertner et al. '03). Using a network infrastructure, users could remotely cooperate on complex problems using an immersive virtual environment. For this purpose, a distributed system architecture was set up using a PC-cluster, an immersive L-shape display (such as a responsive workbench), AV-streaming equipment, and several interaction devices. Enabled by this infrastructure, users could run simulations, manipulate objects, access data from other applications and communicate with each other via videoconference.

Three different input devices were made available: a 12-inch tablet PC, a tracked Polhemus Stylus, and a Cubic Mouse (see section 4.5, (Froehlich '00)). Several studies were performed, which investigated the role of complex but effective system control mechanisms on the flow of action (section 3.4.3) in a hybrid application.

The first question to be asked is: why make use of hybrid interaction? The answer lies in the complexity of many applications that are currently being designed: hybrid interaction can lead to considerable performance increase. The combination of 2D and 3D interactions can have considerable advantages. 2D actions can be performed with relatively high precision, whereas 3D actions are executed at high speeds in specific task situations. As such, a clear speed-accuracy trade-off can be noticed, depending on the task at hand. The usage of 2D devices, specifically handhelds such as PDA’s or transparent props, for 3D interaction has been probed several times before, including (Watsen '99) (Hachet '03) (Schmalstieg '99).

When using hybrid interaction, some general problems can be identified. These problems may seem obvious, but are still regularly occurring pitfalls in many applications. The wrong mapping of techniques on devices is one of the most frequently occurring failures. Developers have to pick from a limited number of interaction techniques that do not always map the input device in use. Also, input and output devices can be wrongly combined. Overall, wrong mapping leads to performance loss, which directly influences flow of action in an application. It is a myth that different input devices are capable of transmitting comparable content (Oviatt and Cohen '00). Another regularly occurring problem is system control overload. To cover the increasing amount of functionality, some developers are simply placing more flying menus in the VE. Often these menus also have to be large, for readability issues. The allocation of menus generally overlaps the main work area, leading to attention problems. Rapid eye and head movements of the user are observable, moving from front (menu) to back (work area) focal planes.

Finally, with the increasing complexity of immersive applications, feedback to the user is of utmost importance. The user needs to know what is the current mode of interaction and if a performed action has been successful. Often, users need to check multiple times to determine if a task has been performed, leading to unnecessary disruptive action loops.

Hypotheses

- Careful combination of different task-specific devices can improve the flow of action in complex applications
- Hybrid interaction can solve system control issues in immersive environments when large menus need to be used

4.6.2 Interface design and development

In order to match the need for a good flow of action, the following concepts were implemented. As a basis for this application, an in-depth task analysis has been performed that specifically looked at task syntax factors affecting flow of action. With respect to the mapping of the functions to the devices, it was analyzed which three-dimensional actions can be performed by the general-task device (the Stylus) and which should be performed by the task-specific device (the Cubic Mouse). A general setup can be seen in Figure 4.24.

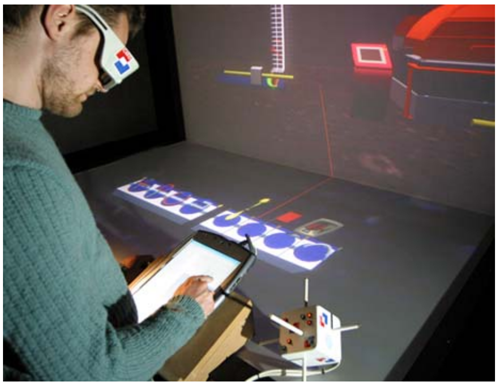


Figure 4.24: *ProViT system setup.*

General task devices often need to make use of some kind of system control technique to change their mode of interaction, whereas task-specific devices regularly integrate the interaction mode into the design of the device. Intrinsically, task-specific devices perform a small amount of actions very well. However, the performance structure of actions performed with the task-specific device regularly includes small repeating interaction loops. These loops are actions that may not be mapped to the device, like selection of a new object, since performance normally decreases. Nevertheless, since the switching of devices disturbs flow action considerably, so called multi-mapping techniques have been applied. Simply said, some actions can be performed by all the devices to avoid device switching. Here the tradeoff between device switching and worse performance is certainly advantage of the latter, since the repeating interaction loops occur rather often.

The two-dimensional tasks, like the session management, are placed on the Tablet PC. In order to avoid some of the device switching, we added a pen-tip to the stylus, thereby being able to directly control the Tablet PC with the stylus. The allocation of the graphical user interface elements has particular effects on the attention of the user. The interaction mode changes of the stylus can be achieved by a hand-oriented menu, whereas all two-dimensional actions (including symbolic input) are placed on the Tablet PC. The Tablet PC is bidirectional synchronized with the VE, showing the current state of action visualized in the application's scenegraph. The Tablet PC is connected directly in front of the user on the Responsive Workbench.

The allocation of the interfaces implies that the user has several focal areas. Foremost, these focal areas are the active work area (the 3D model) and the display of the Tablet PC (Figure 4.25). The advantages of using the Tablet PC are its high readability and low overlap with active work area displayed on the Responsive Workbench. Having two focal areas may have a negative influence on flow of action, since the focus of attention may be changing regularly. In our application, the functions were mapped in such a way that, most of the time, multiple tasks will be performed serially with one device. In this way, there is no direct need to switch between a 3D input device and the

Tablet PC in a parallel way. Some exceptions exist, though. When the user is performing symbolic input, the stylus and the Tablet PC are used in combination.

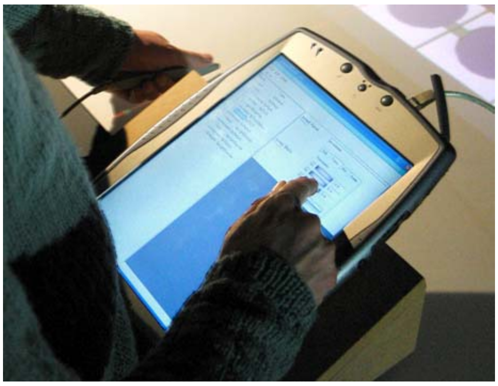


Figure 4.25: *Close-up of the tablet PC interface.*

Also, the Tablet PC can be used for feedback purposes. In order to avoid confusion, having the Tablet PC in the same place with respect to the user has a big advantage. Due to the fixed (mounted) position, the focus of attention of the user is always directed to one spot (directed focus of attention), thereby minimizing unnecessary search behavior. This behavior is often observed when floating menus are applied, since the user has to check through several focal areas to find the desired widget item. Since the user is switching between devices, continuous feedback needs be taken special care of. This feedback can be attached directly to the input device, for example via a small icon at the stylus tip or via a feedback method that is independent of the input device in use. Therefore, we have implemented the concept of cross-device feedback. First of all, we always display the current interaction mode by a small text-icon that is placed at the same focal depth as the active work area, in order to avoid switching between focal planes. Nevertheless, for more complex actions, to communicate the active interaction mode is not enough. Therefore, we have applied a scenegraph-oriented interaction mode on the Tablet PC (Mueller, Conrad et al. '03). At the Tablet PC, the currently active node in the scenegraph is displayed, showing detailed data on this node, and the current action being performed. This implies that users can always fall back to the scenegraph-oriented interface, either to receive detailed feedback or to perform a preferred action directly via discreet input. Even though looking at the Tablet PC to receive feedback implies a change of focus of attention, it resolves user confusion immediately, since the Tablet PC gives a complete state overview (based on node selection) of the last performed action. The possible attention loss is an acceptable tradeoff.

4.6.3 Evaluation and results

The ProViT system was demonstrated to a large number of end-users from different areas, including the car industry (Volkswagen AG, Ford, DaimlerChrysler AG, Audi), suppliers (Faurecia, Brose, Federal-Mogul Sealing Systems Bretten GmbH & Co. KG), and the airspace and ship building industry (Aker Werft, Airbus, Pace). In several demonstrations, users were asked to fill out a basic questionnaire with some basic questions. Most of the users had experience with CAD and simulation systems.

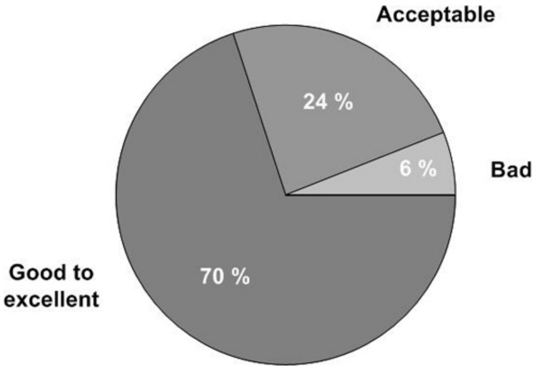


Figure 4.26: *Overall usability satisfaction.*

The overall satisfaction (Figure 4.26) of the application’s usability was found good to excellent by 70% of the users, 24% found the usability still acceptable. Only 6% of the users did not like the software at all.
Through both direct discussions and the questionnaires, it can be stated that most users were satisfied with the usability of the application. 80% found the navigation and manipulation tools to be effective for the tasks at hand. Nonetheless, 44% claimed that the feedback mechanisms worked insufficiently; A good amount of users did not find they were informed well enough on their actions. About 30% claimed that the menu system did not work well enough.

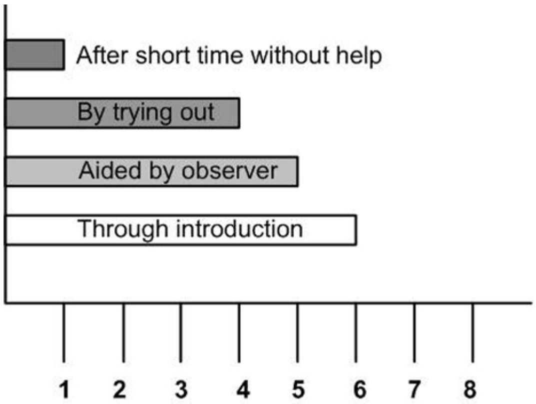


Figure 4.27: *Learning to use the application.*

Finally, most users learned to make use of the application pretty quickly (Figure 4.27). Only 25% had to ask about the interaction mechanisms during usage, all other users were able to operate the application through the knowledge acquired by introductory words or by more or less trying (learning) out themselves.

4.6.4 Reflection

Overall, the test proved though that the hybrid interaction approach is valid, but that improvements are needed. One of the critical issues was, that even with improved feedback mechanisms, users still complained that they were not informed well enough on their actions or system state. One of the reasons can be found in the close coupling between hand-oriented and the Tablet PC menus. Most users, who were first time users, have predominantly tried out the flying menu, not the Tablet PC, quite simply since the latter takes time to understand and use due to its complexity. Users obviously did not completely grasp the interdependency between both menus. It can be expected that a large increase in feedback quality can be achieved when users make use of the system for longer periods of time. The demo sessions did not leave enough space for the end-users to get a thorough picture of the system, a considerably longer and more formal evaluation is needed for this. This kind of test should include the complete cycle of usage, from starting the application up to using most of the functionality, using both 2D and 3D interaction techniques in a more equivalent way. Due to the nature of the demonstration, in the current test users predominantly made use of the 3D techniques.

One final point which seems to be contradictory to the stated feedback problems is the rather well rated learnability of the application. How can it be that users could use the application without much help, but still claimed to be not well informed by the system? An answer to this contradiction was not found.

Reflecting chapter 2 and 3 issues, some concerns need to be stated. After the evaluation, it was clear that during hybrid interaction both 2D and 3D interaction should be more closely coupled. One lesson learned was to observe the different kinds of integration between 2D and 3D interaction. It is important to differentiate between two approaches: serial and parallel integration. Using serial integration, 2D and 3D methods are used in a sequential order, one after each other. In parallel integration, 2D methods are quasi embedded and used directly to control and adapt the data in the immersive environment. An example of parallel integration is to solely make use of a tracked Tablet PC to interact with the VE. Hereby, the 2D (touch screen) and 3D (6DOF tracking) are directly coupled and can be used almost simultaneously. This may have clear advantages for the close integration of hybrid technology and techniques, but has limits too. For example, to carry the device continuously is rather unergonomic and tedious and screens are small, as large screens weight too much. One way around this is the previously presented Control Action Table (section 2.2.3) which integrates the touch screen in a ground-referenced 6DOF construction. Of course, mobility is an issue here, but techniques are clearly closely coupled: subtasks can be better “chunked” together (Buxton '86), leading to better continuity in interaction, since one does not need to switch between devices anymore. Yet another way to more directly integrate hybrid techniques is to limit menu interaction to one method. An example of such a direction is presented in section 4.6.3. By doing so, interaction mixing between 2D and 3D techniques is enforced, since the user has to make use of the Tablet PC to operate the application.

One question remains: does ProViT show any truly unconventional interaction methods? The answer is no, it does not, even when considering the integration of the Cubic Mouse. Its strength, though, should be found in the observations of integrating more unconventional methods into traditional environments. For this purpose, the study provided valuable insights that were used and improved in further studies, as following in the next sections.

4.7 Capsa Arcana

This case study focuses on the design and development of a console display and control device for interaction in virtual environments called Capsa Arcana (freely translated “mysterious box”). The device was initially built as the main I/O device structure for interaction in interactive guided museum tours, within the project DHX (EU-IST 2001-33476). The main aims of this project were to allow non-linear guided tours through reconstructed heritage, possibly in distributed scenarios over network. The interaction mechanism used in the DHX installations was dialogue-based. This mechanisms supported communication with a virtual guide in order to control the tour, and to retrieve information about artifacts. In order to support this interaction, a dialog GUI was developed that needed to be controlled via a direct and robust user interface. Along with the dialog interface, a way of navigating the scenes was required (Kruijff, Conrad et al. '04).

To introduce new interaction mechanisms in museums, more advanced and unconventional interface techniques were conceptualized that could be combined with the rather conventional dialog interface in a modular way. The usage of such interaction techniques could lead to more excitement within the guided museum tour scenarios and therefore, could attract more visitors to museums using such an installation (note that attracting visitors is an important issue for museums to apply VR technology).

The case study produces several useful insights in the design and development of unconventional interfaces and ways to integrate these in conventional environments. Of particular interest is the focus on public space systems, posing strict design boundaries that should be kept in mind. These considerations are presented in the next section, after which the conceptualized and created framework of unconventional controls is handled. An evaluation of the console illuminates some of the presented aspects.

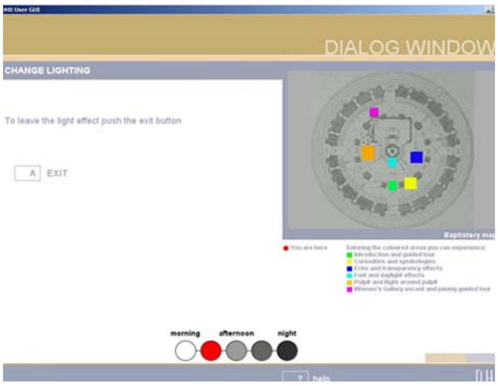
Hypotheses

- Combining conventional and unconventional interaction methods can improve attractivity of a museum demonstration, surpassing simple large-screen walkthrough applications or desktop multimodal applications
- Hybrid interaction is a suitable way to include spectators quickly in an immersive public experience through usage of known interface methods

4.7.1 Interface design and development

The design of public space systems is bound by numerous factors that restrict the choice of interaction devices and techniques. During the design of the Capsa Arcana, several of these issues needed to be regarded. These and further issues are handled in more detail in (Bowman, Kruijff et al. '05).

The first ground rule is robustness: the devices need to be extremely sturdy to withstand rough usage. The construction should be safe, ensuring that a user does not get hurt when bumping into it, or scratches herself on a sharp corner. Kids in particular will try to break devices or take them – theft is unfortunately also a major problem. Hence, input devices should be hard to detach from the construction. Hiding (embedding) cables as far as possible is one way to avoid detachment. A further aspect is the familiarity and ease of interaction. Users should have a short learning curve; Otherwise, the device construction is useless for most of the cases. There are exploratory devices, such as can be useful in museums to learn about mechanisms, but for most applications, users should be working with the devices from the start. One way to go is to integrate traditional technology with unconventional techniques, another is to mimic conventional interaction methods, changing them slightly into a new direction. Finally, the anthropometric differences and capabilities of users of different ages should be regarded. Whereas kids can be accustomed to making use of new devices, conditioned by using game consoles, the elderly might have great problems operating these. Furthermore, all devices should be accessible by the user.



To leave the light effect push the exit button

Figure 4.28: *the DHX dialog window.*

The Capsa Arcana has been built around a familiar concept: touch screens are widely used in museums. This device was the starting point for the design. Seen as a whole, the console infrastructure holds some basic methods of controlling GUI-based applications, communication devices to communicate over the network, and a joystick interface for moving through virtual environments. Hence, as our first approach for controlling the general functionality in a virtual environment, 2D metaphors and techniques were directly mapped on 3D functions.

A robust touch screen was selected, which has a 17” screen and runs at a 1024 x 768 resolution. It has an extremely strong touch area, consisting of 6mm thick secure glass to avoid damage to the screen. In order to guarantee usability, buttons needed to be rather large. Input recognition becomes worse through dirt on the glass plate, and people with lesser visual capabilities require a suitable size. On the touch screen, users could control the DHX dialog window (Figure 4.28), which was used for the dialog oriented interaction with the Virtual Guide, resembling communicative interaction with a museum guide in real life (Kruijff, Conrad et al. '04). Based on a storytelling engine, a hierarchical finite state machine, the user is guided through the museum tour, being offered multiple actions at specific locations. The user can freely modify the tour at will, and is able to access multiple levels of information. Thereby, dialog-like menus were used also for non-dialog control, as can be seen in the Figure above, showing a way of changing light effects in the Baptistery in Pisa.

In order to navigate through virtual worlds, a simple joystick interface was chosen that mimics 2D and 3D movement as used in computer games. The first choice within the development phase, was to use a Logitech Extreme3D joystick, which had the ability to rotate the stick around its axis, so that people could look around while still staying at the same location within the virtual world. The joystick also holds some mini joysticks that were used to control head rotation, ideal for looking up in larger spaces like architectural reconstructions. However, the joystick is not very robust and therefore inappropriate for public installations since it will break within days. Discussions with builders of public VR installations lead to choosing a full metal joystick mounted on a stable console. Industrial joysticks are extremely robust, but most available ones have a limited number of buttons. Therefore, the choice was made to use the Thrustmaster Houtas Cougar, a replica of an F16 flight stick, with multiple buttons and knobs, and completely made from metal (see Figure 4.30). The joystick is robust enough to withstand long-term public usage and supports both navigation in 2D and 3D and the additional functionality of head rotation. The stick of this joystick can not rotate around the axis, but this was found to less important as the needed robustness. Furthermore, the axis rotation could be mapped by one of the several extra directional buttons available on the joystick.

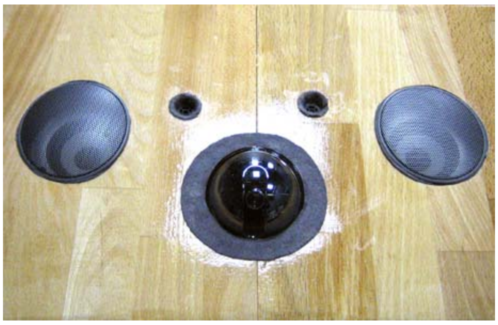


Figure 4.29: *Communication devices.*

Primarily as an outcome of the DHX remote guidance requirements, communication devices were needed, including a microphone, camera and loudspeakers (Figure 4.29). As video camera, the Logitech Quickcam (USB) has been installed. It delivers 640 x 480 videos and holds a pan-tilt unit, which can be controlled automatically using the Logitech computer vision based software that focuses the camera on the user’s face.

