TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Informatikanwendungen in der Medizin & Augmented Reality

Immersive, Interactive and Contextual In-Situ Visualization for Medical Applications

Christoph P. Bichlmeier

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Bernd Brügge, Ph.D.
Prüfer der Dissertation:		
	1.	Univ.-Prof. Dr. Nassir Navab
	2.	apl. Prof. Dr. Ekkehard Euler, Ludwig-Maximilians-Universität München
	3.	Prof. Henry Fuchs, Ph.D., University of North Carolina, USA

Die Dissertation wurde am 18.05.2010 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 02.12.2010 angenommen.

Abstract

English:
The present work addresses the Human Computer Interface Design (HCID) of Medical Augmented Reality (AR) systems, allowing for interactive in-situ visualization. In particular, this thesis aims at two major challenges to adapt a 3D user interface (UI) to the needs of its medical target applications. These challenges are the seamless composition of virtual and real objects and the interaction with them in an AR environment. The proposed solutions are capable of enhancing the user's immersion into the AR scene in order to take full advantage of this technology.

Deutsch:
Die vorliegende Arbeit befasst sich mit der Gestaltung der Mensch-Computer-Schnittstelle bei medizinischen Augmented Reality (AR)-Systemen, die die interaktive in-situ Visualisierung ermöglichen. Um eine AR-basierte 3D-Benutzerschnittstelle an die Anforderungen ihres zugedachten medizinischen Aufgabenfeldes anzupassen wurden zwei Aufgabenstellungen bearbeitet. Zum einen wurden Methoden zur Komposition von realen und virtuellen Inhalten der AR-Szene entwickelt. Zum anderen wurde eine Reihe von Benutzerschnittstellen zur Interaktion mit diesen Inhalten realisiert und evaluiert. Die vorgeschlagenen Lösungen ermöglichen es dem Benutzer nahtloser in die AR-Welt einzutauchen und so deren Potentiale für die medizinische Anwendung der Technologie besser zu nutzen.

Acknowledgements:

I would like to thank my family Franz, Katharina and Kathrin Bichlmeier and Julia Dünzl for their precious support during my studies and my PHD. Second, I want to express my gratitude to Prof. Dr. Nassir Navab and my colleagues and former fellows during my PHD for the best time ever, for precious help and discussion and the great working atmosphere. In particular, I'd like to mention Dr. Tobias Sielhorst, who initiated my passion for Medical Augmented Reality followed by the original core team at NARVIS lab: Prof. Dr. med. Ekkehard Euler, MD Sandro Michael Heining, Dr. Jörg Traub, Dr. Marco Feuerstein, Philipp Stefan and Stefan Wiesner. Thanks also to the more recent NARVIS team consisting of Tassilo Klein, Seyed-Ahmad Ahmadi, Stuart Holdstock, Tobias Blum, Lejing Wang and Oliver Kutter. I wish to thank Prof. Dr. med. Wolf Mutschler for providing me and my colleagues with the interdisciplinary working space (campar.in.tum.de/Chair/NarvisLab) and the Bayerische Forschungsstiftung (BFS) that provided funding to initiate this fruitful collaboration between physicians and computer scientists (narvis.org). Special thanks to Christoph Stapleton for many inspiring discussions during his stay in Munich and my visit at the Institute for Simulation and Training (IST) in Orlando, Florida. I would like to thank all students I worked with in particular Felix Wimmer, Mohammad Rustaee Khoshkbijari, Arno Scherhorn, Latifa Omary, Samuel Kerschbaumer, Irene Faure de Pebeyre, Maxim Kipot and Patrick Wucherer. I also want to express my gratitude to the radiologists and surgeons of Klinikum Innenstadt München for their precious contribution in obtaining medical data and their participation in many user studies. Beside Prof. Dr. med. Wolf Mutschler, Prof. Dr. med. Ekkehard Euler and MD Sandro Michael Heining I would like to mention the following medical collaborators and supporters: Dr. med. Ben Ockert, Dr. med. Simon Weidert, Dr. med. Jürgen Landes and Dr. med. Christian Riquarts at Chirurgische Klinik und Poliklinik - Innenstadt, LMU München and PD Dr. med. Laszlo Kovacs, Dr. med. Maximilian Eder, Dr. med. Bettina Hohlweg-Majert, Dr. med. Steffen Müller, Dr. med. Timm Steiner at Klinikum Rechts der Isar, TU München. I would like to thank our industrial partners Frank Sauer, Ali Khamene, and Sebastian Vogt of Siemens Corporate Research (SCR) for the design, setup, and implementation of the in-situ visualization system RAMP they provided us; Konrad Zürl and Oliver Wenisch of A.R.T. GmbH for providing cameras and software for providing the outside-in tracking system; Nabil Ben Salah of the company Materialise GmbH and the team around Carl Fruth of the company pro-fit.de for having supported the creation of the Visible Korean Human Phantom presented in section 2.2.2. Special thanks to the co-organizers of the AMIARCS workshop 2009, Dr. PJ "Eddie" Edwards, Dr. Jan Klein and Prof. Dr. Kensaku Mori, for introducing me in the organization of an international workshop. I also like to thank the secretary of the CAMP chair Martina Hilla for supporting me at any time with administrative issues. Finally, I wish to thank all visitors of the NARVIS lab who have taken their time to review my demos and to discuss with me future steps of my work.