In certain DHX scenarios, larger groups of people should be able to communicate simultaneously. Hence, an omni directional microphone was needed to grab speech signals from a wider area surrounding the console. The microphone chosen, an Andrea Sound Superbeam, integrates two separate microphones (two-channel microphone array) and can grab sounds within a range of several meters. The microphone has been specifically designed for speech recognition and delivers a very good speech quality. It is connected to the computer using an Andrea USB mini-soundboard and converter.

Finally, two general-purpose loudspeakers (Sony SRS-Z510 active loudspeakers) were selected, that could play the speech from a remote guide or any sound files. To fit all the devices within the console infrastructure, two separate consoles were designed. The first and largest console holds the touch screen, communication devices, and additional controls, and the second, smaller console the joystick. The large console consists of a vertical pillar-like construction holding a barebone computer and a UPS (uninterruptible power supply) unit to keep the installation running during brown-outs. Note that in many countries, power surges are a general occurrence: using a USV is of utmost importance to keep an installation running.

Both consoles (Figure 4.30) have been made from double-layered wood: an outer layer of beech wood glued and screwed to an inner MDF layer, forming a strong and rather heavy construction. The console with the joystick has a floor plate so that it can be mounted onto the floor, since it is much lighter and still rather high and, therefore, less stable.

Figure 4.30: *The Capsa Arcana console.*

During the design specific attention was placed on the height of both consoles, and the tilting level of the touch screen. As discussed in section 4.5.1, an important issue when dealing with hybrid interfaces is to analyze the different focal planes of the user. Users will need to focus their attention on multiple areas that generally do not lie in the same depth plane. As such, the switch between different focal planes should be limited where possible. Their focus will be on the GUI displayed on the touch screen and on the large stereoscopic projection wall placed several meters behind the console. Consequently, head rotation and eye focus should be carefully analyzed to see if different attention areas can be laid close together without negatively affecting factors like visual overlap. The angle of the touch screen and the placement of the console in front of the projection screen were adopted, such that the focal plane differences were as limited as possible without occluding the projection screen. Furthermore, the joystick could be operated in “eyes-off” modus; The user did not need to look at the joystick to control the navigation. Since the user would hold the joystick in the (dominant) hand most of the time, there was no need to search for it during interaction, even when touch screen interaction was mixed with navigation at a frequent rate. Even in darker environments, users could easily find the joystick without wasting too much focal attention on it. This statement may sound trivial, but it often occurs that users standing in projection systems have difficulties finding controls due to the limited illumination. Furthermore, the familiarity with using joysticks avoided most operating issues, since most users were used to control this kind of device. The control-body linkage was carefully investigated, observing the torque and force of arm movements in a standing pose. As a result, the height of the joystick console was such that users could ergonomically control navigation for longer times, without straining (fatiguing) the arm or wrist. Thereby, the arm is positioned at an approximately 110 degree angle, allowing the needed movements in an ergonomic way. Furthermore, the hand rests upon the joystick when not being used, relaxing the arm.

4.7.2 Adding unconventional controls using MIDI sensors

During the analysis and design of the console, one factor that came forward was the ability to extend the functionality of the console. Hence, the construction was laid out to not only support the basic remote guidance actions via standard input methods, but to further allow the application of more unconventional techniques. Furthermore, extensibility would help support reusability of the console in other application scenarios.
In order to support extensibility, a component-based input architecture (Figure 4.31) was developed. The aims of this architecture were:

- To mix or possibly replace traditional input methods by freely selectable new and possibly unconventional controls
- To extend the interaction framework, allowing not only voluntary, but also involuntary control methods

By extending the control possibilities through allowing usage of less conventional interaction methods, it was hypothesized that users would be more excited while using the console, thereby possibly promoting increased learning of these users. Furthermore, within the DHX project, so called virtual storytelling mechanisms are applied that could potentially be triggered by the involuntary controls.
As a result, the console would allow three kinds of control: “normal” control, extended control using unconventional input methods, and input using involuntary (bio)control methods.

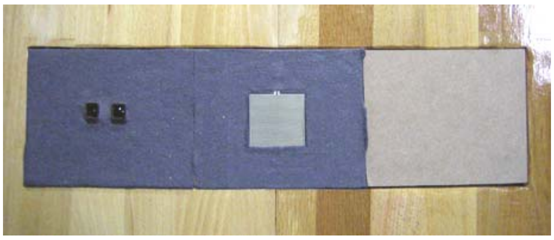


Figure 4.31: *MIDI component infrastructure.*

The component-based architecture is a rather simple structure of small box-like parts that can be placed in the console and are held steady by the weight of the “hood” of the wooden construction. Boxes can hold anything from a garage-interface (see section 3.7) to any analogue or digital input. To support a wide range of inputs, a MIDI interface was chosen. MIDI inputs are widely available, cheap, and can be bought in industrial (robust) quality. MIDI controls have been widely used by musicians, making arbitrary musical controls out of them. In the current console, a variety of sensors from Infusion Systems is integrated (Figure 4.32). This firm offers a large number of sensors and actuators that can be easily connected. For this purpose, an 8-port miniDIG analogue to MIDI converter is placed inside the hood of the console, to which the MIDI sensors can easily be connected. The miniDIG is connected to a MIDI-converter (Midisport 4x4) which sends the MIDI signals to a USB port. From there they are fed into the MIDI software interface of the AVANGO software environment (Tramberend '01) (Conrad, Krueger et al. '04).

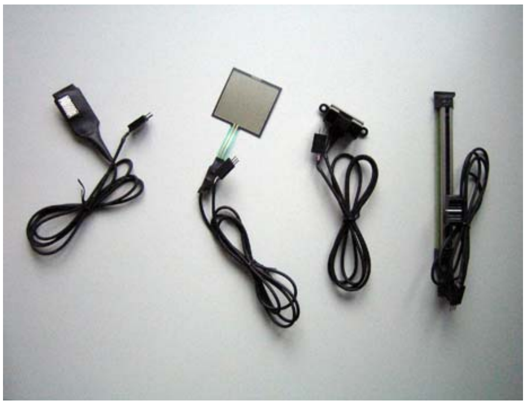


Figure 4.32: *Sensors from Infusion Systems.*

The different directly controlled sensors that are currently supported in the console allow for basically two kinds of actions.

- *General control*: different knobs and sliders allow for the more traditional way of providing a numerical value, which can be interpreted for a variety of purposes. These knobs or sliders can theoretically be hidden in larger objects to allow for more playful interaction. An example is “Duplo-like” controls which are large and colorful sturdy objects that have sensors inside. Such controls are regularly used in public space installations for kids. Besides these mechanical controls, users can make use of IR-based proximity sensors to perform gestural interaction. By moving the hand to and from the sensors, the distance between the hand and the sensor can be used to provide numerical input. The numerical values gathered by any of these sensors can be mapped to a multitude of different actions. One example is to define height during navigation by moving the hand up and down. Another is to make use of two IR-proximity sensors to mimic an airplane by moving the left and right hand independently up and down (quasi like the wing of the airplane) to turn left or right (such as also applied by Doulis et al (Doulis, Zwimpfer et al. '06).

- *Pressure sensitive control:* in addition to general manipulation or navigation actions, users can make use of two kinds of sensors that allow for pressure-sensitive control. The first sensor is a force sensitive resistor which can sense forces ranging from 1 to 100N in continuous mode. It can sense force changes over time (pushing) or individual taps. The second sensor is a bend-sensor, which senses flex angle of a small piece of tape. Even though not directly intended for pressure-based input, a specific device was made, which allowed for surface-oriented interaction (see “exploring haptic interfaces”).

The next kind of action, involuntary control, is currently based on a single sensor, a temperature and humidity sensor (thermistor). When held in contact to the user's skin, the sensor can sense the body temperature and sweat level changes. As handled in section 2.2.6, monitoring these values can be used to detect the user's psychological state, such as sensing the level of stress (Healey and Picard '05). The level of stress can be used to make assumptions on level of interest of the user, ranging from being bored up to excited. Having this information can be an extremely useful input for a storytelling engine, such as the one used in DHX (Conrad, Kruijff et al. '03). This engine makes use of a hierarchical finite state machine to structure actions in an application, storing behavior that can initiate specific kinds of actions. As such, receiving information on the mental state of a user can trigger new actions that are focused on changing this state. This can be exemplified by an action game scenario: when it is sensed that the user is bored, visual effects, music, and maybe a couple of extra monsters might increase the excitement level of a user. A similar model can be used to increase excitement to promote learning: by restructuring the presented information or by providing additional features, users might be tempted to further explore an application (environment) in order to learn more.

Exploring simple haptic interfaces

With the previously mentioned bend sensor, several directions were explored to allow for interesting tactile / haptic input and output. An initial idea was to place the bend-sensor inside a flexible object, such as a tennis ball to allow for squeezing actions. The deformation of the surface of the flexible object brought some interesting ideas to mind. One of these ideas was tried out by making use of a flexible surface which could be pressed and sensed.

Figure 4.33: *Simple haptic interface.*

By mounting a piece of flexible material (a part of a kitchen rubber glove) inside one of the exchangeable interaction blocks of the component-based infrastructure, a surface was made available that could be pressed rather far (Figure 4.33). By placing the bend sensor in the middle of the surface, the penetration depth can be measured. The sensation created by pressing the surface is both strange and rather exciting. The sensor can be used for manipulation actions, for example to deform an object. On the other hand, the sensor proves an interesting alternative for a button that can be pressed. Even though the multiple levels of pressure put on the surface may not make any sense, for kids it can mean great pleasure.

After using the flexible pressure sensitive surface control for a while, a new idea came to mind, visualized in the concept drawing below. For specific application areas, it makes sense to be able to simulate the stiffness of a surface when pressing it. An example is to simulate the stiffness of skin in medical scenarios. The idea behind the conceptualized device (Figure 4.34) is actually rather straightforward: by using an actuator (a multiphase motor), the corners of a surface can be pulled outwards (downwards), changing the stiffness of the surface itself. The harder the motor pulls, the stiffer the surface gets. Such a device can easily be used in connection with a bend sensor, to sense the depth of penetration as a result of the pressure of the user's finger on the surface. A second, more advanced possibility is to make use of a grid of potentiometers laid out below the surface to detect exactly where the user is pressing, and to which extent. The grid of potentiometers would resemble devices such as the Feelex from Iwata et al (Iwata, Yano et al. '01).

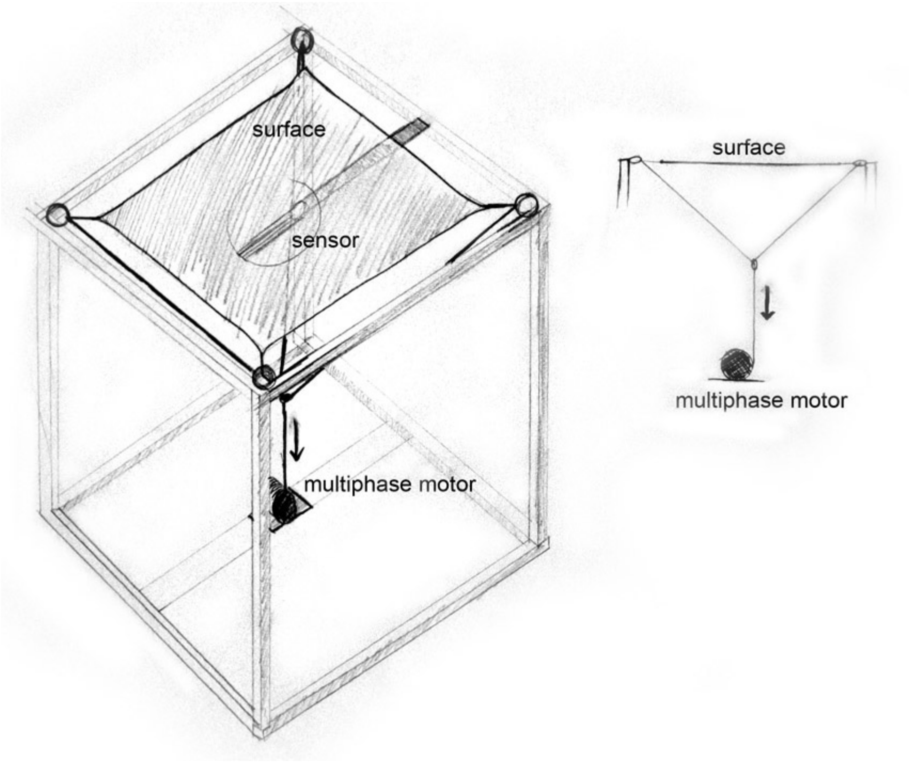


Figure 4.34: *Concept drawing of haptic surface device.*

4.7.3 Experiment setup

The interaction concepts of DHX were presented in several public demonstrations, showing four out of five developed heritage scenarios. At the Natural History Museum of Crete, 210 people experienced the DHX scenarios. At CNR-ISTI's Science Week, around 100 people participated. The largest demonstration in which the Capsa Arcana itself was also used, took place at the Piccolo Theatre in Milan (Figure 4.35). The

Piccolo Theatre is one of the most prominent theatres worldwide and the display was shown to between 400 and 500 people. The evaluated demonstrations were not using distributed mechanisms, which had been demonstrated outside these evaluation sessions.

The users (subjects) had a wide variety of backgrounds, differing from museum directorates, general users, students, up to computer scientists. Demonstrations lasted around 10-20 minutes. The evaluation results were retrieved from user observations (notes) and questionnaires using a 5-point Likert Scale.

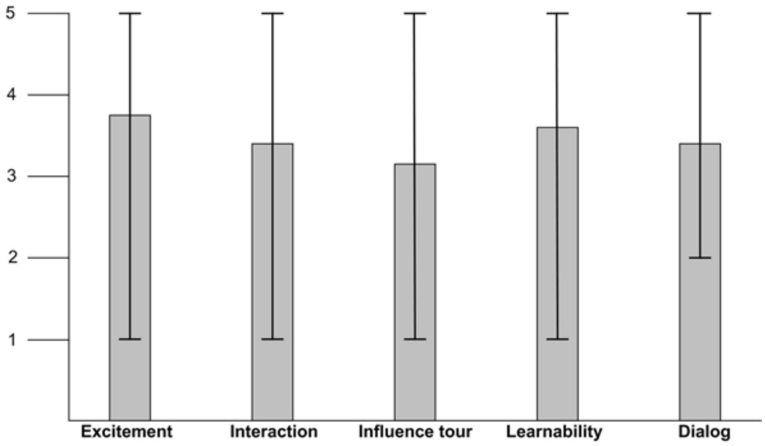
Figure 4.35: *The console at the Piccolo Theatre demonstration.*

During the demonstrations, the basic setup was used, consisting of a stereoscopic projection screen, touch screen control, and the joystick. Thus, no special sensors were tested during the evaluation - the focus was more on the guided interaction principles and the basic usability of the console.

4.7.4 Evaluation and results

A total of 105 questionnaires were collected from users and rated. The questionnaire included 6 questions, though for this analysis, one question concerning the content of the demonstration was left out, since it did not have direct effect on the interaction. An overview of the results can be found in Figure 4.36. Users were pretty excited about the demonstrations (avg. 3,78 / Stdev 1,14), showing the overall quality of the installation. Most users were satisfied with the interaction mechanisms (avg. 3,45 / Stdev. 1,25). Only few users (5,72 %) felt that they could not sufficiently control the application. Almost half of the people rated the interaction with good or even excellent (20% with note 4, 23,81% with note 5). A good number of the users had the feeling they could influence the tour (42,86% score 3, 28,57% score 4), though some users also gave worse scores (1,9% with score 1, 14,29% with score 2). These lower scores are easily explained by the fact that not all of the questioned users were able to make use of the interaction possibilities, since only a single user could control the application.

The majority of the users reported positively, up to very positively on the learnability of the demonstrations, even though some users of a specific scenario did not find the dialogs extensive enough. Thereby, the dialog mechanism was rated fairly high – 92,38% were satisfied with the offered method of information retrieval, with 40% of the users rating the mechanisms with a 4 or 5. This rating was supported by individual evaluation by museum guide experts from the German Federal Organization for Museum Pedagogy, which rated it as excellent.



	Avg	**Min**	**Max**	**Stdev**	**1**	**2**	**3**	**4**	**5**
Excitement	*3,46*	*1*	*5*	*1,25*	*2,86%*	*2,86%*	*43,81%*	*20,00%*	*23,81%*
Interaction	*3,61*	*1*	*5*	*1,02*	*1,90%*	*7,62%*	*42,86%*	*20,00%*	*25,71%*
Influence	*3,16*	*1*	*5*	*1,02*	*1,90%*	*14,29%*	*42,86%*	*28,57%*	*6,67%*
Learnability	*3,63*	*1*	*5*	*0,99*	*2,86%*	*6,67%*	*35,24%*	*32,38%*	*20,95%*
Dialog	*3,55*	*2*	*5*	*0,88*	*0%*	*5,71%*	*52,38%*	*20,00%*	*20,00%*

Figure 4.36: *Evaluation results of the demonstrations.*

4.7.5 Reflection

The design of public space systems poses several interesting constraints on the development of interactive techniques and I/O devices. The design of a console may seem straightforward, but do to the mixing with a second screen, the stereo wall, factors like focal attention need to be closely investigated. The usage of an off-the-shelve touch screen console would only have worked out to a limited extent in the scenarios it was applied in.

As reported in the previous section, the evaluation of the device and basic interaction mechanisms worked out fine. Users could easily make use of the dialog-based interaction mechanisms for controlling the application; Due to the resemblance with real-life communication with a museum guide, the interaction mechanism proved easy to learn and use, thereby supporting a short learning curve for the user. Even with the expert user watching over the actions of the museum visitors, users did not require much help.

In respect to chapter 3 issues, it should be stated that the evaluation results, in regard to the interaction mechanisms, were rather generalized. No comparative evaluation was held with other museum exhibition methods, and during the demonstrations, none of the unconventional methods were applied. Nonetheless, people were very much attracted by the interaction possibilities, which may seem to favorable point of the installation in comparison to general museum exhibition methods. Hence, the evaluation solely shows the quality of the hybrid interface, but not of any true unconventional interaction methods, even when the usage of virtual storytelling methods as basis for interaction is slightly experimental.

As such, the usage of the unconventional controls never left the experimental stage. To this respect, the case study primarily shows principles of designing and developing unconventional controls using simple sensors and some garage interface design methods. An evaluation comparing different kinds of unconventional sensors is intended, but not yet performed.

4.8 Eye of Ra

This case study focuses on the design of hybrid interaction techniques and a new input device for a virtual liver operation planning system. This system aims at supporting interactive visualization and manipulation of medical datasets for surgery planning based on a hybrid VR / Tablet PC user interface. The goal of the system is to facilitate efficient visual inspection and correction of surface models generated by automated segmentation algorithms based on x-ray computed tomography scans, needed for planning surgical resections of liver tumors (Bornik, Beichel et al. '06). The study specifically introduces concepts and devices that illuminate how unconventional technology can be integrated in conventional work environments.

4.8.1 Background and related work

Typically, the first step in liver surgery planning is segmentation of the individual structures, required to plan the surgical intervention. This task is normally done manually. However, this is tedious and time consuming, since it involves drawing contours on several hundred slices. A fully automated segmentation of the liver is difficult to achieve, because the shape of the human liver highly varies. Hence, in advanced automatic segmentation algorithms, the segmentation problems are usually limited to local errors, while most areas of the liver boundary can be correctly found using the automatic algorithms. A radiologist’s task can, therefore, be simplified from manual contour specification to interactively correcting errors in segmented datasets. This segmentation refinement approach is expected to be much less time consuming in most cases.

Normally, the segmentation is composed of three subtasks, namely model inspection, error marking and error correction. At a first glance 3D segmentation refinement tools seem to favor VR techniques; Stereoscopic visualization provides good 3D perception of the dataset, whereas tracked input devices allow for direct 3D interaction with the dataset. However, 2D screens have a much higher resolution than their 3D counterparts, and an inexpensive optical mouse easily outperforms high-end tracking devices in terms of accuracy, when precision input in 2D is required. In the medical field, where imprecision may have dire consequences, the virtues of established 2D techniques should not be discarded lightly. Moreover, physicians are used to desktop interfaces, and, in particular for system control, VR interfaces are not yet mature.

These considerations lead to the design of a hybrid user interface that combines multiple display and interaction techniques in order to match the work processes at hand. The objective of the hybrid user interface is to pair 3D perception and direct 3D interaction with 2D system control and precise 2D interaction. For such an interface, it is important that the flow of action of working between 2D and 3D visualization and interaction techniques are not disturbed. Both the different views and the interaction with the data need to be handled coherently. To ease the transition between the interface modalities, a hybrid input device, which can be conveniently used in all 2D and 3D tasks, was designed and developed. A focus was placed on analyzing the differences between action performance in the 2D and 3D domain, leading to a more extensive human factors study. A more detailed description of this process and related tools can be found in (Bornik, Beichel et al. '06).

Hypotheses

- The tools developed for the liver planning application would benefit from an integrated 2D and 3D control device
- The control device would require a carefully designed form (grip) in order to function properly
- Users would be able to integrate 2D and 3D actions in a hybrid setup

4.8.2 Interface design and development

The hardware setup consists of two main parts (Figure 4.37), the 3D (VR) system and the 2D system. The VR system’s display is a large stereo wall (stereoscopic back projection screen, 375cm diameter, 1280x1024 pixels) viewed with shutter glasses. A Barco Galaxy 3-chip DLP projector provides high quality active stereo rendering with very good channel separation, which is important when displaying virtual objects close to the user. The stereo wall is driven by a PC workstation (dual 3GHz Xeon, NVIDIA Quadro FX 3400). Optical tracking of the user’s head and the input device is done using a 4-camera infrared system from ART. The desktop system is a Tablet PC (Toshiba Portege M200, 1.8 GHz CPU, GeForce Go 5200 graphics card, 12-inch TFT touch screen at 1400x1050 pixels). The Tablet PC is placed on a desk approximately 2 meters in front of the screen. The user is seated at the desk, so that both stereo wall and Tablet PC are within the field of view as shown in the Figure below.

Due to the different locations of the desktop display and the stereo wall in relation to the user, switching between desktop and spatial interaction (for example during mode change) necessarily results in a change of visual focus. Best practice demands that head rotation and focal plane difference are as limited as possible, without the desktop display occluding the stereo wall. In the current hardware setup, the Tablet PC is placed on a table, and tilted towards the user. The user can conveniently use the touch screen for selecting menu items or manipulating objects. The user’s arm may be placed on the table to reduce fatigue. The table is placed at a specific distance from the stereo wall, so that stereo objects are viewed in a depth plane that seems to be above or just behind the visuals viewed at the desktop screen. Consequently, both the angular movements of the head are limited, as well as the change of focus between depth planes. There are still field of view differences. Combining a smaller stereo wall with a larger touch screen in an L-shape like configuration may improve this issue. Nonetheless, the field of view issues, which have also been addressed in section 4.6.1, have largely been addressed.

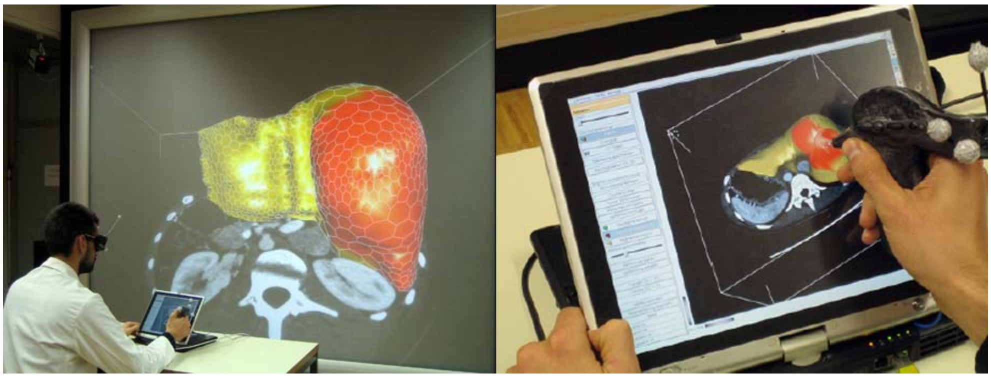


Figure 4.37: *Virtual liver planning system setup.*

To perform the surgical planning, there are several tools that support the user in a variety of actions.

- *Model Inspection* - The user tries to locate errors in the surface model by comparing raw CT data to the boundary of the surface. Thereby, clipping planes and different visualization modes can be made use of.

- *Error Marking* - Regions of the surface model that were found to be erroneous in the inspection step are marked for further processing. This allows restricting the following correction step to the erroneous regions and avoids accidentally modifying correct regions. To mark the erroneous parts, a resizable brush is available.

- *Error Correction* - Marked regions are corrected using special correction tools based on mesh deformation. Tools include sphere and plane-based deformation, and point-dragger tools.

The combination of 2D and 3D interactions can have considerable advantages. 2D actions can be performed with relatively high precision, whereas 3D actions are executed at high speeds in specific task situations. As such, a clear speed-accuracy trade-off can be noticed, depending on the task at hand. In that respect, the virtual liver planning application contains actions that are inherently 2D (like contour editing or point-based segmentation refinement) or 3D (including visual inspection of mixed data, or approximation of surfaces).

Performing the different actions in desktop or spatial mode necessarily leads to different kinds of input. In the spatial environment, most actions are coarse, mixed with some more fine-grained actions, whereas at the desktop, all actions are fine-grained. The different kinds of performance characteristics, and the necessity to make use of a pen-like device to control the touch screen lead to different kinds of dynamical coupling between hand and device. This is mostly caused by the different kinds of grips on the device that match the precision needed to perform the task.

Ergonomic device design

A new input device was sought that could combine high resolution and exact input with lower resolution but fast and probably more intuitive interaction. Hence, it should perform both spatial and constrained interaction, suitable for the variety of tasks in the liver planning system. In order to state the design parameters of the new device, a close analysis of the tasks and, therefore, hand-device couplings and movements was made. The usage of a two-handed device infrastructure was abandoned since the initial test setup mainly made use of the second device for reasons of having more buttons that could easily be fit into a single device.

As a result of the ergonomic study, it was clear to see that the device would require a form that allows for both power and precision grasps. The power grasp performs well for course actions, whereas the precision grasp is needed for fine-grained actions. For example, rotational movements for placing clipping plane or CT data plane are generally performed in high-speed / lower accuracy (sweeping task), whereas the usage of some tools, like deformation, are lower speed / higher accuracy and better performed in precision grip. Hereby, different interplays of the skeletal-muscular activities in wrist and forearm (power grasp) and foremost pointing finger and thumb (precision grasp) can be noticed (Balakrishnan and MacKenzie '97).

A finer task-movement analysis showed that for sweeping tasks like moving a plane in a dataset, users perform movements with the forearm (supination and pronation), flexion and extension, and some gliding movements of the wrist (see section 2.2 for more information). As such, the device should not only support a power grip, but also the right offset of the initial orientation for connection of (digital) tools to the device in the VE, in order to support the full rotation of the action. In practice this would mean that for example a clipping plane should be tilted to the right angle in relation to the top of the device, in order to perform the sweeping task in an ergonomic way (Goldstein '02) (Zhai '98a).

To get an idea of a basic form of the device and, therefore, the control-body linkage, a general device comparison was made, to which the movement and rotation patterns were matched. The general conclusion was that the hand activity consisted of characteristics normally matched by either a flying mouse or a pen-like device. Especially the latter was important, since the device needed to function as a pen-input device for the Tablet PC. However, the mix of both devices would probably lead to a special-purpose input device, with general purpose aspects, resulting in some specific design problems (Paley '98).

The approach followed was basically to morph the flying mouse and pen device forms into one single design, thereby allowing for an unobtrusive switch between power and precision grasps. Numerous plasticine models were made, ranging from pistol-form like devices up to strange forms fitting in the palm of the hand. Hereby, a small user group of different people with different sizes of hands, and hand preference (left or right-handed) were continuously used to evaluate the form. Also, the form would need to be large enough to fit in electronics. Hence, a search was made for small electronics, preferably with a wireless connection to reduce cabling. The final solution was found by using the EZ5 Wireless 5D Optical Pen Mouse, having a very small circuit board. The wireless connection was tested and found suitable when combined with a longer antenna in the device.

After much exploration, a device form was found (Figure 4.38) that allowed for easy switching between flying mouse and pen mode, within the ergonomic boundaries set. Due to the visual form of the device, it was nicknamed Eye of Ra. By pronating the forearm, and slightly changing the position of the fingers (mostly moving the thumb), the user could easily change between the different modes. This allows for dynamic coupling between device and hand without the user actively noticing it. One of the trade-offs of this design was, however, that different devices were needed for left and right-handed users. In order to create first prototypes of the devices, the left and right-handed plasticine models were fine-tuned and laid in plaster to make negative forms. The negative forms were used to make the final casing, by using carbon and fiberglass mats in a layered way, in combination with epoxy. This resulted in an extremely light but sturdy construction (also section 3.6). In the device, the button casing from the EZ5 was let in the housing, in order to make a stable connection between device casing and electronics.

Figure 4.38: T*he two grips afforded by the Eye of Ra.*

4.8.3 Experiment setup

The overall goal of the evaluation was to investigate the validity of the hybrid interface for liver surgery planning by comparing spatial (3D) and constrained (2D) tasks. Therefore, the evaluation was performed in two steps. In this case study, only the first evaluation step is reported on.

1. The first step examined the general spatial manipulation tools
2. In the second step specific constrained tasks like segmentation refinement based on local contour drawing are evaluated.

The different modes of the system, namely desktop, spatial interaction, or hybrid mode were tested in a comparative study. The evaluation included mostly empirical testing, with some analytical methods, using a variety of data collection methods. The evaluation addressed several relevant issues in complex interaction tasks, in particular learning curve effects and mode switching. The evaluation included both qualitative and quantitative components, in which the user attitude and psycho-physiological abilities were collected and analyzed. All results were cross-compared to see if there were any notable differences between the user attitude towards the system and the data collected through observation and recording.

The qualitative component of the evaluation was dominated by the subjective measurements obtained from the questionnaires, and the quasi thinking aloud protocols. The thinking aloud protocol was more or less a notification of the thoughts that were expressed by the subjects. The subjects were asked to speak, but not forced. As such, results from the thinking aloud protocols differed between users, since expression (explanation) levels differed between persons. The questionnaire was mostly focused on the user satisfaction, by validating 17 questions. The main factors included were user learning curve, attitude towards tools, ease of use and effectiveness of tools, user comfort, including fatigue and device ergonomics, and attention. Hence, the questionnaires focused on the main issues specified in the hybrid interaction methodology applied. The quantitative data was collected from the external observer's notes, the quality of the final liver model delivered by the subjects, and the logging files that tracked duration and changes of interaction modes. The observer noted all question asked by the user as well as user behavior (grasps, observable dexterity in fingers and wrist, arm-hand steadiness, attention to desktop and projection screen). Furthermore, the work-flow was observed and later compared with logs. Finally, a comparison was performed between the data produced by the subjects, and the best-practice model provided by an expert user.

The different steps of the evaluation were as follows: Subjects were first introduced to the system by the instructor, taking a 10 minute tour through the software. After the introduction, users could make use of the system for 12 minutes and ask questions. Next, the tests were performed. Users were instructed to only ask questions when absolutely necessary. All questions were recorded and analyzed to get an impression of the learning behavior.

The evaluation consisted of a basic set of functions, focusing on the visual inspection and segmentation of a liver dataset. The overall goal was to perform segmentation refinement on pre-segmented liver models that had artificially induced segmentation errors. The errors were designed to be obvious, even to novices, after introduction to the problem domain and system function. Three steps had to be performed: visual inspection, failure marking, and segmentation refinement at the identified erroneous spot. Hence, the task consisted of navigation (zooming and camera movement actions), mesh marking, and segmentation modification (mesh contour clipping, mesh freezing and unfreezing, plane and sphere deformation, and surface mesh deformation). Overall, the task can be categorized as highly complex. The tools to perform the task were either used in desktop mode, or spatial mode, or in a mixed way. Tool selection always took place in the menu on the desktop. Zooming, camera movement, and the placement of the cutting plane in spatial mode were possible by pressing a button on the input device.

4.8.4 Evaluation and results

Eighteen subjects (11 female, 7 male) aged between 21 and 46 participated in the evaluation. All users had a medical background. Most of them were students of human medicine, whereas some of the users already had more extensive medical skills. Their experience with computer systems differed widely, from only incidental usage to expert users. Most users had only used a mouse before; About half of the users had also used a touch screen. None had any real experience with immersive environments. A complete evaluation session took about one hour per subject. The questionnaires used a 7 point Likert scale. In order to get an impression of the users' attitudes, both averages and standard deviations were calculated. The results (Figure 4.39) were compared with the protocol of the observer that included thinking-aloud statements and the analysis of dataset results.

For the task set given to them, the subjects generally preferred the 3D above the 2D tool set. Overall, 37% were very satisfied with the ease of use of the 2D tool set, whereas 63% were satisfied with the 3D tool set. The mean values showed that most tools were rated in the range of very much acceptable up to very good (means between 5 and 6). Only a limited amount of tools performed poorly in either of the modalities. However, this user preference data is highly biased by the tasks evaluated, which were inherently 3D tasks. Tools for less complex actions (visual inspection tools) were rated considerably higher than for more complex ones (segmentation refinement tools). Clearly the deformation with the plane was not very well rated in both modalities, whereas deformation with the sphere was very well rated in 3D. The error marking task was perceived as being able to be performed very well in both 2D and 3D, which was confirmed by external observations. Figure 4.39 gives an overview the percentage of user satisfied with individual tools and the average tool rating concerning ease of use and effectiveness. In a direct comparison of separate tool preference, 70% voted for 3D, 16.4% for desktop, and only 13.6% for mixed tools. The overall preference for complete tool sets was almost exactly split between mixed and 3D tool sets, being in line with the expectation that the 3D tool set would be preferred above 2D tool set. When interpreting the results, it should be taken into consideration that the tasks used in the evaluation have strong 3D characteristics. Therefore they illustrate only certain aspects of the complex overall liver surgery planning system.

	Easiness				***Effectiveness***			
Tool	***2D***		***3D***		***2D***		***3D***	
Zooming	*64%*	*5.7*	*89%*	*6.6*	*53%*	*5.2*	*78%*	*6.2*
Camera move	*62%*	*5.1*	*88%*	*6.5*	*35%*	*4.6*	*71%*	*5.9*
Move C-plane	*27%*	*4.5*	*69%*	*6.4*	*27%*	*4.5*	*71%*	*5.8*
Mark surface	*64%*	*5.7*	*78%*	*6.0*	*57%*	*5.3*	*71%*	*6.0*
Plane def.	*9%*	*3.0*	*37%*	*4.8*	*18%*	*3.4*	*33%*	*4.6*
Sphere def.	*53%*	*5.2*	*94%*	*6.2*	*36%*	*4.5*	*61%*	*5.6*
Free def.	*25%*	*3.7*	*40%*	*5.2*	*13%*	*4.1*	*30%*	*4.9*
Point dragging	*14%*	*3.4*	*33%*	*4.6*	*13%*	*3.8*	*27%*	*3.9*

Figure 4.39: *Results (averages) of the questionnaires, percentages showing very satisfied users.*

Tool performance

Looking at the actual performance of tools, users noted few problems on basic actions like rotation or translation. The mean values for object rotation (avg. 5.33 in 2D and avg. 5.94) and translation (avg. 5.00 in 2D, avg. 5.78 in 3D) where rather high. Comparing truly satisfied users revealed that performance of atomic actions was far more appreciated in 3D (around 70% truly satisfied) than in 2D (only around 31% truly satisfied). With respect to precision of performance, 3D was marked much higher than 2D interaction. A highly diverse mix of user feedback could be noted: 44% were truly satisfied with the precision in 3D, against only 6% in 2D. Average marks were mediocre: a rating of avg. 3.22 for the 2D environment, and avg. 4.67 for the spatial setup. The mediocre marks for the 2D precision probably were caused by the size of the desktop display, and it can be expected to increase considerably by using a larger touch screen display. Looking at the level of visual details, 53% were very satisfied with the desktop display, whereas the stereo wall got 83% full satisfaction. Furthermore, the better marks for the precision of 3D tools could have been biased by the ease of interaction. Through user experience and by using a larger 2D display, marks would be expected to improve and become more level in comparison to each other.

User observation

User observations showed a rather diverse image of performance: some users could correct the error extremely fast in 3D. Well performing users obviously made use of strategy taught to them in the introduction. Some users had too many problems with the complexity of the tasks and tools, such that they could not even apply the strategy. Regarding flow of action during action performance, some users had problems with switching between navigation and manipulation modes in the desktop interface. This disturbance was caused since users forgot to click on a specific button that was placed inconsistently in the user interface – a problem that can easily be solved.

Due to the complexity of the evaluated task, it was expected that understanding the task and its tools would greatly influence the performance of using the tools. Learning how to use the tools turned out to be not so easy, but not too hard; Both user observations and user satisfaction showed learning curve issues. 35% were completely satisfied with the speed of learning the 2D interface, in the 3D interface this was even 56%. Average marks reflected the complexity though. 2D scored an average of avg. 4.47, whereas 3D scored avg. 5.17. Subjects often clearly noted that they could have used the more complex tools rather easily, if they had been given more practicing time, which was confirmed by user observations. Though the users did not always quickly learn how to perform an action, they seemed well enough informed on what they were doing, stating only little mode errors or problems with feedback. 63% were completely satisfied with feedback in 2D, against 78% in 3D (avg. 5.59 in 2D, avg. 5.89 in 3D). Experienced computer users did not always learn to use the tools faster and did not necessarily perform better. As such, the classic experience-performance tradeoff was not always fully true in this experiment, possibly affected by the complexity of the task and necessity of understanding the medical data and performance strategies.