CONTENTS

CHAPTER

ONE

INTRODUCTION

The main purpose of visualizing anatomy has always been the generation of a mental model of the human body in order to improve health care. Physicians have to apply excellent cognitive, performance and communication skills within a highly organized, team oriented working environment. Knowledge about the patient's anatomy serves here as the fundamental bases and has to be learned, organized, filtered and applied in time-critical situations. From this finding, researchers derived already in the early days of medicine a fundamental question, which is still today highly topical.

Which is the best way to visually communicate anatomy in order to create the physicians' sustained, entire mental model of the patient's body?

Until today, dissection courses are inherent instructional units of the medical formation. Students explore the human body and exercise surgeries to prepare for their professional life. Unfortunately, learned information bases on a dead organism with limited variety of pathology. In addition, organizing dissection courses is cost-intensive and requires complex logistics. As a consequence, this lesson is reserved to only a small elite at only few points in their career. In order to record their findings and share it with a broad audience and later generations, early anatomists began to invent methods to efficiently document their research from dissections. Anatomic illustrations considered as an addition to text descriptions became more and more important to better transfer increasing knowledge of the "system of the systems". With the level of detail, also the variety of anatomy was steadily augmented. Starting from the *homo perfectus* combining objectivity, symmetry and vitality of healthy anatomy also pathological counterparts, its topology and the functional interconnectivity and physiology of the body have been illustrated.

Along with the media used to visually describe the physical structure of the human body, the art of presenting this subject has progressed over centuries. Artists, educationalists and physicians together explored intuitive presentation methods with the highest educational value to sustainable transfer and embed this complex knowledge. Two of their primary objectives were first to portray the human body, not only in terms of its physical properties, but also as a living system and second, to show anatomic parts not as standalone objects but within the context of surrounding anatomy, in the best case, within the context of the whole body. For example, Jacopo Berengario da Carpi (1460-1530) is one of the pioneer anatomists who determined the importance of combining science and art, and illustrated complete bodies in a living pose within a natural environment. Also wax modelers from the era of medical ceroplasty beginning from the late 18th century designed their wax models rather as living and sleeping persons than dead bodies. Since 1996, Gunther von Hagens' exhibition *Körperwelten* presents complete cadavers in living situations, for instance when performing sports, to individually highlight specifically affected anatomy and allow the visitor to intuitively map this information onto his/her own body.

In contrast to all previous attempts of anatomy illustration, the combination of modern imaging modalities, e.g. Computer Tomography (CT) or Magnetic Resonance Imaging (MRI), and state-of-the-art CPUs[1] and GPUs [2] allows for capturing and visualizing parts of a patient specific anatomic situation close to the time of operation. Computer scientists from the field of computer graphics have invested time, effort and creativity to transform huge amounts of acquired raw data to a more intuitive presentation form. Such realistically looking 3D/4D, real-time rendering of an anatomic region of interest is accessible on 2D monitors, which is standard equipment in today's clinical working environments. However, in many cases data is still presented as 2D gray scale printouts or monitor projections. The complex task of mentally mapping such 2D information presented in one workspace onto the 3D patient being in a second workspace has motivated researchers to apply the relatively new technology of Augmented Reality to medical procedures.