Focal attention

The majority of users did not have a problem of switching between focusing on the large stereo wall and the desktop. 50% were completely satisfied, and 56% noted they noticed no problems at all with changing between desktop and projection screen for interaction purposes. The averages for both issues were at 5.1. This mark shows that there is still space for improvement.

Input device acceptance

The new input device was well accepted. Users rated the weight of the device as being excellent (almost 100% satisfaction rate). This rating was especially good when compared to the actual duration of usage of the device (also in free-air) of about 40 minutes. About 56% were highly satisfied with the device ergonomics (average of 5.56). The mark might be biased by the fact that most of the users did only use a mouse before. Most users needed to get used to the new form of the device, since it is slightly unusual. Observations showed that users seemed to be performing well with the device, when focusing on the grips they used, and the amount of re-grasping, being a possible sign of problems. Most users seemed to handle the device very naturally, using dynamic coupling. Switching between flying mouse and pen-mode did not seem to cause any problems, and users often took an intermediate grip between two grip-modes, supported by the form of the device. The intermediate grip did not seem to be uncomfortable for the users.

Users noted some hand discomfort after 40 minutes of usage. As could be expected, for desktop interaction, more users were satisfied (59%) than for spatial operation (28%). The marks were highly variable in the rating of spatial interaction. Some tools in 3D (notably the plane deformation tool) were difficult to use, which definitely explained some of the worse results. However, the significance of the marks is unclear, since no comparison with any other 3D input device was made. Also, close observations of the users' hands by the external observer showed that most (but not all) users had very steady hands during operation. Hence, since device ergonomics were rated well, we tend to believe that users did not have significant problems with fatigue. Also, all users did not use 3D input devices before, so usage was probably very strange to them, making a comparative rating difficult.

4.8.5 Reflection

Visual inspection and segmentation refinement can be successfully performed using the developed system. The users' attitude towards the spatial tool set was better than to the desktop tool set, as expected. Less complex tools were rated higher on ease of use and effectiveness than more complex tools. Nonetheless, several of the tools did perform below acceptance level (below a mark of avg. 4.5), notably the plane and point deformation tools. The plane deformation tool seems to produce some ergonomic hand rotation problems that would require a redesign of the tool to make it more useful. It was interesting to see how users seemed to interpret precision in relation to the size of the display, and not in terms of the resolution of input. Even though the input on the touch screen was steadier, users felt better in control when working with the large model.

Continuing the focus on interaction flow factors of chapter 3, the majority of users did not have a problem mixing the modalities, even though it should be stated that most users predominantly worked in one of the modalities. Hence, effects were lower than if they had changed between the modalities continuously. Some users had problems with switching between modes, both at interaction and at focal level. All of these users also expressed learning problems, which may indicate that they were not used to working in a 3D environment. In particular, we observed problems with using the 2D system controls correctly. Most of these problems were related to learning deficiencies, since the problem was not the actual selection of the menu item, but rather which item needed to be selected.

In general, learning effects affected the outcome of the test to a large extent. To all users, working in a VE was new, and many users did not have much computer experience either. Hence, learning how to use the interfaces affected the usability of the system considerably. Users who quickly grasped the concept of solving the task could effectively solve the task within time, even sometimes much faster than expected. Several users solved the problem within 7 minutes, which was extremely fast. Through longer duration evaluation sessions, learning effects would affect the marks in a positive way.

Looking back at the higher level design issues discussed in section 4.6.1., one effect that was visible again was that user's generally like to stick to a specific modality, even if another modality might prove to be more effective. That is, some users would stick to a 3D tool, even when the 2D tool was known to perform better. Hence, clever strategies with which the actions can be performed in hybrid interfaces need to be learned by a user. To which extend the spatial capabilities of a user limit the want for switching between modalities is unknown. It can be assumed that switching between different kinds of representations requires “rethinking” to interpret the information correctly, which may be unwanted by the user.

Finally, even though the users did not switch regularly between modalities to perform manipulation actions, the integration of 2D system control with spatial interaction seemed to work out fine. Since the users were forced to make use of the menu at the Tablet PC, specific feedback was automatically noticed by the user, and not ignored or unseen, as happened with the ProViT application.

4.9 Summary

Throughout this chapter, seven case studies have reflected a multitude of factors handled in chapters 2 and 3. This section provides a synopsis of these tests and their results.

Somatic and kinesthetic feedback: In the Shockwaves, BioHaptics, and Tactylus studies, several different (experimental) haptic or pseudo-haptic feedback methods are presented. Shockwaves introduces the usage of sound and air-based shockwaves to provide pseudo-haptic feedback to groups of users. Sound-based shockwaves using low frequency range loudspeakers and in-floor vibration speakers produced suitable results. The usage of air-based shockwaves posed distinct problems, since the air propulsion was too low to be fully noticed.

BioHaptics focuses on the usage of electric stimuli to trigger muscle responses. Small muscular contractions could be established, which were interpreted by users as haptic-like feedback. A more advanced system could potentially lead to largely replicating haptic feedback of grounded devices, but would require suitable model of muscular behavior and better calibration methods. Furthermore, next to the triggering of muscles, the electric stimulation of skin receptors would be both interesting and suitable.

Tactylus shows how visual audio and vibrotactile cues can replace (substitute) traditional haptic feedback, using grounded devices applied in collision detection and texture recognition scenarios.

Biopotential systems: In BioHaptics, the topic of human biopotential is addressed by triggering the motor nerves (muscle endings) of a user for feedback purposes, in contrast to receptor-level stimulation applied in general haptic systems. It is envisioned that the triggering of muscle endings for feedback could be well combined with muscle-based control of applications, defining a whole new kind of haptic I/O system.

Multisensory processing: In the Tactylus study, multisensory feedback is presented that combine visual, auditory and vibrotactile information, showing how these methods can bias or strengthen perception. The test confirms that integration of visual, auditory, and tactile feedback is preferred by users for more complex collision detection tasks. Furthermore, in the texture recognition test, the most prominent result is that audio can potentially bias visual perception.

Sensory and control substitution, addition and integration: In Shockwaves and Tactylus, feedback methods are presented that substitute haptic feedback. Shockwaves demonstrates how bone-conduction and the vibration of human organs can lead to haptic sensations other than caused by grounded devices. Particular for the Tactylus study is, how sensory channels can be combined to support integrated feedback perception.

Flow of action and feedback mechanisms: In the ProViT and Capsa Arcana studies, interaction flow is observed, focusing on how different input and output devices can be combined, among aiming at the reduction of cognitive load and problems with focal attention. Particularly ProViT shows that hybrid interfaces can potentially increase flow of action in more complex applications.

General application and transfer issues: Especially in ProViT, Capsa Arcana, and Eye of Ra, general application and transfer issues are addressed, as a result of the close connection to usage domains. All three installations make use of a hybrid setup, combining more traditional interaction methods and new, partly unconventional and/or spatial control methods. Evaluations show that hybrid setups are well usable, especially when first-time users interact with functionally less complex spatial display systems.

Social and ethical issues: The BioHaptics study shows that rather experimental characteristics of interfaces (in this case small electric stimuli) can still be acceptable to users. The experiment environment was rather private and game-like, but the majority of users reported potential further usage outside this setting.

Garage Interface design methods: The Shockwaves, Tactylus, Capsa Arcana and Eye of Ra studies demonstrate how new devices can be built out of simple material, products and devices, or toolkits. Rapid prototyping methods or simple woodcutting methods were used to produce the housing for the devices.

Evaluation: In the Cubic Mouse study, specific focus was put on the performance evaluation of 3D control devices, represented by a trajectory analysis tool. The tool was successfully used in a performance test comparing fine-grain and coarse actions of multiple control devices (Cubic Mouse versus Polhemus Stylus and Pinchgloves).

CHAPTER 5
Conclusion

The ultimate goal is to find out how the potential of the human body can be used to design, develop and analyze new spatial interaction methods that surpass performance or application possibilities of currently available techniques

5.1 Reflection

Within this dissertation, a variety of issues have been presented that illuminate the design, development, and analysis of unconventional 3DUIs from different perspectives. The core of the content has been formed by three major blocks: the potential of the human I/O system with its incredible breath of possibilities for application (chapter 2), specific factors that affect the design and development of more unconventional or experimental interfaces (chapter 3), and finally a series of seven case studies that focused on several factors, including the usage of human potential or the experimental design of devices using garage interface design techniques (chapter 4).
Throughout this dissertation, several key 3DUI directions have been addressed (also see section 1.5):

- *More advanced I/O devices*: in this dissertation, many new and/or technologically advanced devices were presented, both from other researchers (chapter 2), and as a result of our work (chapter 4). Our techniques (devices) specifically focus on increased ergonomics (Tactylus, Eye of Ra), the usage of sensory substitution methods (Shockwaves, Tactylus), or the combination of 2D and spatial techniques in hybrid interfaces (see next point).
- *Mixing 2D and 3D techniques*: three case studies focus on the combination of 2D and spatial techniques in so called hybrid interaction techniques (chapter 3 and 4, ProViT, Capsa Arcana, Eye of Ra) in order to advance interaction in tasks that are not solely of spatial nature.
- *Making VEs more realistic*: several techniques are presented that focus on the combination of multiple sensory or control systems in order to create more “vivid” interactive environments (specifically Shockwaves, Tactylus, Capsa Arcana, and BioHaptics).
- *Making 3D user interaction “easier and better”:* this dissertation focused specifically on making interaction better by looking at specific factors such as the advantages, disadvantages and problems of multisensory processing (Tactylus), flow of action in complex applications (ProViT), and the development of advanced feedback mechanisms (Shockwaves, BioHaptics, Tactylus, Capsa Arcana). This included the work at *specialized interaction techniques* such as those focusing on the exploration of textures (Tactylus). Guidelines and /or explanations were included that make it easier for other researcher to replicate results (chapter 3).

One major question, which runs through all chapters is: what is actually unconventional? Or, is there actually something like *unconventionalism*? Even though introduced in the preface (see: possible axes of unconventionalism), to the author's impression, unconventionalism is in the eye of the beholder. A technique or device may seem highly unconventional to one user, whereas for the other, it may be the only or least-unnatural way to interact. One such area that has thrived on unconventional techniques that fit well within the human potential view is the field of assistive technology, with techniques such as biopotential interfaces. In any case, the majority of techniques presented in this work have not yet hit the "mainstream" of application in VEs or in those interfaces used by the public in daily usage. Hence, they show, each in their own respect, a high(er) level of experimentalism.

Another problem with describing the techniques presented in this work as unconventional is the effect of time. Over the last years, technological developments have been incredible, and some devices have found their way into the daily life of the user faster than can be imagined. A mobile phone resembling a small multimedia computer is just one example and techniques like the EyeToy another.

Hence, what has also been noticed in the case studies themselves: not all studies exhibit the same level of unconventionalism. Whereas some of the hybrid interfaces show techniques that are partly also known in mainstream developments within the 3DUI area, others like the BioHaptics study are more on the border of experimentalism.

A second issue which should be handled in the conclusion is the "*completeness*" of this work. This dissertation has never been intended to give a complete overview of all possibilities and devices that would follow out of observing all developments from a human I/O potential point of view. New devices and ideas for techniques keep appearing on a daily basis and could never be captured within a single study. However, the studies can be a great incentive for further studies, providing a good starting place for developing new techniques. Especially since this work is based on *human potential*, its content will have a longer validity. Basically, the psycho-physiological possibilities will not change too quickly, quite simply because of the speed of human evolution. Hence, looking at what is possible by the human body may yield a great basis for creating new techniques. Even more important, I hope that 3D user interfaces can be further enhanced, in ways that interaction becomes better and "richer" (accurate, exciting) through more advanced control and feedback possibilities. I foresee new ways of interaction that are currently hardly covered, as well as new or changed fields of application.

A third issue not to be disregarded when developing new and possibly unconventional techniques is *evaluation*. This dissertation shows a large amount of evaluation results, both out of other studies (especially in chapters 2 and 3) and through our studies (chapter 4). Some of the presented case studies made use of a basic evaluation that only tested the affinity of the users towards the technology or device, whereas in others (like the Eye of Ra study) exhibit extensive user studies taking a detailed look at user performance and other human factors. Hence, some case studies should be seen as incentive for further studies. For example the BioHaptics study exhibits great complexity which needs to be researched within a large study (or studies), which would be outside the boundaries of this work.

5.2 Contributions

- The **design space** of spatial interfaces has been extended by identifying new possibilities for creating (unconventional) 3DUI techniques. This is achieved by providing a comprehensive **investigation of human I/O potential**, analyzing the psycho-physiological background, available technology and possible application areas matching this potential.

- Based on our experience and studies, evaluations and background investigations, **guidelines for designing and developing unconventional 3DUI techniques** have been provided, while also providing ways for **porting** these interfaces to general / more traditional work environments through hybrid interface methods.

- **A new input device** (*Tactylus*) has been presented which **combines visual, vibrotactile, and auditory feedback** to successfully support collision detection and texture recognition. The design of the device is based on the premise of sensory substitution, replacing haptic feedback through multisensory binding of alternate feedback methods, in this case vibrotaction and audio coupled to visual output.

- Two techniques have been presented that make use of the potential of the human body to sense **haptic feedback via alternative methods**. Both methods currently deliver partly pseudo-haptic feedback by not fully replicating common haptic feedback methods. The *Shockwaves* study shows how pseudo-haptic sensations can be generated in groups of people by using the vibration capacities of organs and bone structures. The study also showed a new idea of using air-propulsion for haptic feedback. Current efforts were not successful, hindered more by technical nature, rather than by the actual effect, which should be reproducible by a suitable device.
  The second technique, *BioHaptics*, makes use of neuroelectrical muscular stimulation to contract muscles, in order to provide partial haptic feedback to users. A first evaluation showed a positive tendency, but the study requires further evaluation, especially focusing on triggering the right combination of muscles to obtain an actual involuntary movement of the arm, as would normally be caused by an external device. What can be taken from the test is, that the technique has good potential as warning mechanism in applications that require more “extreme” forms of feedback, for example for safety reasons.

- A **performance study** comparing the *Cubic Mouse* with a Stylus and gloves showed the strength and preference (83%) of users and for using the “prop” device for controlling fine-grain actions, but also illuminated its deficiencies for coarse actions. The study also presented **a new trajectory analysis method** using 3D movement paths logged during the evaluation.

- Several studies focused on integrating unconventional interaction methods in more traditional, possibly desktop work environments by using **hybrid interface techniques**. The *ProViT* study focused predominantly at flow of action factors, including the chunking of actions, device switching influences, and focal attention factors. The results of this study were used in the second device construction, the Capsa Arcana, investigating the usage of MIDI sensors in a more traditional console form for usage in public space. Hereby, several ideas for haptic (-like) sensors that are easily built using garage interface design techniques were presented. Both ProViT and *Capsa Arcana* were tested with end-users, showing very satisfactory results. Finally, with the *Eye of Ra*, a device is presented that integrates 2D and 3D functionality in a rather radical physical (device-) shape. The hybrid input device was developed for controlling a surgical planning application and showed excellent results in an extensive qualitative and quantitative evaluation.

5.3 Road map

This section provides some future directions that can be derived from the body of work presented in this dissertation. The directions are centered on the questions stated in introductory section.

What is the potential of the human input and output channels from a human-computer interaction perspective?

The overview showed that besides hands and eyes, the human body allows for much more control and feedback possibilities than is currently used. One field that will certainly create more attention is full-body interfaces. Though already existing since such installations as Videoplace (Krueger, Gionfriddo et al. '85) and present in many art installation such as shown at SIGGRAPH, only recently setups like EyeToy show larger interest in this field. The full-body interface also holds the great promise for “fully” connecting to the human potential. For now, a key direction seems to be the follow-up of sensory substitution to see which task can also be performed or perceived by an alternative body part or sensory system.
A second area that has received much attention lately is the field of biopotential. Mostly driven by brain-computer interfaces, biopotential may lead to radical changes on how human-computer interaction mechanisms are perceived, extending or even replacing the views that have dominated the research community since the mid eighties.

How and why can we use this potential in 3DUIs?

The potential of the human body can be used to create more effective and vivid interfaces, surpassing the “traditional” 3DUI, as is used in the majority of mixed reality environments. Furthermore, the potential can be used in demanding situations, in which sensory or motor channels are blocked or overloaded, or to support those that are unable to perform actions due to physical deficiencies. Of course, unconventionalism can also be used out of plain fun or for artistic expressions.

How does human potential drive the design of new and unconventional (hardware) interfaces?

Human potential drives the demand for new hardware that can actually read the output of the human body or trigger specific muscle-joint constructions, receptors, nerves or brain sections. Much of this hardware is not available and will need to be developed or considerably advanced. As previously mentioned, one area that has a high innovation potential is bio-interfaces. Especially when the direction of implants is further explored, currently unforeseen possibilities may be found. Not only will the integration of sensors and actuators close or even in the body expand, the embedding of sensor technology in everyday objects will probably spread out soon. This direction will probably be driven by such fields like ambient technology and tangible user interfaces. One further, largely open field is the direction of behavior-oriented interaction, which may demand a whole range of new devices. Finally, once new hardware device has been created, current market demand often requires integration in existing devices. Hence, some of the new technology will most likely need to be ported after the initial versions are available.

Which implications derive from the usage of this potential in spatial computer generated or adapted environments?

One factor has already been mentioned; The view on information processing in human computer interfaces will change through the availability of new techniques, such as based on biopotential. A second issue, which more or less forms the basis for this change, is actual “hardcore” evaluation. Evaluation often leads to identification of new problems and new potential, possibly starting up the whole cycle up again, creating new, currently unforeseen unconventional interfaces.

REFERENCES

3DCONNEXION (2005). Spacemouse website, available at: http://www.3dconnexion.com/spacemouseplus.htm. Last accessed September 15, 2005.

ABC (2006). Advanced Bionics Corporation website, available at: http://www.bionicear.com. Last accessed July 14, 2006.

ACTUALITY (2006). Actuality website, available at: http://www.actuality.com. Last accessed July 14, 2006.

ALLEN, J. (1995). *The TRAINS Project: A Case Study in Building a Conversational Planning Agent*. Journal of Experimental and Theoretical AI (JETAI) 7: 7-48.

ALON, G. and G. SMITH (2005). *Tolerance and conditioning to neuro-muscular electrical stimulation within and between sessions and gender*. Journal of Sports Science and Medicine 4: 395-405.

ALONSO, R., H. BLAIR-SMITH and A. HOPKINS (1963). *Some Aspects of the Logical Design of a Control Computer: A Case Study*. IEEE Transactions on Electronic Computers EC-12(6): 687-697.

ALTMANN, J. (2001). *Acoustic Weapons - A Prospective Assessment*. Science & Global Security 9: 165-234.

AMBIENCE (2005). Ambience project website, available at: http://www.hitech-projects.com/euprojects/ambience/. Last accessed October 21, 2005.

AMTA (2006). American Music Therapy Association website, available at: http://www.musictherapy.org/. Last accessed July 17, 2006.

ANIMAZOO (2005). Animazoo website, available at: http://www.animazoo.com. Last accessed September 28, 2005.

ARK, W., D. DRYER and D. LU (1999). *The Emotion Mouse*. Proceedings of HCI International '99.