The combination of AR technology and the power of today's medical imaging modalities results in a novel approach of presenting anatomy that has the potential to overcome many problems of previous media forms. AR takes over the task of mentally mapping the patient specific anatomic information onto the patient. However, the advantage of transferring this mental task to a computer system strongly depends on the quality of visually embedding the patient specific anatomy into his body. In more technical terms, this addresses the quality of composing real and virtual entities of the medical AR scene. The objective is to generate the most beneficial view into the human body being intuitive and interactive to let the user fully immerse himself into the AR environment. Correct design of the AR scene is supposed to let the user examine and treat deep seated disease as easy as superficial wounds.

The present work *Immersive, Interactive and Contextual In-Situ Visualization for Medical Applications* addresses two major aspects how to make AR acceptable and beneficial to be applied to the medical workflow. These aspects are first the composition of the virtual and real parts of the human anatomy to correctly perceive their relative spatial position and second, the interaction with the AR environment within the intraoperative workspace.

The thesis begins with an historical review (section 1.1) on different attempts how to design the view into the human body. The review serves as a motivation to design the most intuitive and interactive 3D user interface based on AR technology. It covers the whole spectrum from antique 2D illustrations to 3D/4D rendering on modern computer graphic cards and is followed by the separated section 1.2 providing a closer look at state-of-the-art AR systems having been developed for medical applications.

[1] Central Processing Unit

[2] Graphics Processing Unit

Chapter 2 addresses the AR system, the corresponding development environment and human perception. First, the hard- and software components of the used AR system are described in section 2.1. As a contribution of this thesis section 2.2.2 presents a novel testbed that has been created to simulate more realistically clinical conditions for the target oriented development of medical AR systems. Section 2.3 addresses the human perception system with a strong focus on the preconditions defined by the AR hardware setup and its target applications. In chapter 3 two novel solutions are proposed how to correctly compose virtual and real objects of a medical AR scene. Chapter 4 presents the interactive component of the present 3D user interface. Section 4.1 evaluates three methods to enable the user to control parameters for adjusting the AR environment. Section 4.2 describes and evaluates a tangible user interface being integrated into AR environments in order to improve navigation and exploration tasks. Chapter 5 starts with an overview on medical target applications for the proposed genre of AR system, i.e. see-through head worn displays. Sections 5.1.2 and 5.3 introduce two promising applications that have been determined and evaluated together with partnering surgeons. After a discussion of the present results and an extensive outlook on future work in chapter 6, chapter 7 concludes the thesis.

1.1 The Historical View into the Human Body

The objective of scientific visualization is the translation of large amounts of complex empirical or simulated data into a visual format that can be intuitively perceived by humans. One popular example of today is the measured meteorologic data in combination with simulated predicted data that has to be prepared as the daily weather forecast to be easily understood by a broad audience. Already early scientists and artists determined the importance of intuitive visual presentation of information. The documentation and visualization of the landscapes of the world was one of the first applications to safely guide former seafarers through the oceans of the world. The human body has been determined as another world that had to be explored and mapped.

This section provides a short historical review of some of the most important trends and steps of systematically and topologically visualizing macroscopic anatomy. The initial intent of different media forms visually explaining the human body was to document expert's findings mainly gained from dissection for future generations and to better communicate an extremely complex and versatile subject. Today's renderings of volumetric, patient specific imaging data, e.g. CT, are further used for supporting surgeons in diagnostics and intraoperative guidance tasks. The dimensionality of shown anatomy as well as the ability to interact with the presented subject strongly relate to the provided media. The review will highlight different media forms that have been invented to present anatomy.

The first known anatomists Aristotle (BC 384-322) and Galenus (AD 131-201) gathered their knowledge primarily from dissecting human cadavers and animals. Dissections allow the learner to study the 3D anatomy in a 3D environment quasi at first hand. However, this learning method, then and today, is accessible only for a small academic elite and involves intensive logistic and ethical issues. In addition, only dead bodies can be explored with a limited amount of pathological cases in a limited period of time. The initial anatomic situation can not be reproduced after the lesson to provide similar instructional bases for many scholars. Public dissections had been established in the past for a paying audience. However, such events never became an instructional standard to teach the public, mainly due to ethical concerns.

The transformation of the 3D human body to a 2D presentation has been the first attempt to preserve the anatomists' knowledge from human dissection and make it accessible for a broad audience and future generations. The de Gregoriis brothers published in Venice in 1492 the anatomic book *Fasciculus medicinae* that contains for the first time in history descriptive illustrations of the human body. The original woodcuts of the illustrations belonged to Johannes de Ketham [136] who has been identified by some literature as the German physician Johannes von Kirchheim having practiced in Vienna around 1460. Already this early work showed the full body figures in a living pose revealing all views of their anatomy. In contrast to the design of Ketham's woodcuts, Leonardo da Vinci's (1452-1519) anatomical drawings around 1510 show a comparably high level of detail and were labeled with plenty of annotations. Similar to the objectives of his designs of mechanical devices, his anatomic illustrations depict the "workings of animate mechanisms" [178] in order to objectively explain topology and function of the anatomy. Leonardo da Vinci identified the importance of realistically copying his observations from own dissections and topologically presenting his findings. Figure 1.1(b) shows a famous example of his work presenting the vascular system within the context of whole body. Leonardo da Vinci was not a member of the scientifically educated upper class at his time and had hardly any access to the scientific network to publish and share his knowledge. His work stayed unappreciated for a long period of time and was rediscovered not earlier than in the turn of the nineteenth century [178]. A few decades later Andreas Vesalius, a Flemish professor of anatomy at the University of Padua (1514-1646), published in 1536 a set of large anatomical woodcuts know as the *Tabulae Six*. This work has been extended and republished in 1543 as *De Humani Corporis Fabrica*, which superseded any earlier works in the science and art of illustrating anatomy and became a standard for many generations of physicians. Vesalius's work is based on pure empiric research from dissecting human cadavers, which allowed him to correct some of Aelius Galenus' assumptions that were inherently part of the medical curriculum for hundreds of years. Vesalius closely collaborated with the Flemish artists Jan Stefan van Kalkar to achieve high quality documentations of his dissections in terms of detail, appeal and instructional value. Also Jacobo Berengario da Carpi (1460-1530), son of a physician, aspired to combine science and art in order to efficiently transfer anatomical knowledge and presented with the help of artists "the bodies of the dead depicted as protagonists of moods and morals" [178]. Di Carpi proposed his work to "surgeons, as a guide to vulnerable tendons, and to artists, as an aid in the correct design of figures" [178].

Vesalius and many further anatomists and artists invented several forms to visualize anatomy. Anatomists of the Renaissance were inspired by the cultural and social influences of their time and "dramatized, beautified, and moralized" [188] their works. Images were composed of upright, living bodies being sometimes in emotional state and positioned in natural landscapes. The intent of their image compositions was to prevent observers from thinking of death and let them exclusively focus on the subject of anatomy [178].