ARSELECTRONICA (2005). Ars Electronica Humphrey II website, available at: http://www.aec.at/en/center/project.asp?iProjectID=12280. Last accessed September 25, 2005.

ASCENSION (2005). Ascension Technologies website, available at: http://www.ascension-tech.com. Last accessed September 20, 2005.

ATC (2004). American Technology Corporation website on hypersonic sound systems, available at: http://www.atcsd.com/hss.html. Last accessed June 12, 2004.

BALAKRISHNAN, R. and G. KURTENBACH (1999b). *Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces*. Proceedings of the 1999 ACM Conference on Human Factors in Computing Systems (CHI'99), ACM Press.

BALAKRISHNAN, R. and C. MACKENZIE (1997). *Performance Differences in the Fingers, Wrist and Forearm in Computer Input Control*. Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems, ACM Press.

BAYLISS, J. and D. BALLARD (2000). *A Virtual Reality Testbed for Brain-Computer Interface Research*. IEEE Transactions on Rehabilitation Engineering 8(2).

BEATY, B. (2004). Ring Vortex (smoke ring) Launchers, available at: http://www.amasci.com/amateur/vortgen.html. Last accessed December 17, 2004.

BECKHAUS, S., K. BLOM and M. HARINGER (2005). *Intuitive, Hands-free Travel Interfaces for Virtual Environments*. Proceedings of the 3D user interface workshop, IEEE Virtual Reality Conference (VR2005).

BECKHAUS, S. and E. KRUIJFF (2004). *Unconventional Human Computer Interfaces*. Course at SIGGRAPH 2004.

BERKHOUT, A. (1988). *A holographic approach to acoustic control*. Journal of Audio Engineering Society 36: 977-995.

BERSAK, D., G. MCDARBY, N. AUGENBLICK, P. MCDARBY, D. MCDONELL, B. MCDONALD and R. KARKUN (2001). *Intelligent Biofeedback using an Immersive Competitive Environment*. Proceedings of UbiComp 2001.

BIELIKOVÁ, M. (2002). *A Body-Monitoring System with EEG and EOG Sensors*. ERCIM News 51.

BIER, E., M. STONE, K. PIER, B. BUXTON and T. DEROSE (1993). *Toolglass and Magic Lenses: The See-Through Interface*. Proceedings of SIGGRAPH'93, ACM Press.

BIGGS, S. J. A. M. A. S. (2002). *Haptic Interfaces*. Handbook of Virtual Environments. K. Stanney, Lawrence Erlbaum**:** 93-115.

BILLINGHURST, M. (1998). *Put That Where? Voice and Gesture at the Graphic Interface*. Computer Graphics 32(4): 60-63.

BIOCONTROL (2005). Biocontrol website, available at: http://www.biocontrol.com. Last accessed October 5, 2005.

BIOSOMNIA (2005). Oxford Biosignals Biosomnia system website, available at: http://www.oxford-biosignals.com/solutions/prod_somnia.asp. Last accessed October 10, 2005.

BLAKE, R., K. SOBEL and W. JAMES (2004). *Neural Synergy Between Kinetic Vision and Touch*. Psychological Science 15(6).

BOBICK, A. (1997). *Movement, Activity and Action: the Role of Knowledge in the Perception of Motion*. Philosophical Transactions of the Royal Society 352: 1257-1265.

BOLT, R. (1980). *"Put-That-There": Voice and Gesture at the Graphics Interface*. Proceedings of SIGGRAPH'80, ACM Press.

BORNIK, A., R. BEICHEL, E. KRUIJFF, B. REITINGER and D. SCHMALSTIEG (2006). *A Hybrid User Interface for Manipulation of Volumetric Medical Data*. Proceedings of the 2006 Symposium on 3D user interfaces (3DUI 2006), IEEE Virtual Reality Conference (VR2006).

BOWDITCH, S., K. COX and J. NIPARKO (2003). *Auditory Rehabilitation: Hearing Aids*. Clinical Neurotology : Diagnosing and Managing Disorders of Hearing, Balance, and the Facial Nerve. L. Lustig and J. Niparko, Isis Medical Media**:** 277-290.

BOWMAN, D., E. KRUIJFF, J. LAVIOLA and I. POUPYREV (2001). *An Introduction to 3D User Interface Design*. Presence: Teleoperators and Virtual Environments 10(1).

BOWMAN, D., E. KRUIJFF, J. LAVIOLA and I. POUPYREV (2005). *3D user interfaces: theory and practice*, Addison-Wesley.

BOWMAN, D. A., D. KOLLER and L. F. HODGES (1997). *Travel in Immersive Virtual Environments: an Evaluation of Viewpoint Motion Control Techniques*. Proceedings of the 1997 IEEE Virtual Reality Annual International Symposium (VRAIS'97), Albuquerque, New Mexico.

BRAINOPERA (2004). Brain Opera website, available at: http://brainop.media.mit.edu/. Last accessed July 24, 2004.

BRESCIANI, J.-P., M. ERNST, K. DREWING, G. BOUYER, V. MAURY and A. KHEDDAR (2004). *Feeling what you Hear: Auditory Signals can modulate Tactile Tap Perception*. Experimental Brain Research 162: 172-180.

BRINDLEY, G. and W. LEWIN (1968). *The sensations produced by electrical stimulation of the visual cortex*. Journal of Physiology 196: 479-493.

BRIP (2004). Boston Retinal Implant Project website, available at: http://www.bostonretinalimplant.org. Last accessed June 15, 2004.

BULLINGER, H., W. BAUER and M. BRAUN (1997). *Virtual Environments*. Handbook of Human Factors and Ergonomics. G. Salvendy, John Wiley & Sons.

BULLINGER, H., P. KERN and M. BRAUN (1997). *Controls*. Handbook of Human Factors and Ergonomics. G. Salvendy, John Wiley & Sons.

BUOGUILA, L., M. ISHII and M. SATO (2000). *Multi-Modal Haptic Device for large-scale Virtual Environments*. Proceedings of the 2000 ACM international conference on multimedia.

BURDEA, G. C. (1996). *Force and Touch Feedback for Virtual Reality*, Wiley Interscience.

BUSSO, C., Z. DENG, S. YILDIRIM, M. BULUT, A. KAZEMZADEH, S. LEE, U. NEUMANN and S. NARAYANAN (2004). *Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information*. Proceedings of the 6th International Conference on Multimodal interfaces (ICMI'04).

BUXTON, B. (2005). Input Device Sources & Resources, available at: http://www.billbuxton.com/InputSources.html. Last accessed September 14, 2005.

BUXTON, W. (1986). *Chunking and Phrasing and the Design of Human-Computer Dialogues*. IFIP World Computer Congres, Dublin.

BUXTON, W. and B. MYERS (1986). *A Study in Two-handed Input*. Proceedings of the 1986 ACM Conference on Human Factors in Computing Systems (CHI'86), ACM Press.

CAIROS (2005). Cairos tracking systems website, available at: http://www.cairos.com/sports/index.php. Last accessed September 23, 2005.

CAMBRIDGE (2005). 2005. Last accessed September 13, 2005.

CAO, W., H. GAERTNER, S. CONRAD, E. KRUIJFF, D. LANGENBERG and R. SCHULTZ (2003). *Digital Product Development in a Distributed Virtual Environment*. VRAI, Proceedings of the SPIE.

CARD, S., J. MACKINLAY and G. ROBERTSON (1990). *The Design Space of Input Devices*. Proceedings of the 1990 ACM Conference on Human Factors in Computing Systems (CHI'90), ACM Press.

CARLIN, A., H. HOFFMAN and S. WEGHORST (1997). *Virtual Reality and Tactile Augmentation in the Treatment of Spider Phobia: A Case Report*. Behavior Research and Therapy 35(2): 153-159.

CAULKINGS, T., E. CORTEEL and O. WARUSFEL (2003). *Wave Field Synthesis Interaction with the Listening Environment, Improvements in the Reproduction of Virtual Sources situated inside the Listening Room*. Proceedings of the 6th Int. Conference on Digital Audio Effects (DAFx-03).

CHAN, C., M. LYONS and N. TETSUTANI (2003). *Mouthbrush: Drawing and Painting by Hand and Mouth*. ACM ICMI 2003.

CHANCE, S., F. GAUNET, A. BEALL and J. LOOMIS (1998). *Locomotion Mode Affects the Updating of Objects Encountered During Travel: The Contribution of Vestibular and Proprioceptive Inputs to Path Integration.* Presence: Teleoperators and Virtual Environments 7(2): 168-178.

CHEN, D. and R. VERTEGAAL (2004). *Using Mental Load for Managing Interruptions in Phsyiologically Attentive User Interfaces*. Proceedings of the Conference on Human Factors in Computing Systems (CHI'04).

CHENG, L.-T., R. KAZMAN and J. ROBINSON (1996). *Vibrotactile Feedback in delicate Virtual Reality Operations*. Proceedings of the Fourth ACM International Conference on Multimedia.

CHESKY, K. and D. MICHEL (1991). *The Music Vibration Table (MVTtm): Developing a technology and conceptual model for pain relief.* Music Therapy Perspectives 9: 32-38.

CIRCULAFLOOR (2004). CirculaFloor website at SIGGRAPH Emerging Technologies, available at: http://www.SIGGRAPH.org/s2004/conference/etech/index.php?pageID=conference. Last accessed Juli 14, 2006.

CLEVELAND, N. (1994). *Eyegaze Human-Computer Interface for People with Disabilities*. Proceedings of the First Automation Technology and Human Performance Conference.

COGAIN (2005). Overview of eye typing systems at the COGAIN project website, available at: http://www.cogain.org/links/eyetyping/. Last accessed September 14, 2005.

COHEN, I. and M. LEE (2002). *3D Body Reconstruction for Immersive Interaction*. Second International Workshop on Articulated Motion and Deformable Objects, Palma de Mallorca, Spain.

CONRAD, S., H. KRUEGER and M. HARINGER (2004). *Live Tuning of Virtual Environments: The VR-Tuner*. Proceedings of the Eurographics Symposium on Virtual Environments (EGVE04).

CONRAD, S., E. KRUIJFF, M. SUTTROP, F. HASENBRINK and A. LECHNER (2003). *A Storytelling Concept for Digital Heritage Exchange in Virtual Environments*. Proceedings of the 2003 International Conference on Virtual Storytelling.

COOK, J., D. FIELY and M. MCGOWAN (1995). *Nonlethal Weapons, Technologies, Legalities, and Potential Policies.* Airpower Journal Special Edition 1995.

COSTANZA, E., A. PERDOMA, S. INVERSO and R. ALLEN (2005). *Toward Subtle Intimate Interfaces for Mobile Devices using an EMG Controller*. Proceedings of the 2005 ACM Conference on Human Factors in Computing Systems (CHI2005).

CUTLER, L., B. FROEHLICH and P. HANRAHAN (1997). *Two-Handed Direct Manipulation on the Responsive Workbench*. Proceedings of the 1997 ACM Symposium on Interactive 3D Graphics (I3D'97), ACM Press.

CYBERKINETICS (2005). BrainGate (Cyberkinetics Neurotechnology Systems) website, available at: http://www.cyberkineticsinc.com/content/clinicaltrials/braingate_trials.jsp. Last accessed October 11, 2005.

CYBERLINK (2005). Cyberlink website, available at: http://www.brainfingers.com/index.html. Last accessed October 10, 2005.

CYBORG (2005). Cyborg 2.0 website at the University of Reading, available at: http://www.reading.ac.uk/KevinWarwick/html/project_cyborg_2_0.html. Last accessed October 25, 2005.

DARKEN, R. and H. CEVIK (1999a). *Map Usage in Virtual Environments: Orientation Issues*. Proceedings of IEEE Virtual Reality '99, IEEE Press.

DARKEN, R., W. COCKAYNE and D. CARMEIN (1997). *The Omni-Directional Treadmill: A Locomotion Device for Virtual Worlds*. Proceedings of the 1997 ACM Symposium on User Interface Software and Technology (UIST'97), ACM Press.

DAVIES, C. and J. HARRISON (1996). *Osmose: Towards Broadening the Aesthetics of Virtual Reality*. Computer Graphics 30(4): 25-28.

DEFANTI, T. A. D. S. (1977). *Final Report to the National Endowment of the Arts,* University of Illinois at Chicago.

DELIGIANNIDIS, L. and R. JACOB (2006). *The VR Scooter: Wind and Tactile Feedback Improve User Performance*. Proceedings of the 2006 Symposium on 3D user interfaces (3DUI 2006), IEEE Virtual Reality Conference.

DELSYS (2005). DelSys Incorporated products website, available at: http://www.delsys.com/products/products.htm. Last accessed October 10, 2005.

DIGIWALL (2006). DigiWall website, available at: http://www.digiwall.se. Last accessed July 14, 2006.

DINH, H., N. WALKER, C. SONG, A. KOBAYASHI and L. F. HODGES (1999). *Evaluating the Importance of Multi-Sensory Input on Learning and the Sense of Presence in Virtual Environments*. Proceedings of IEEE Virtual Reality 1999.

DISNEYQUEST (2006). DisneyQuest website, available at: http://disneyworld.disney.go.com/wdw/entertainment/entertainmentDetail?id=DisneyQuestIndoorInteractiveThemeParkEntertainmentPage&bhcp=1. Last accessed April 3, 2006.

DOBELLE, W. (2000). *Artificial vision for the blind by connecting a television camera to the visual cortex.* American Society of Artificial Internal Organs (ASAIO) 46: 3-9.

DOHERTY, G. and M. MASSINK (1999). *Continuous Interaction and Human Control*. Proceedings of the European Conference on Human Decision Modeling and Manual Control, Loughborough, Group-D Publications.

DOMBOIS, F. (2002). *Using Audification in Planetary Seismology*. 7th International Conference on Auditory Display, Espoo, Finland.

DOULIS, M., V. ZWIMPFER, J. PFLUGER, A. SIMON, C. STERN, T. HALDIMANN and C. JENNI (2006). *SpaceActor - Interface Prototypes for Virtual Environments*. 3D User Interfaces (3DUI'06).

DOURISH, P. (2001). *Where The Action Is: The Foundations of Embodied Interaction*, MIT Press.

DRS (2006). Design Research Society website, available at: http://www.designresearchsociety.org/. Last accessed August 2, 2006.

DUBOST, G. and A. TANAKA (2002). *A Wireless, Network-based Biosensor Interface for Music*. Proceedings of the 2002 International Computer Music Conference (ICMC'02),.

DUCHOWSKI, A. T. (2003). *Eye Tracking Methodology: Theory and Practice*. London, Springer-Verlag.

DURIC, Z., W. GRAY, R. HEISHMAN, F. LI, A. ROSENFELD, M. SCHOELLES, C. SCHUNN and H. WECHSLER (2002). *Integrating Perceptual and Cognitive Modeling for Adaptive and Intelligent Human-Computer Interaction*. Proceedings of the IEEE 90(7): 1272-1289.

EAGLEEYES (2005). EagleEyes project website, available at: http://www.bc.edu/schools/csom/eagleeyes/eagleeyes/. Last accessed September 13, 2005.

ECKEL, G. (1999). *Applications of the Cyberstage Spatial Sound Server*. Audio Engineering Society 16th International Conference on Spatial Sound Reproduction.

EEGSPECTRUM (2005). Therapeutic uses of biofeedback, available at EEG Spectrum: http://www.eegspectrum.com/Applications/InfoNetwork. Last accessed October 14, 2005.

EKSTROM, A., J. CAPLAN, E. HO, K. SHATTUCK, I. FRIED and M. KAHANA (2005). *Human Hippocampal Theta Activity During Virtual Navigation*. In press.

ELSENAAR, A. and R. SCHA (2002). *Electric Body Manipulation as Performance Art: A Historical Perspective*. Leonardo Music Journal 12: 17-28.

EMED-X (2005). Emed-X website, available at: http://www.novel.de/nav2/nav_2.htm. Last accessed September 23, 2005.

ERNST, M. and M. BANKS (2002). *Humans integrate visual and haptic information in a statistically optimal fashion*. Nature 415(6870): 429-433.

EYETOY (2005). EyeToy website, available at: http://www.eyetoy.com. Last accessed September 30, 2005.

FAKESPACE (2005). Fakespace pinchgloves website, available at: http://www.fakespace.com/pinch.htm. Last accessed September 23, 2005.

FELS, S. (1994). *Glove-Talk II: Mapping Hand Gestures to Speech using Neural Networks: An Approach to Building Adaptive Interfaces*, PhD Dissertation, University of Toronto.

FELS, S. and F. VOGT (2002). *Tooka: Explorations of Two Person Instruments*. 2nd International Conference on New Interfaces for Musical Expression (NIME02).

FELS, S., S. YOHANAN, S. TAKAHASHI, Y. KINOSHITA, K. FUNAHASHI, Y. TAKAMA and G. CHEN (2005). *Swimming Across the Pacific*. IEEE Computer Graphics and Applications 25(1): 24-32.

FERGUSON, S. and G. DUNLOP (2002). *Grasp Recognition from Myoelectric Signals*. Proceedings of the 2002 Australasian Conference on Robotics and Automation.

FIORENTINO, M., G. MONNO and A. UVA (2005). *The Senstylus: a Novel Rumble-Feedback Pen Device for CAD Application in Virtual Reality*. Proceedings of the 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2005, WSCG 2005.

FITTS, P. M. (1954). *The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement*. Journal of Experimental Psychology 47: 381-391.

FITZMAURICE, G., H. ISHII and W. BUXTON (1995). *Bricks: Laying the Foundations for Graspable User Interfaces*. Proceedings of the 1995 ACM Conference on Human Factors in Computing Systems (CHI'95), ACM Press.

FLETCHER, R. (1996). *Force transduction materials for human-technology interfaces*. IBM Systems Journal 35.

FOFONOFF, T., S. MARTEL, C. WISEMAN, R. DYER, I. HUNTER, L. HATSOPOULOS and J. DONOGHUE (2002). *A highly Flexible Manufacturing Technique for Microelectrode Array Fabrication*. Proceedings of the Second Joint EMBS/BMES Conference.

FOGSCREEN (2005). Fogscreen website: http://www.fogscreen.com. Last accessed July 5, 2005.

FOLEY, J., A. VAN DAM, S. FEINER, AND J. HUGHES (1996). *Computer Graphics: Principles and Practice*, Addison Wesley Publishing Company.

FOLGHERAITER, M., G. GINI and D. VERCESI (2005). *A Glove Interface with Tactile feeling Display for Humanoid Robotics and Virtual Reality systems*. International Conference ICINCO.

FONO, D. and R. VERTEGAAL (2005). *EyeWindows: Evaluation of Eye-Controlled Zooming Windows for Focus Selection*. Proceedings of the 2005 ACM Conference on Human Factors in Computing Systems (CHI 2005).

FORMAN, E. A. C. L. (2003). *Building Physical Interfaces: Making Computer Graphics Interactive*. SIGGRAPH Course #30.

FORSBERG, A., K. HERNDON and R. ZELEZNIK (1996). *Aperture Based Selection for Immersive Virtual Environments*. Proceedings of the 1996 ACM Symposium on User Interface Software and Technology (UIST'96), ACM Press.

FROEHLICH, B., J. PLATE, J. WIND, G. WESCHE and M. GOEBEL (2000). *Cubic-Mouse-Based Interaction in Virtual Environments*. IEEE Computer Graphics & Applications(July 2000).

FROEHLICH, B. A. J. P. (2000). *The Cubic Mouse: A New Device for Three-Dimensional Input*. Proceedings of the 2000 ACM Conference on Human Factors in Computing Systems (CHI 2000), ACM Press.

FROMHERZ, P. (2003). *Neuroelectronic Interfacing: Semiconductor Chips with Ion Channels, Nerve Cells, and Brain*. Nanoelectronics and Information Technology. R. Waser, Wiley-VCH Verlag Berlin**:** 781-810.

FUKUMOTO, M. and Y. TONOMURA (1999). *Whisper: a Wristwatch Style Wearable Handset*. Proceedings of the 1999 ACM Conference on Human Factors in Computing Systems (CHI '99).

G.TEC (2005). Guger Technologies (g.tec) products website, available at: http://www.gtec.at/products/product_overview.htm. Last accessed October 5, 2005.

GAVRILA, D. (1999). *The Visual Analysis of Human Movement: A Survey*. Computer Vision and Image Understanding 73(1): 82-99.

GERBERT-HIRT, S., W. HAUSER, W. RATHJEN and F. BREITSAMETER (2004). *Leben mit Ersatzteilen*, Deutsches Museum.

GILLEADE, K., A. DIX and J. ALLANSON (2005). *Affective Videogames and Modes of Affective Gaming: Assist me, Challenge Me, Emote Me*. Proceedings of DiGRA 2005.

GOBLE, J., K. HINCKLEY, R. PAUSCH, J. SNELL and N. KASSELL (1995). *Two-Handed Spatial Interface Tools for Neurosurgical Planning*. Computer 28(7): 20-26.

GOLDSTEIN, E. (2002). *Sensation and Perception*, Brooks Cole.

GOTHE, J., S. BRANDT, K. IRLBACHER, S. RÖRICHT, B. SABEL and B.-E. MEYER (2002). *Changes in Visual Cortex excitability in Blind Subjects as demonstrated by Transcranial Magnetic Stimulation*. Brain 125(3): 479-490.

GRANGER, M., J. LITTLE, E. ADAMS, C. BJÖRKMAN, D. GOTTERBARN, D. JUETTNER, C. MARTIN and F. YOUNG (1997). *Using Information Technology to integrate Social and Ethical Issues into the Computer Science and Information Systems Curriculum: Report of the ITiCSE '97 Working Group on Social and Ethical Issues in Computing Curricula*. ACM SIGCUE Outlook 25(4).

GRASSO, M., D. EBERT and T. FININ (1998). *The Integrality of Speech in Multimodal Interfaces*. ACM Transactions on Computer-Human Interaction 5(4): 303-325.

GREENBURG, S. A. C. F. (2001). *Phidgets: Easy Development of Physical Interfaces Through Physical Widgets*. Proceedings of the 2001 ACM Symposium on User Interface Software and Technology (UIST 2001), ACM Press.

GREENBURG, S. A. M. B. (2002). *Customizable Physical Interfaces for Interacting with Conventional Applications*. Proceedings of the 2002 ACM Symposium on User Interface Software and Technology (UIST 2002), ACM Press.

GREENHALGH, C. and S. BENFORD (1995). *MASSIVE: a Collaborative Virtual Environment for Teleconferencing*. ACM Transactions on Computer-Human Interaction (TOCHI 2(3).

GROSSMAN, T., R. BALAKRISHNAN, K. SINGH (2003). *An Interface for Creating and Manipulating Curves using a High Degree-of-Freedom Curve Input Device*. Proceedings of 2003 ACM Conference on Human Factors in Computing Systems (CHI 2003), ACM Press.

GUEST, S., C. CATMUR, D. LLOYD and C. SPENCE (2002). *Audiotactile Interactions in Roughness Perception*. Experimental Brain Research 146: 161-171.

GUGER, C., W. HARKAM, C. HERTNAES and G. PFURTSCHELLER (1999). *Prosthetic Control of an EEG-based Brain-Computer Interface (BCI)*. Proceedings of the AAATE 5th European Conference for the Advancement of Assistive Technology.

GUIARD, Y. (1987). *Symmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model*. The Journal of Motor Behaviour 19(4): 486-517.

GUISEPPI-ELIE, A. (2002). *Introduction to Biosensors. ENGR 645 Biosensors and Bioelectric Devices course material.*

GUNETTI, D. and C. PICARDI (2005). *Keystroke Analysis of Free Text*. ACM Transactions on Information and System Security (TISSEC) 8(3): 312-347.

HACHET, M., P. GUITTON, AND P. REUTER (2003). *The CAT for Efficient 2D and 3D Interaction as an Alternative to Mouse Adaptations*. Proceedings of the 2003 ACM Symposium on Virtual Reality Software and Technology (VRST 2003), ACM Press.

HAMNES, K. (2002). *Smelly Interfaces, a Brief Review of the Application of Smell in User Interfaces*. Fornebu, Future Media Group, Telenor R&D.

HAQUE (2005). Usman Haque homepage, available at: http://www.haque.co.uk/. Last accessed September 7, 2005.

HARRIS, L., M. JENKIN and D. ZIKOVITZ (1999). *Vestibular Cues and Virtual environments: Choosing the Magnitude of the Vestibular Cue*. Proceedings of IEEE Virtual Reality'99, IEEE Press.

HARRIS, L., M. JENKIN and D. ZIKOVITZ (2000). *Visual and Non-Visual Cues in the Perception of Linear Self Motion*. Experimental Brain Research 135.

HASAN, L., N. YU and J. PARADISO (2002). *The Termenova: A Hybrid Free-Gesture Interface*. Proceedings of the 2nd International Conference on New Interfaces for Musical Expression (NIME02).

HATANA, K., D. MASUI and H.-W. YEN (2003). *The Dimension Book*. Proceedings of ACM SIGGRAPH 2003.

HEALEY, J. and R. PICARD (2005). *Detecting Stress During Real-World Driving Tasks using Physiological Sensors*. IEEE Transactions on Intelligent Transportation Systems 6(2).

HINCKLEY, K., R. PAUSCH, J. GOBLE and N. KASSELL (1994). *A Survey of Design Issues in Spatial Input*. Proceedings of the 1994 ACM Conference on User Interface Software Technology (UIST `94), Marina del Rey, CA, USA, ACM.

HINCKLEY, K., R. PAUSCH, D. PROFFITT, J. PATTEN and N. KASSELL (1997b). *Cooperative Bimanual Action*. Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems (CHI'97), ACM Press.

HINCKLEY, K., R. PAUSCH, J. GOBLE, AND N. KASSELL (1994). *Passive Real-World Interfaces Props for Neurosurgical Visualization*. Proceedings of the 1994 ACM Conference on Human Factors in Computing Systems (CHI'94), ACM Press.

HIX, D. and H. HARTSON (1993). *Developing User Interfaces: Ensuring Usability Through Product & Process*. New York, John Wiley and Sons.

HOFFMAN, H. (1998). *Physically touching Virtual Objects using Tactile Augmentation enhances the Realism of Virtual Environments*. Proceedings of the 1998 IEEE Virtual Reality Annual International Symposium (VRAIS'98), IEEE.

HOLLERBACH, J. (2002). *Locomotion Interfaces*. Handbook of Virtual Environments. K. Stanney, Lawrence Erlbaum**:** 239-254.

HRL (2004). Holosonics Research Labs website, available at: http://www.holosonics.com. Last accessed June 14, 2004.

HUBBOLD, R. (2002). *Collaborative Stretcher carrying: a Case Study*. Proceedings of the Eight Eurographics Workshop on Virtual Environments (EGVE 2002).

HUITT, W. (1999). *Conation as an Important Factor of Mind*. Educational Psychology Interactive, available at: http://chiron.valdosta.edu/whuitt/materials/conation.html.

HUITT, W. (2003). *A Systems Model of Human Behavior*. Educational Psychology Interactive, available at: http://chiron.valdosta.edu/whuitt/materials/sysmdlo.html.

IGA, S. (1999). *Approximate Interaction: User Interface for the Ambiguous World*. Proceedings of SCI / ISAS '99.

IMMERSION (2005). Immersion website, available at: http://www.immersion.fr. Last accessed September 15, 2005.

INTEC (2005). Intec Wireless controllers website, available at: http://www.inteclink.com/inteclink/products/wireless/wireless.asp. Last accessed September 23, 2005.

ISHII, H. and B. ULLMER (1997). *Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms*. Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems (CHI'97), ACM Press.

ISHII, M., AND M. SATO (1994). *A 3D Spatial Interface Device Using Tensed Strings*. Presence: Teleoperators and Virtual Environments 3(1): 81-86.

ISO (2003). ISO Standard 2631 - Whole-body vibration, available at: http://www.iso.org/iso/en/CatalogueListPage.CatalogueList?ICS1=13&ICS2=160&ICS3=&scopelist=. Last accessed February 1, 2005.

ISTANCE, H., C. SPINNER and P. HOWARTH (2003). *Providing Motor-Impaired Users with Access to Standard Graphical User Interface (GUI) Software via Eye-Based Interaction*. Proceedings of the 1996 European Conference on Disability, Virtual Reality & Associated Technology (ECDVRAT'96).

IWATA, H. (2001). *GaitMaster: A Versatile Locomotion Interface for Uneven Virtual Terrain*. Proceedings of IEEE Virtual Reality 2001, IEEE Press.

IWATA, H., T. MORIYA, T. UEMURA and H. YANO (2003). *Food Simulator*. ACM SIGGRAPH Emerging Technologies.

IWATA, H., H. YANO, F. NAKAIZUMI and R. KAWAMURA (2001). *Project FEELEX: adding Haptic Surface to Graphics*. Proceedings of ACM SIGGRAPH 2001.

JACOB, R., J. SIBERT, D. MCFARLANE and M. PRESTON MULLEN (1994). *Integrality and seperability of input devices*. ACM Transaction on Computer-Human Interaction 1(1): 3-126.

JACOB, R. A. L. S. (1992). *The Perceptual Structure of Multidimensional Input Devices*. Proceedings of the 1992 ACM Conference on Human Factors and Computing Systems (CHI'92), ACM Press.

JACOB, R. J. K. (1995). *Eye Tracking in Advanced Interface Design*. Virtual Environments and Advanced Interface Design. W. Barfield and T. A. Furness. New York, Oxford University Press**:** 258-288.

JACOB, R. J. K. (1996). *Input Devices and Techniques*. The Computer Science and Engineering Handbook. A. B. Tucker, CRC Press**:** 1494-1511.

JACOBS, J. (1988). *Social Implications of Computers: Ethical and Equity Issues*. ACM SIGCUE Outlook 20(1).

JACOBSON, D. (1996). *Talking Tactile Maps and Environmental Audio Beacons: An Orientation and Mobility Development Tool for Visually Impaired People*. ICA Commission on Maps and Graphics for Blind and Visually Impaired People.

JAMESON, A., B. GROSSMANN-HUTTER, L. MARCH, R. RUMMER, T. BOHNENBERGER and F. WITTIG (2001). *When Actions Have Consequences: Empirically Based Decision Making for Intelligent User Interfaces*. Knowledge-Based Systems 14: 75-92.

JDR (2006). Journal of Design Research website, available at: http://www.inderscience.com/index.php. Last accessed August 2, 2006.

JOHNSON, M. (2001). *Transcutaneous Electrical Nerve Stimulation*. Electrotherapy: Evidence based practice. S. Kitchen, Churchill Livingstone**:** 259-286.

JOHNSON, M., A. WILSON, C. KLINE, B. BLUMBERG and A. BOBICK (1999). *Sympathetic Interfaces: Using a Plush Toy to Direct Synthetic Characters*. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '99).

JONES, L., M. NAKAMURE and B. LOCKYER (2004). *Development of a Tactile Vest*. HAPTICS'04, Chicago, USA.

JORGENSEN, C., K. WHEELER and S. STEPNIEWSKI (2000). *Bioelectric Control of a 757 Class High Fidelity Aircraft Simulation*. Proceedings of the World Automation Conference.

JUST, C., A. BIERBAUM, A. BAKER and C. CRUZ-NEIRA (1998). *VR Juggler: A Framework for Virtual Reality Development*. 2nd Immersive Projection Technology Workshop (IPT98).

KABBASH, P., B. BUXTON and A. SELLEN (1994). *Two-Handed Input in a Compound Task*. Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems, ACM Press.

KACZMAREK, K., J. WEBER, P. RITA and W. TOMPKINS (1991). *Electrotactile and vibrotactile displays for sensory substitution systems*. IEEE Transactions on Biomechanical Engineering 38(1).

KAJIMOTO, H., N. KAWAKAMI, T. MAEDA and S. TACHI (1999). *Tactile Feeling Display Using Functional Electrical Stimulation*. Proceedings of the 9th International Conference on Artificial Reality and Telexistence.

KANDEL, E., J. SCHWARTZ and T. JESSELL (2000). *Principles of Neural Science*, McGraw-Hill Education.

KATO, S., Y. YAMAKI, S. TAKEDA, S. ONO, M. IKEMOTO, N. TAKAGI and T. MUSYA (2002). *Application of the Eye Movement to the Life Support System.* Technology And Persons With Disabilities Conference 2002.

KAYE, J. (1999). *Symbolic Olfactory Display*. Brain & cognitive science department. Boston, Massachusetts Institute of Technology.

KAYE, J. (2004). *Making Scents: Aromatic Output for HCI*. ACM Interactions(January / February issue).

KEEFE, D., D. ACEVEDO, T. MOSCOVICH, D. LAIDLAW, AND J. LAVIOLA (2001). *CavePainting: A Fully Immersive 3D Artistic Medium and Interactive Experience*. Proceedings of the 2001 Symposium on Interactive 3D Graphics (I3D 2001), ACM Press.

KIERAS, D. (1997). *Task Analysis and the Design of Functionality*. The Computer Science and Engineering Handbook**:** 1401-1423.

KIRSCH, D. (1999). *The Affective Tigger: A Study on the Construction of an Emotionally Reactive Toy*. School of Architecture and Planning, Massachusetts Institute of Technology.

KITAMURA, Y., Y. ITOH and F. KISHINO (2001). *Real-time 3D Interaction with the ActiveCube*. Proceedings of the 2001 Conference on Human Factors in Computing Systems (CHI'2001).

KLATZKY, R., J. LOOMIS, A. BEALL, S. CHANCE and R. GOLLEDGE (1998). *Spatial Updating of Self-Position and Orientation During Real, Imagined and Virtual Locomotion*. Psychological Science 9: 29-298.

KONTRARINIS, D. and R. HOWE (1995). *Tactile display of vibrotactile information in teleoperation and virtual environments*. Presence: Teleoperators and Virtual Environments 4(4): 387-402.

KORITZINSKY, I. (1989). *New Ways to Consistent Interfaces*. Coordinating User Interfaces for Consistency. J. Nielsen, Academic Press.

KRAMER, J. (1991). *Communication System for Deaf, Deaf-Blind and Non-Vocal Individuals Using Instrumented Gloves, Patent No. 5,047,952.*

KREPKI, R., B. BLANKERTZ, G. CURIO and K. MUELLER (2003). *The Berlin Brain-Computer Interface (BBCI) - Towards a new Communication Channel for Online Control of Multimedia Applications and Computer Games*. Proceedings of the 9th International Conference on Distributed Multimedia Systems (DMS'03).

KRUEGER, M., T. GIONFRIDDO and K. HINRICHSEN (1985). *VIDEOPLACE - An Artificial Reality*. Proceedings of the 1985 ACM Conference on Human Factors in Computing Systems (CHI'85), ACM Press.

KRUGER, W. A. B. F. (1994). *The Responsive Workbench*. IEEE Computer Graphics & Applications 14(3): 12-15.

KRUIJFF, E., S. CONRAD and A. MUELLER (2003). *Flow of Action in Mixed Interaction Modalities*. Proceedings of HCI International.

KRUIJFF, E., S. CONRAD, P. PALAMIDESE, P. MAZZOLENI, F. HASENBRINK, M. SUTTROP and Y.-M. KWON (2004). *Remote Virtual Guidance in Immersive Museum Applications*. Proceedings of the 2004 Conference on Virtual Systems and Multimedia (VSMM 2004).

KRUIJFF, E. and A. PANDER (2005). *Experiences of using Shockwaves for Haptic Sensations*. Proceedings of 3D user interface workshop, IEEE Virtual Reality Conference (VR2005).

KRUIJFF, E., D. SCHMALSTIEG and S. BECKHAUS (2006). *Using Neuromuscular Electrical Stimulation for Pseudo-Haptic Feedback*. Proceedings of the ACM Symposium on Virtual Reality Software & Technology 2006 (VRST 2006).

KRUIJFF, E., G. WESCHE, K. RIEGE, G. GOEBBELS, M. KUNSTMAN and D. SCHMALSTIEG (2006). *Tactylus, a Pen-Input Device exploring Audiotactile Sensory Binding*. Proceedings of the ACM Symposium on Virtual Reality Software & Technology 2006 (VRST 2006).

KURZWEIL (2006). Kurzweil AI net, available at: http://www.kurzweilai.net. Last accessed February 10, 2006.

LALOR, E., S. KELLY, C. FINUCANE, R. BURKE, R. REILLY and G. MCDARBY (2004). *Brain Computer Interface based on the Steady-State VEP for Immersive Gaming Control*. Proceedings of the 2nd International Brain-Computer Interface Workshop and Training Course.

LAMBERTI, F., B. MONTRUCCHIO, A. SANNA and C. ZUNINO (2003). *A Web-based Architecture Enabling Multichannel Telemedicine Applications*. Journal of Systemics, Cybernetics and Informatics 1(1).

LATOSCHIK, M. (2001). *A General Framework for Multimodal Interaction in Virtual Reality Systems: PrOSA.* Proceedings of the IEEE Virtual Reality 2001, Yokohama, Japan.

LAU, C., S. CHURCHILL, J. KIM, F. MATSEN and Y. KIM (2002). *Asynchronous Web-Based Patient Centered Home Telemedicine System*. IEEE Transactions on Biomedical Engineering 49(12).

LAVIOLA, J. (2000a). *A Discussion of Cybersickness in Virtual Environments*. SIGCHI Bulletin 32(1): 47-56.

LAVIOLA, J., D. KEEFE, D. ACEVEDO, AND R. ZELEZNIK (2004). *Case Studies in Building Custom Input Devices for Virtual Environment Interfaces*. Proceedings of the VR 2004 Workshop on Beyond Wand and Glove Interaction.

LEDERMAN, S., A. MARTIN, C. TONG and R. KLATZKY (2004). *Relative Performance using Haptic and/or Touch-Produced Auditory Cues in a Remote Absolute Texture Identification Task*. Proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (IEEE HAPTICS'03).

LEDERMAN, S., G. THORNE and B. JONES (1986). *Perception of Texture by Vision and Touch: Multidimensionality and Intersensory Integration.* Journal of Experimental Psychology: Human Perception & Performance 12(2): 169-180.

LEEB, R., R. SCHERER, F. LEE, H. BISCHOF and G. PFURTSCHELLER (2004). *Navigation in Virtual Environments through Motor Imagery*. 9th Computer Vision Winter Workshop (CVWW'04).