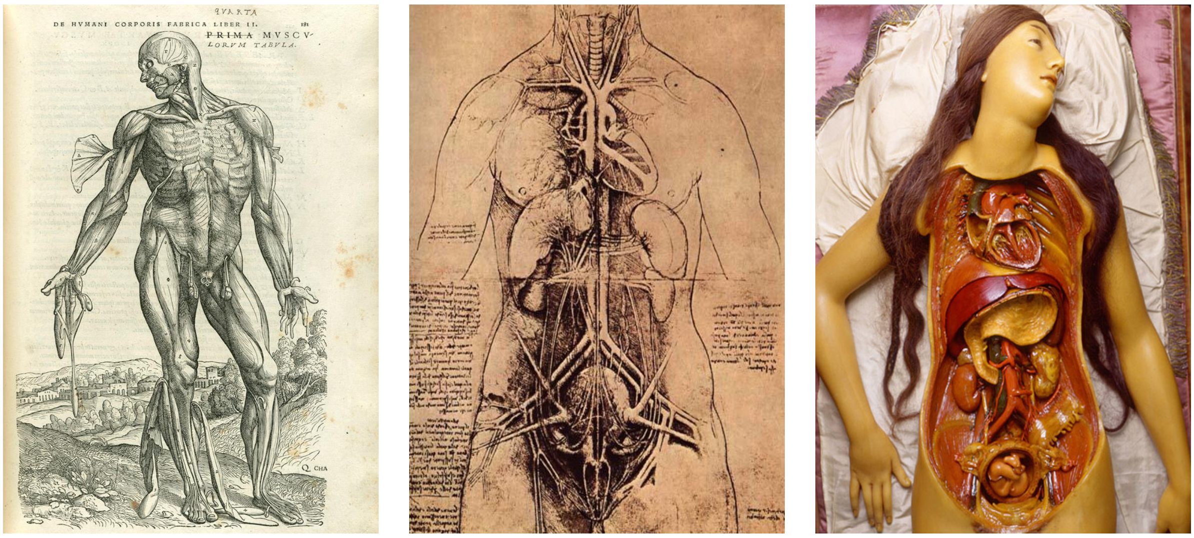
Illustrations were arranged systematically to discover the anatomy from the outer skin layer to the deep seated organs without loosing the link to the whole body. Vesalius issued along with the *De Humani Corporis Fabrica* a flap system, the *Epitome*, consisting of prints from large woodcuts that could be used to successively cover and uncover layers of the body. Exact overlay of gradually augmented flaps allowed the reader to visually dissect the body and better understand its topology.

Figures primarily showing the muscles of the body with pealed off skin were classified as écorchés. The écorchés, impressively documented the intent of obscuring the death and presenting the anatomy within a living context. Many écorchés have been designed as spreading the skin with their own hands to reveal the view onto their viscera. For instance Giulio Cesare Casserio's and Adriaan can den Spiegehel's figure TAB XII in *De humani corporis fabrica libri decem* (Venice, 1627) or different figures in Juan Valverde de Amuscos' *Anatomia del corpo humano* show standing figures exposing body holes to offer the view into their body. Following the example of Vesalius' *Epitome*, the lack of interactivity due to the print media had been targeted by the invention of anatomic flip books. Their intent was rather to amuse a broad audience than providing high level of correctness and deep insights into the anatomic subject. Two examples to be pointed out are the *Yaggy's Anatomical Study* [249] published in 1886 and the *Anatomie Modèle Femme* [225] from 1937 giving the observer the possibility to cover and uncover body layers to contextually explore the deeper seated regions.