LEGANCHUK, A., S. ZHAI and B. BUXTON (1999). *Manual and Cognitive Benefits of Two-Handed Input: An Experimental Study*. Transactions on Computer-Human Interaction 5(4): 326-359.

LEGO (2005). LEGO Mindstorms website, available at:http://mindstorms.lego.com. Last accessed September 5, 2005.

LEIBE, B., T. STARNER, W. RIBARSKY, Z. WARTELL, D. KRUM, J. WEEKS, B. SINGLETARY and L. HODGES (2000). *Toward Spontaneous Interaction with the Perceptive Workbench*. IEEE Computer Graphics & Applications 20(6): 54-65.

LEIKAS, J., H. STRÖMBERG, A. VÄÄTÄNEN and L. CLUITMANS (2004). *An Intuitive Game in an Intelligent Ubiquitous Environment*. ERCIM News 57.

LENAY, C., O. GAPENNE, S. HANNETON, C. GENOUËLLE and C. MARQUE (2003). *Sensory Substitution: Limits and Perspectives*. Touching for Knowing. Y. Hatwell, A. Streri and E. Gentaz**:** 275-292.

LIANG, J. (1994). *JDCAD: A Highly Interactive 3D Modeling System*. Computers and Graphics 18(4): 499-506.

LILEG, E., G. WIESSPEINER and H. HUTTEN (1999). *Evaluation of the EOG for Communication through Eye Movements*. Proceedings of the 10th Conference on Eye Movements.

LINDEMAN, R., J. SIBERT, C. LATHAN and J. VICE (2004). *The Design and Deployment of a Wearable Vibrotactile Feedback System*. Proceedings of the 8th IEEE International Symposium on Wearable Computers.

LISTEN (2005). Listen website, available at: http://listen.imk.fraunhofer.de/index2.html. Last accessed August 10, 2005.

LOOMIS, J. (2003). *Sensory Replacement and Sensory Substitution: Overview and Prospects for the Future*. Converging technologies for improving human performance. M. Roco and W. Bainbridge, Kluwer Academic Publishers.

LUCZAK, H. (1997). *Task Analysis*. Handbook of Human Factors and Ergonomics. G. Salvendy, John Wiley & Sons.

LUMELSKY, V., M. SHUR and S. WAGNER (2001). *Sensitive Skin*. IEEE Sensors Journal 1(1): 41-51.

LUSTED, H. and R. KNAPP (1996). *Controlling Computers with Neural Signals*. Scientific American 275(4): 82-87.

LYNCH, K. (1960). *The Image of the City*. Cambridge, MA, MIT Press.

LYONS, M., M. HAEHNEL and N. TETSUTANI (2001). *The Mouthesizer: a Facial Gesture Musical Interface*. Proceedings of ACM SIGGRAPH 2001.

MADCATZ (2006). MadCatz Bioforce article at Newsfactor, available at: http://www.newsfactor.com/perl/story/12528.html. Last accessed March 6, 2006.

MAEDA, T., H. ANDO and M. SUGIMOTO (2005). *Vection by Galvanic Vestibular Stimulation in a Virtual Reality Environment*. Proceedings of the IEEE Virtual Reality 2005.

MAES, P., T. DARELL, B. BLUMBERG and A. PENTLAND (1996). *The ALIVE System, Wireless, Full-Body Interaction with Autonomous Agents*. ACM Multimedia Systems(Special Issue on Multimedia and Multisensory Virtual Worlds).

MAKINWA, K. and J. HUISING (2001). *A Smart Wind-Sensor Based on Thermal Sigma-Delta Modulation*. Proceedings of TRANSDUCERS '01.

MANABE, H., A. HIRAIWA and T. SUGIMURA (2003). *A Ring-Shaped EMG Measurement System for Applying to User Interface*. Proceedings of IEEE EMBS, Cancun, Mexico.

MANN, S. (1996). *Smart Clothing: The Shift to Wearable Computing*. Communications of the ACM(August 1996).

MAPES, D. and J. MOSHELL (1995). *A Two-Handed Interface for Object Manipulation in Virtual Environments*. Presence: Teleoperators and Virtual Environments 4(4): 403-416.

MARRAS, W. (1997). *Biomechanics of the Human Body*. Handbook of Human Factors and Ergonomics. G. Salvendy, John Wiley & Sons.

MASLIAH, M. and P. MILGRAM (2000). *Measuring the Allocation of Control in a 6-Degree-of-Freedom Docking experiment*. Proceedings of the 2000 ACM Conference on Human Factors in Computing Systems (CHI'2000).

MASSIE, T. H. (1993). *Design of a Thee Degree of Freedom Force Reflecting Haptic Interface*, Massachusetts Institute of Technology.

MASSIMINO, M. and T. SHERIDAN (1993). *Sensory Substitution for Force Feedback in Teleoperation*. Presence: Teleoperators and Virtual Environments 2(4).

MAYNES-AMINZADE, D. (2005). *Edible Bits: Seamless Interfaces between People, Data and Food*. Proceedings of the 2005 ACM Conference on Human Factors in Computing Systems (CHI'2005).

MAYNES-AMINZADE, D. and H. RAFFLE (2003). *You're In Control: A Urinary User Interface*. Proceedings of the 2003 ACM Conference on Human Factors in Computing Systems (CHI'2003).

MCCAIG, G. and S. FELS (2002). *Playing on Heart-Strings: Experiences with the 2Hearts System*. Proceedings of the 2nd International Conference on New Interfaces for Musical Expression (NIME02).

MCELLIGOTT, L., M. DILLON, K. LEYDON, B. RICHARDSON, M. FERNSROM and J. PARADISO (2002). *ForSe FIElds - Force Sensors for Interactive Environments*. Proceedings of UbiComp 2002, Springer-Verlag Berlin Heidelberg.

MCMILLAN, G., R. EGGELSTON and T. ANDERSON (1997). *Nonconventional Controls*. Handbook of Human Factors and Ergonomics. G. Salvendy. New York, John Wiley and Sons: 729-771.

MCTEAR, M. (2002). *Spoken Dialogue Technology: Enabling the Conversational User Interface.* ACM Computing Surveys 34(1): 90-169.

MENASHE, I., O. MAN, D. LANCET and Y. GILAD (2003). *Different noses for different people*. Nature Genetics 34(2).

MERL (2004). Haptic Stylus project summary, available at: http://www.merl.com/projects/hapticstylus/. Last accessed June 1, 2005.

MERRILL, D. (2004). *FlexiGesture: A Sensor-Rich Real-Time Adaptive Gesture and Affordance Learning Platform for Electronic Music Control*. Media Arts and Sciences, Massachusetts Institute of Technology.

METROVISION (2005). Metrovision website, available at: http://www.metrovision.fr. Last accessed October 10, 2005.

MICROVISION (2005). Microvision website, available at: http://www.microvision.com. Last accessed July 10, 2005.

MIGNONNEAU, L. and C. SOMMERER (2005). *Nano-Scape: Experiencing Aspects of Nanotechnology through a Magnetic Force-Feedback Interface*. Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology.

MILGRAM, P. and F. KISHINO (1994). *A Taxonomy of Mixed Reality Visual Displays*. IECE Transactions on Information and Systems E77-D(12): 1321-1329.

MINDPEAK (2005). MindPeak website, available at: http://www.mindpeak.com. Last accessed October 11, 2005.

MINE, M. (1995). *ISAAC: A Virtual Environment Tool for the Interactive Construction of Virtual Worlds*, Dept. of Computer Science, University of North Carolina at Chapel Hill,TR-95-020.

MINE, M. (1995a). *Virtual Environment Interaction Techniques*, Dept. of Computer Science, University of North Carolina at Chapel Hill,TR95-018.

MINE, M., F. BROOKS and C. SEQUIN (1997a). *Moving Objects in Space: Exploiting Proprioception in Virtual Environment Interaction*. Proceedings of SIGGRAPH'97, ACM Press.

MITOXYGEN (2005). MIT Oxygen project website, available at: http://oxygen.lcs.mit.edu/Overview.html. Last accessed October 21, 2005.

MOHSANI, P., K. NAJAFI, S. ELIADES and X. WANG (2005). *Wireless Multichannel Biopotential Recording Using an Integrated FM Telemetry Circuit*. IEEE Transactions on Neural Systems and Rehabilitation Engineering 13(3).

MORGAN, R. and D. HEISE (1988). *Structure of Emotions*. Social Psychology Quarterly 51(1): 19-31.

MORIE, J., K. IYER, K. VALANEJAD, R. SADEK, D. MIRAGLIA, D. MILAM, D. LUIGI, J. LESHIN and J. WILLIAMS (2003). *Sensory Design for Virtual Environments*. Proceedings of ACM SIGGRAPH 2003.

MORRIS, S. and J. PARADISO (2002). *Shoe-integrated Sensor System for Wireless Gait Analysis and Real-time Feedback*. IEEE Second joint EMBS/BMES conference.

MOTIONANALYSIS (2005). MotionAnalysis website, available at: http://www.motionanalysis.com/. Last accessed September 28, 2005.

MUELLER, A., S. CONRAD and E. KRUIJFF (2003). *Multifaceted Interaction with a Virtual Engineering Environment using a Scenegraph-oriented Approach*. WSCG, Plzen.

MUELLER, F. (2002). *Exertion Interfaces: Sports over a Distance for Social Bonding and Fun*. Media Arts and Sciences, Massachusetts Institute of Technology.

MURAYAMA, J., L. BOUGRILA, L. LUO, K. AKAHANE, S. HASEGAWA, B. BHIRSBRUNNER and M. SATO (2004). *SPIDAR G&G: A two handed haptic interface for bimanual VR interaction*. Proceedings of EuroHaptics.

NASA (2005). TSAS website, available at: http://www.ihmc.us/research/projects/TSAS/. Last accessed September 5, 2005.

NEWMAN, J., J. ROSENBACH, K. BURNS, B. LATIMER, H. MATOCHA and E. VOGT (1995). *An Experimental Test of "the Mozart Effect": does Listening to his Music improve Spatial Ability?* Perception and Motor Skills 81: 1379-87.

NIELSEN, J. (1993). *Usability Engineering*. Boston, Academic Press.

NIELSEN, J. and R. MOLICH (1992). *Heuristic Evaluation of User Interfaces*. Proceedings of the 1992 ACM Conference on Human Factors in Computing Systems (CHI'92), ACM Press.

NINJOUJI, T., H. MANABE, T. MASHIKO, S. SUZUKI and T. SUGIMURA (2003). *Biological Information Interface Technology*. NTT Technical Review Online, available at: http://www.ntt.co.jp/tr/0311/special.html 1(8).

NINTENDO (2005). Nintendo Donkey Kong Jungle Beat website, available at: http://www.nintendo.com/gamemini?gameid=4a6a4f98-6bcb-414a-b2cf-f8e112c704b9. Last accessed September 16, 2005.

NOLAN (2005). Mindset (Nolan Computer Systems) website, available at: http://www.mindset-eeg.com/. Last accessed October 11, 2005.

NORMAN, D. (1990). *The Design of Everyday Things*. New York, Doubleday.

NTT (2004). NTT Docomo Fingerwhisper website, available at: http://www.nttdocomo.com/corebix/ubiquity/fingerwhisper.html. Last accessed June 15, 2004.

ODGAARD, E., Y. ARIEH and L. MARKS (2004). *Brighter Noise: Sensory Enhancement of Perceived Loudness by Concurrent Visual Stimulation*. Cognitive, Effective, & Behavioral Neuroscience 4(2).

OKAMURA, A., J. DENNERLEIN and R. HOWE (1998). *Vibration feedback models for virtual environments*. IEEE International conference on Robotics and Automation, Leuven, Belgium.

OPENEEG (2005). OpenEEG list of biopotential technology and software, available at: http://openeeg.sourceforge.net/doc/links-biopsy.html. Last accessed October 14, 2005.

OVIATT, S. and P. COHEN (2000). *Multimodal Interfaces that Process What Comes Naturally*. Communications of the ACM 43(3): 45-51.

PAI, D. (2003). *Multisensory Interaction: Real and Virtual*. Proceedings of the International Symposium on Robotics Research.

PAINSTATION (2005). PainStation website, available at http://www.painstation.de. Last accessed August 1, 2005.

PAIVA, A., R. PRADA, R. CHAVES, M. VALA, A. BULLOCK, G. ANDERSSON and K. HÖÖK (2003). *Towards Tangibility in Gameplay: Building a Tangible Affective Interface for a Computer Game*. Proceedings of the International Conference on Multimodal interfaces (ICMI'03).

PALANKER, D. (2006). Retinal prosthesis developments at Stanford, available at: http://www.stanford.edu/~palanker/lab/retinalpros.html. Last accessed July 11, 2006.

PALEY, W. (1998). *Designing Special-Purpose Input Devices*. Computer Graphics(November 1998).

PANTIC, M. and L. ROTHKRANTZ (2000). *Automatic Analysis of Facial Expressions: The State of the Art*. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12).

PARK, C., H. KO, J. KIM, S. AHN, Y.-M. KWON, H. KIM and T. KIM (2003). *Gyeongju VR Theater: 'A Journey into the Breadth of Seorabol'*. Presence: Teleoperatore and Virtual Environments 12(2).

PATTICHIS, C., E. KYRIACOU, S. VOSKARIDES, M. PATTICHIS, R. ISTEPANIAN and C. SHIZAS (2002). *Wireless Telemedicine Systems: An Overview*. IEEE Antennas & Propagation Magazine 44(2): 143-153.

PAUSCH, R., J. SNODDY, R. TAYLOR, S. WATSON and E. HASELTINE (1996). *Disney's Aladdin: First Steps toward Storytelling in Virtual Reality*. Proceedings of SIGGRAPH'96, ACM Press.

PEDALPANNER (2005). Padel Panner website, available at: http://hct.ece.ubc.ca/research/pedal/index.html. Last accessed September 30, 2005.

PENFIELD, W. and T. RASMUSSEN (1950). *The Cerebral Cortex of man - a Clinical Study of Localization of Function*. New York, The Macmillan Comp.

PERLIN, K., S. PAXIA, AND J. KOLLIN (2000). *An Autostereoscopic Display*. Proceedings of SIGGRAPH 2000, ACM Press.

PICARD, R. (1997). *Affective Computing*, MITPress.

PICARD, R. and J. SCHEIRER (2001). *The Galvactivator: A Glove that Senses and Communicates Skin Conductivity*. Proceedings of the 9th International Conference on Human Computer Interaction.

PIERCE, J., A. FORSBERG, M. CONWAY, S. HONG, R. ZELEZNIK and M. MINE (1997). *Image Plane Interaction Techniques in 3D Immersive Environments*. Proceedings of the1997 ACM Symposium on Interactive 3D Graphics (I3D'97), ACM Press.

PIERCE, J., B. STEARNS and R. PAUSCH (1999). *Two Handed Manipulation of Voodoo Dolls in Virtual Environments*. Proceedings of the 1999 Symposium on Interactive 3D Graphics (I3D'99).

POLSON, P., C. LEWIS, J. RIEMAN and C. WHARTON (1992). *Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces*. International Journal of Man-Machine Studies 36: 741-773.

PORCARI, J., J. MILLER, K. CORNWELL, C. FOSTER, M. GIBSON, K. MCLEAN and T. KERNOZEK (2005). *The effects of neuromuscular electrical stimulation training on abdominal strength, endurance, and selected anthropometric measures*. Journal of Sports Science and Medicine 4: 66-75.

POUPYREV, I., M. BILLINGHURST, S. WEGHORST and T. ICHIKAWA (1996a). *The Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR*. Proceedings of the 1996 ACM Symposium on User Interface Software and Technology (UIST'96), ACM Press.

POUPYREV, I. and E. KRUIJFF (2000). *20Th Century 3DUI Bib: Annotated Bibliography of 3D User Interfaces of the 20th Century*.

POUPYREV, I., N. TOMOKAZU, AND S. WEGHORST. (1998a). *Virtual Notepad: Handwriting in Immersive VR*. Proceedings of the 1998 IEEE Virtual Reality Annual International Symposium (VRAIS'98), IEEE Press.

POUPYREV, I., S. WEGHORST, M. BILLINGHURST and T. ICHIKAWA (1998b). *Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques*. Computer Graphics Forum, EUROGRAPHICS'98 Issue 17(3): 41-52.

POWERLAB (2005). PowerLab / ADInstruments website, available at: http://www.adinstruments.com. Last accessed October 6, 2005.

PREECE, J. (2002). *Interaction design: beyond human-computer interaction*, John Wiley & Sons.

PRIDE (2005). Pride Mobility Products Corp. website, available at: http://www.pridemobility.com/quantum/. Last accessed September 12, 2005.

QUASAR (2005). QUASAR products website, available at: http://www.quasarusa.com/tp.html. Last accessed October 13, 2005.

RAFFLE, H., A. PARKES and H. ISHII (2004). *Topobo: A Constructive Assembly System with Kinetic Memory*. Proceedings of the 2004 Conference on Human Factors in Computing Systems (CHI'2004).

RASK, J., R. GONZALEZ and T. BUCHANAN (1999). *Servo-Motor Control of Human Arm Kinematics in Virtual Reality Modeling*. Proceedings of the 1999 Bioengineering Conference (BIO'99).

RASKIN, J. (2000). *The Humane Interface: New Directions for Designing Interactive Systems*, ACM Press.

RAUSCHER, F. and G. SHAW (1993). *Music and Spatial Task Performance*. Nature 365: 611.

REGEN (2005). Regenerative Brainwave Music website, available at http://www.eyetap.org/about_us/people/fungja/regen.html. Last accessed October 14, 2005.

REKIMOTO, J. (2002). *SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces*. Proceedings of the 2002 ACM Conference on Human Factors in Computing Systems (CHI'2002).

REKIMOTO, J. and N. MATSUSHITA (1997). *Perceptual Surfaces: Towards a Human and Object Sensitive Interactive Display*. Proceedings of the 1997 ACM Workshop on Perceptual User Interfaces (PUI'97).

REKIMOTO, J. and H. WANG (2004). *Sensing GamePad: Electrostatic Potential Sensing for Enhancing Entertainment Oriented Interactions*. Proceedings of the 2004 ACM Conference on Human Factors in Computing Systems (CHI 2004).

RICHARDSON, D. and M. SPIVEY (in press (a)). *Eye-Tracking: Characteristics and methods*. Encyclopedia of Biomaterials and Biomedical Engineering. G. Bowlin, Marcel Dekker Inc.

RICHARDSON, D. and M. SPIVEY (in press (b)). *Eye-Tracking: Research Areas and applications*. Encyclopedia of Biomaterials and Biomedical Engineering. G. Wnek and G. Bowlin, Marcel Dekker Inc.

ROSENFELD, R., D. OLSEN and A. RUDNICKY (2001). *Universal Speech Interfaces*. Interactions VIII(6): 33-44.

RPP (2004). Retinal Prosthesis Project website, available at: http://www.icat.ncsu.edu/projects/retina/index.htm. Last accessed June 15, 2004.

SALEM, C. and S. ZHAI (1997). *An Isometric Tongue Pointing Device*. Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems (CHI'97).

SALVENDY, G. (1997). *Handbook of Human Factors and Ergonomics*, John Wiley & Sons.

SAWAN, M. (2004). *Wireless Smart Implants dedicated to Multichannel Monitoring and Microstimulation*. International Conference on Pervasive Services (ICPS'2004).

SCAPONE, G. (2003). *The Pipe: Explorations with Breath Control*. Proceedings of the 3rd International Conference on New Interfaces for Musical Expression (NIME'03).

SCHMALSTIEG, D., A. FUHRMANN, G. HESINA, Z. SZALAVARI, L. ENCARNACAO, M. GERVAUTZ and W. PURGATHOFER (2002). *The Studierstube Augmented Reality Project*. Presence: Teleoperators and Virtual Environments 11(1): 32-54.

SCHMALSTIEG, D., L. M. ENARNACAO, AND Z. SZALAVARI (1999). *Using Transparent Props For Interaction with The Virtual Table*. Proceedings of the 1999 ACM Symposium on Interactive 3D Graphics (I3D'99), ACM Press.

SCHMALSTIEG, D. and G. REITMAYR (2001). *An Open Software Architecture for Virtual Reality Interaction*. Proceedings of the 2001 ACM Symposium on Virtual Reality Software & Technology (VRST 2001).

SEKULER, R., A. SEKULER and R. LAU (1997). *Sound alters Visual Motion Perception*. Nature 385.

SELLEN, A., G. KURTENBACH and W. BUXTON (1992). *The Prevention of Mode Errors through Sensory Feedback*. Human Computer Interaction 7(2): 141-164.

SENSATEX (2005). Sensatex website, available at: http://www.sensatex.com. Last accessed October 11, 2005.

SERRUYA, M., N. HATSOPOULOS, L. PANINSKI, M. FELLOWS and J. DONOGHUE (2002). *Instant Neural Control of a Movement Signal*. Nature 416: 141-142.

SHAW, J. (2005). Jeffrey Shaw website, available at: http://www.jeffrey-shaw.net. Last accessed September 22, 2005.

SHEPHERD, G. (1994). *Neurobiology*, Oxford University Press.

SHERMAN, B. and A. CRAIG (2003). *Understanding Virtual Reality*, Morgan Kauffman Publishers.

SHIMOJO, S. and L. SHAMS (2001). *Sensory Modalities are not Separate Modalities: Plasticity and Interactions*. Current Opinion in Neurobiology 11: 505-509.

SHNEIDERMAN, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction, 3rd Edition*, Addison-Wesley.

SHNEIDERMAN, B. (2000). *Creating Creativity: User Interfaces for Supporting Innovation*. ACM Transaction on Computer-Human Interaction 7(1): 114-138.

SHNEIDERMAN, B. (2000). *The Limits of Speech Recognition*. Communications of the ACM 43(9): 63-65.

SIMON, A. (2005). *First-person Experience and Usability of Co-located Interaction in a Projection-based Virtual Environment*. Proceedings of the 2005 ACM Symposium on Virtual Reality Software and Technology (VRST 2005).

SLATER, M., M. USOH and A. STEED (1994). *Depth of Presence in Virtual Environments*. Presence: Teleoperators and Virtual Environments 3(2): 130-144.

SLATER, M., M. USOH and A. STEED (1995). *Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality*. ACM Transactions on Computer-Human Interaction 2(3): 201-219.

SMARTFINGER (2005). SmartFinger website, available at http://www.star.t.u-tokyo.ac.jp/projects/smartfinger/. Last accessed August 1, 2005.

SMITH, B., J. HO, W. ARK and S. ZHAI (2000). *Hand Eye Coordination Patterns in Target Selection*. Proceedings of the 2000 ACM Eye Tracking Research & Applications Symposium.

SMITH, J., T. WHITE, C. DODGE, D. ALLPORT, J. PARADISO and N. GERSHENFLED (1999). *Electric Field Sensing for Graphical Interfaces*. Computer Graphics and Applications 18(3): 54-61.

SMITH, T. and K. SMITH (1987). *Feedback-Control Mechanisms of Human Behavior*. Handbook of Human Factors. G. Salvendy, John Wiley and Sons**:** 251-293.

SOARES, L., L. NOMURA, M. CABRAL, L. DULLEY, M. GUIMARÃES, R. LOPES and M. ZUFFO (2004). *Virtual Hang-gliding over Rio de Janeiro*. Proceedings of the 2004 IEEE Virtual Reality Workshop “Virtual Reality for Public Consumption”.

SÖRNMO, L. and P. LAGUNA (2005). *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Elsevier Academic Press.

SPACESYNTAX (2005). Space Syntax website, available at: http://www.spacesyntax.net/. Last accessed September 23, 2005.

SPENCE, C. and S. SQUIRE (2003). *Multisensory Integration: Maintaining the Perception of Synchrony*. Current Biology 13: 519-521.

SRINIVASAN, P., D. BIRCHFIELD, G. QIAN and A. KIDANÉ (2005). *A Pressure Sensing Floor for Interactive Media Applications*. Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE 2005).

STANNEY, K., Ed. (2002). *Handbook of Virtual Environments*, Lawrence Erlbaum and Associates.

STANNEY, K., R. MOURANT and R. KENNEDY (1998). *Human Factors Issues in Virtual Environments: A Review of the Literature.* Presence: Teleoperators and Virtual Environments 7(4): 327-351.

STEFANI, O. and J. RAUSCHENBACH (2003). *3D Input Devices and Interaction Concepts for Optical Tracking in Immersive Environments*. Immersive Projection Technology and Virtual Environments (ACM IPT).

STEFFIN, M. (1998). *Virtual Reality Therapy of Multiple Sclerosis and Spinal Cord Injury: Design Considerations for a Haptic-Visual Interface*. Virtual Reality in Neuro-Psycho-Physiology. G. Riva. Amsterdam, Ios Press.

STEFFIN, M. (2005). *Virtual Reality Biofeedback in Chronic Pain and Psychiatry*. eMedicine Journal, available online at: http://www.emedicine.com/neuro/topic466.htm.

STEPHANIDIS, C., C. KARAGIANNIDIS and A. KOUMPIS (1997). *Decision Making in Intelligent User Interfaces*. Proceedings of the 1997 ACM International Conference on Intelligent User Interfaces (IUI'97).

STEUER, J. (1992). *Defining Virtual Reality: Dimensions Determining Telepresence*. Journal of Communication 42(4).

STREITZ, N., J. GEISSLER, T. HOLMER, S. KONOMI, C. MUELLER-TOMFELDE, W. REISCHL, P. REXROTH, P. SEITZ and R. STEINMETZ (1999). *i-LAND: An interactive Landscape for Creativity and Innovation*. ACM Conference on Human Factors in Computing Systems (CHI '99).

STURMAN, D., D. ZELTZER and S. PIEPER (1989). *Hands-On Interaction with Virtual Environments*. Proceedings of the 1989 ACM Symposium on User Interface Software and Technology (UIST'89), ACM Press.

SUI (2005). Strawlike user interface website, available at: http://www.hi.mce.uec.ac.jp/inami-lab/. Last accessed September 6, 2005.

SUN, Y., N. SEBE, M. LEW and T. GEVERS (2004). *Authentic Emotion Detection in Real-Time Video*. HCI/ECCV 2004.

SUTHERLAND, I. (1965). *The Ultimate Display*. Proceedings of the IFIP Congress.

SYKES, J. and S. BROWN (2003). *Affective Gaming - Measuring Emotion through the Gamepad*. Proceedings of the 2003 ACM Conference on Human Factors in Computing Systems (CHI 2003).

SYSTEM16 (2005). Taito Real Puncher at System16 Arcade Museum website, available at http://www.system16.com/taito/hrdw_unknownpost90.html. Last accessed September 20, 2005.

SZALAVARI, Z. and M. GERVAUTZ (1997). *The Personal Interaction Panel - a Two-Handed Interface for Augmented Reality*. Computer Graphics Forum 16(3): 335-346.

TAN, H. and A. PENTLAND (1997). *Tactual displays for wearable computing*. IEEE ISWC.

TAN, H., A. SLIVOVSKY and A. PENTLAND (2001). *A Sensing Chair Using Pressure Distribution Sensors*. IEEE/ASME Transactions on Mechatronics 6(3): 261-268.

TAYLOR, R., T. HUDSON, A. SEEGER, H. WEBER, J. JULIANO and A. HELSER (2001). *VRPN: A Device-Independent, Network-Transparent VR Peripheral System*. Proceedings of the 2001 ACM Symposium on Virtual Reality Software and Technology (VRST 2001).

TECHNOLOGYWATCH (2005). Technology Watch report from DFKI on Intelligent User Interface research and developments, available at: http://www.dfki.de/fluids/Intelligent_User_Interfaces.html. Last accessed October 17, 2005.

THERAFIN (2005). Therafin Sip 'n Puff website, available at: http://www.therafin.com/sipnpuff.htm. Last accessed September 12, 2005.

THERMOMETRICS (2005). Thermometrics website, available at: http://www.thermometrics.com. Last accessed October 10, 2005.

THORNDYKE, P. and B. HAYES-ROTH (1982). *Differences in Spatial Knowledge Obtained from Maps and Navigation*. Cognitive Psychology 14: 560-589.

THOUGHTTECHNOLOGY (2005). Thought Technology Ltd website, available at: http://www.thoughttechnology.com. Last accessed October 10, 2005.

TIDWELL, M., R. S. JOHNSTON, D. MELVILLE, AND T. A. FURNESS (1995). *The Virtual Retinal Display - A Retinal Scanning Imaging System*. Proceedings of Virtual Reality World'95.

TOLLMAR, K., D. DEMIRDJIAN and T. DARELL (2003). *Gesture + Play: Exploring Full Body Interaction for Virtual Environments*. Proceedings of the 2003 ACM Conference on Human Factors in Computing Systems (CHI 2003).

TOUMAZ (2005). Toumaz Sensium platform website, available at: http://www.toumaz.com. Last accessed July 17, 2005.

TRAMBEREND, H. (2001). *Avango: A Distributed Virtual Reality Framework*. Proceedings of ACM Afrigraph'01.

TREJO, L., K. WHEELER, C. JORGENSEN, R. ROSIPAL, S. CLANTON, B. MATTHEWS, A. HIBBS, R. MATTHEWS and M. KRUPKA (2002). *Multimodal Neuroelectric Interface Development*. IEEE Transactions on Neural Systems and Rehabilitation Engineering (Special Issue on BCI2002).

TUFTE, E. (1990). *Envisioning Information*. Cheshire, Graphics Press.

TURK, M. and G. ROBERTSON (2000). *Perceptual User Interfaces*. Communications of the ACM 42(2): 33-34.

TURNER, A. (1996). Biosensors: Past, Present and Future. Last accessed October 6, 2005.

ULLMER, B. and H. ISHII (1997). *The metaDesk: Models and Prototypes for Tangible User Interfaces*. Proceedings of the 1997 ACM Symposium on User Interface Software and Technology (UIST'97), ACM Press.

USOH, M., ARTHUR, K., WHITTON, M.C., BASTOS, R., STEED, A., SLATER, M., BROOKS, F.P. JR (1999). *Walking > Walking-in-Place > Flying in Virtual Environments*. Proceedings of SIGGRAPH '99, ACM.

USPTO (2005). Sony patents 6,729,337 and 6,536,440 in US patent database, available at : http://www.uspto.gov/patft/index.html. Last accessed September 6, 2005.

VÄLKKYNEN, P., J. HEINILÄ, S. LAINIO, S. LAKANIEMI and A. VÄÄTÄNEN (2001). *Using Exercise Cycle as a Haptic Input Device in a Virtual Environment*. Proceedings of the Joint IPT/EGVE ’01 Workshop Fifth Immersive Projection Technology Workshop together with the Seventh Eurographics Workshop on Virtual Environments.

VERAART, C., C. RAFTOPOULOS, J. MORTIMER, J. DELBEKE, D. PINS, G. MICHAUX, A. VANLIERDE, S. PARRINI and M. WANET-DEFALQUE (1998). *Visual Sensations produced by Optic Nerve Stimulation using an Implanted Self-Sizing Spiral Cuff Electrode*. Brain Research 813: 181-186.

VIEGA, J., M. CONWAY, G. WILLIAMS and R. PAUSCH (1996). *3D Magic Lenses*. Proceedings of the 1996 ACM Conference on User Interface Software and Technology (UIST'96).

VIRTUSPHERE (2005). VirtuSphere products website, available at: http://www.virtusphere.net/product.htm. Last accessed September 23, 2005.

VO-DINH, T. (2004). *Biosensors, Nanonsensors and Biochips: Frontiers in Environmental and Medical Diagnostics*. 1st International Symposium on Micro & Nano Technology.

VOGT, F., G. MCCAIG, A. ALI and S. FELS (2002). *Tongue 'n' Groove*. Proceedings of the 2nd International Conference on New Interfaces for Musical Expression (NIME 2002).

WALL, C., D. MERFELD, S. RAUCH and F. BLACK (2002). *Vestibular Prostheses: the Engineering and Biomedical Issues*. Journal of Vestibular Research 12: 95-113.

WANG, H., H. PRENDINGER and T. IGARASHI (2004). *Communicating Emotions in Online Chat Using Physiological Sensors and Animated Text*. ACM Conference on Human Factors in Computing Systems (CHI'04).

WARE, C. and D. JESSOME (1988). *Using the Bat: a Six-Dimensional Mouse for Object Placement*. IEEE Computer Graphics&Applications 8(6): 65-70.

WARE, C. and S. OSBORNE (1990). *Exploration and Virtual Camera Control in Virtual Three Dimensional Environments*. Proceedings of the 1990 ACM Symposium on Interactive 3D Graphics (I3D'90), ACM Press.

WARWICK, K., M. GASSON, B. HUTT, I. GOODHEW, P. KYBERD, B. ANDREWS, P. TEDDY and A. SHAD (2003). *The Application of Implant Technology for Cybernetic Systems*. Archives of Neurology 60(10): 1369-1373.

WASHBURNE, D. and L. JONES (2004). *Could Olfactory Displays Improve Data Visualisation?* IEEE Computing in Science & Engineering(November / December).

WATSEN, K., R. DARKEN, AND M. CAPPS (1999). *A Handheld Computer as an Interaction Device to a Virtual Environment*. Proceedings of the Third Immersive Projection Technology Workshop, Stuttgart, Germany.

WATT, A. and M. WATT (1992). *Advanced Animation and Rendering Techniques: Theory and Practice*. New York, ACM Press.

WEARABLEMOTHERBOARD (2005). Georgia Tech Wearable Motherboard website, available at: http://www.gtwm.gatech.edu/index/nutshell.html. Last accessed October 11, 2005.

WEISENBERGER, J. and G. POLING (2004). *Multisensory Roughness Perception of Virtual Surfaces: Effects of Correlated Cues*. Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environments and Teleoperators Systems (HAPTICS'04).

WEISER, M. (1991). *The Computer for the 21st Century*. Scientific American 265(3): 66-75.

WELCH, G., G. BISHOP, L. VICCI, S. BRUMBACK, K. KELLER and D. COLUCCI (1999). *The HiBall Tracker: High-Performance Wide-Area Tracking for Virtual and Augmented Environments*. Proceedings of the 1999 ACM Symposium on Virtual Reality Software and Technology (VRST'99), ACM Press.

WHEELER, K. (2003). *Device Control Using Gestures Senses From EMG*. Proceedings of the 2003 IEEE International Conference on Soft Computing in Industrial Applications.

WHEELER, K. and C. JORGENSEN (2003). *Gestures as Input: Neuroelectric Joysticks and Keyboards*. IEEE Pervasive Computing 2(2).

WICAB (2005). BrainPort Technologies website, available at: http://www.wicab.com. Last accessed September 7, 2005.

WIENER, N. (1948). *Cybernetics, or Control and Communication in the Animal and the Machine*. New York, Wiley.

WILDDIVINE (2005). Wild Divine website, available at: http://www.wilddivine.com. Last accessed October 14, 2005.

WILDER, J., G.K. HUNG, M.M. TREMAINE, AND M. KAUR (2002). *Eye Tracking in Virtual Environments*. Handbook of Virtual Environments. K. Stanney, Lawrence Erlbaum and Associates**:** 211-222.

WILSON, G., J. LAMBERT and C. RUSSELL (1999). *Performance enhancement with real-time physiologically controlled adaptive aiding*. Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting.

WISE, K., D. ANDERSON, J. HETKE, D. KIPKE and K. NAJAFI (2004). *Wireless implantable microsystems: high-density electronic interfaces to the nervous system*. Proceedings of the IEEE 92(1): 76-97.

WLOKA, M. and E. GREENFIELD (1995). *The Virtual Tricorder: A Uniform Interface for Virtual Reality*. Proceedings of the 1995 ACM Symposium on User Interface Software and Technology (UIST'95), ACM Press.

WOLPAW, J., N. BIRBAUMER, D. MCFARLAND, G. PFURTSCHELLER and T. VAUGHAN (2002). *Brain-Computer Interfaces for Communication and Control*. Neurophysiology 113: 767-791.

WOOD, S. (2004). *Automated Behavior-Based Interaction Customization for Military Command and Control*. Proceedings of 2004 ACM Conference on Intelligent User Interfaces 2004 (IUI 2004).

WREN, C., S. BASU, F. SPARACINO and A. PENTLAND (1999). *Combining Audio and Video in Perceptive Spaces*. 1st International Workshop on Managing Interactions in Smart Environments, Dublin, Ireland.

YANAGIDA, Y., H. KAWATO, H. NOMA, A. TOMONO and N. TETSUTANI (2003). *A nose-tracked, personal olfactory display*. ACM SIGGRAPH.

YANG, U., Y. JANG and G. KIM (2002). *Designing a Vibro-Tactile Wear for "Close-Range" Interaction for VR-based Motion Training*. Proceedings of ICAT2002.

YIN, K. and D. PAI (2003). *FootSee: an Interactive Animation System*. Eurographics/SIGGRAPH Symposium on Computer Animation.

ZELEZNIK, R., J. LAVIOLA, D. ACEVEDO, AND D. KEEFE (2002). *Pop-Through Buttons for Virtual Environment Navigation and Interaction*. Proceedings of IEEE Virtual Reality 2002, IEEE Press.

ZHAI, S. (1995). *Human Performance in Six Degree of Freedom Input Control*. PhD Dissertation, Dept. of Computer Science, University of Toronto.

ZHAI, S. (1998a). *User Performance in Relation to 3D Input Device Design*. Computer Graphics 32(4): 50-54.

ZHAI, S., P. MILGRAM and W. BUXTON (1996). *The Influence of Muscle Groups on Performance of Multiple Degree-Of-Freedom Input*. Proceedings of the 1996 ACM Conference on Human Factors in Computing Systems (CHI'96), ACM Press.

ZHAO, W., R. CHELLAPPA, P. PHILLIPS and A. ROSENFELD (2003). *Face Recognition: A Literature Survey*. ACM Computing Surveys 35(4): 399-458.

ZIMMERMAN, T. (1996). *Personal Area Networks: Near-Field Intrabody Communication*. IBM Systems Journal 35(3 & 4).

ZIMMERMAN, T., J. LANIER, C. BLANCHARD, S. BRYSON, AND Y. HARVILL (1987). *A Hand Gesture Interface Device*. Proceedings of CHI+GI'87, Human Factors in Computing Systems and Graphics Interface, ACM Press.