In the 18th century a new type of documenting and presenting the human anatomy had been invented. Anatomical ceroplastics are wax models of full bodies or parts of it with the intent to present gathered information from dissection with a long lasting, realistic looking, three-dimensional media. The birthplace of anatomical ceroplasty was located in Bologna with Gaetano Giulio Zumbo (1656-1701) as the most important and first modeler who used colored wax for his medical wax models [55]. The workshop, of Bologna was followed by the Florentine workshop having been installed in the buildings of the Florentine museum *La Specola*, which was initially called the *Imperial Regio Museo di Fisica e Storia Naturale*. The museum and the workshop were directed by Felice Fontana (1730-1805) who had been positioned by Peter Leopold von Habsburg-Lothringen (1747-1792) a benefactor and connoisseur of natural sciences. At the Florentine workshop anatomists dissected the cadavers to prepare templates for the ceroplasty figures produced by wax modelers. This interdisciplinary collaboration guaranteed high quality results combining aesthetic, realism and anatomic correctness. Due to missing means to conserve the dead bodies, sometimes over 200 cadavers had to be dissected to create only one wax counterpart. Fontana's vision was to replace cadaver dissection for the medical formation by anatomical ceroplasty [55]. Education of professionals and the public motivated Fontana to further augment the selection and to distribute his models through the early scientific centers of todays Europe. Also other anatomists began to understand that there was a market for "moral, high minded people wanting to study anatomy but repel[led] by dissection" [16]. For example, Louis Auzoux, a French surgeon (1797-1880) founded a company to produce wax models "with the intent to replace dissection" [16]. Even though high realistic models have been developed showcasing all kind of anatomic variates and diseases, in fact dissection of cadavers, which is still today an inherent part of the medical professional curriculum, never could be replaced. The construction of wax models was extremely expensive and time-consuming. Many cadavers had to be dissected serving as sample in order to reach the targeted level of detail and realism. It was unthinkable to explore interactively the wax anatomy or simulate a surgery (which would have destroyed the wax figure) in the same ways as the dissection of cadavers. Similar to cadaver dissection anatomical ceroplasty presented the anatomy again in 3D within a 3D environment. Further anatomists understood that the material wax allows for an exceptional realistic replication of human tissue. "Its plasticity, instability, [..] and thermal sensitivity suggest the imagination of flesh" [55]. True-scale wax models were realistically colored, showed fine details and were posed as living bodies "to achieve a precise and a lifelike representation" [143]. Medical ceroplastics also document the attempt to combine science and entertainment in order to educate and "delight" the public [16]. Marjorie Winslow the creator of a series of gynecologic ceroplastics for Queens's University, Kingston stated in 1940 that ceroplastics form a "very real art in reproduction of life's drama" [143]. In particular one category of models known as *Medical or Anatomical Venuses* became popular to attract the publics's interest on the subject. The majority of Venuses, in most cases designed as pregnant woman, have a modular structure. Topologically arranged layers and modules can be removed to get access to deeper regions of the body. One of the most famous examples of this genre is Clement Susini's Medical Venus (see Fig. 1.1(c) that is today exhibited at the museum *La Specola*. Due to their aesthetic appeal, Venuses effectively helped to obscure "the links between death, dissection and anatomical knowledge" [16]. Anatomy can be explored "without exciting the feeling of horror that men usually have on seeing corpses" (Guillaume Desnoues, 1650-1735) for the "instruction and delight both of the learned and the unskill'd in Anatomy" [16]. At that time, the presentation of Venuses as living bodies also helped to avoid the taboo of exposing the public with dead bodies [42] (see also info table 1.1).

A new form of presenting real anatomy of dead bodies in 3D within 3D environments has been developed by Gunther von Hagens in 1977. His method called plastination allows the odorless, long term preservation of cadavers. In contrast to embedding cadaveric specimen into plastic material for preservation, the plastination impregnates the tissue from inside and allows for direct access of the anatomic structures. Gunther von Hagens' ambition to present anatomy in a novel manner came under criticism as soon as he launched his first exhibition *Körperwelten* in Japan in 1996. In particular, opponents have criticized the public exposition and commercialization of the dead bodies. However, according to the programme managers of *Körperwelten*, the primary objective of the exhibition is health education [230]. Gunther von Hagens' exhibitions aim at effectively transferring information to a broad audience that has previously been accessible only for a small group of experts. First the exposed subjects shall raise the visitor's interest on anatomy and health care and second it shall enhance the awareness of the sensitivity of their own body and motivate a more responsible lifestyle. When collocating real pathology, for example a smoker's, lung next to a healthy counterpart, this information transfer is driven by much more immersive fascination than presenting the same information with classic instructional media such as 2D printed illustrations or plastic manikins [3]. Posing the exhibits in a

[3]"Von echten Präparaten geht zudem eine sehr viel eindringlichere Faszination aus als von Kunststoffmodellen." [230]



(a) Écorché by Andreas Vesalius (b) Vascular system, in-situ, by Leonardo da Vinci (c) Medical Venus by Clemente Susini

Figure 1.1: Early media to visually describe anatomy. Image 1.1(c) courtesy of Museo Zoologico e di Storia Naturale della Specola, Italy

living situation can help to individually highlight specific anatomy, for example muscular structures, and map this information to one's own body [4]. Similar to the objectives of anatomical ceroplasty, the vivid posing of cadavers avoids a shock effect that might hinder the knowledge transfer and rather uses this aesthetic element to stimulate the visitor's imagination [5].

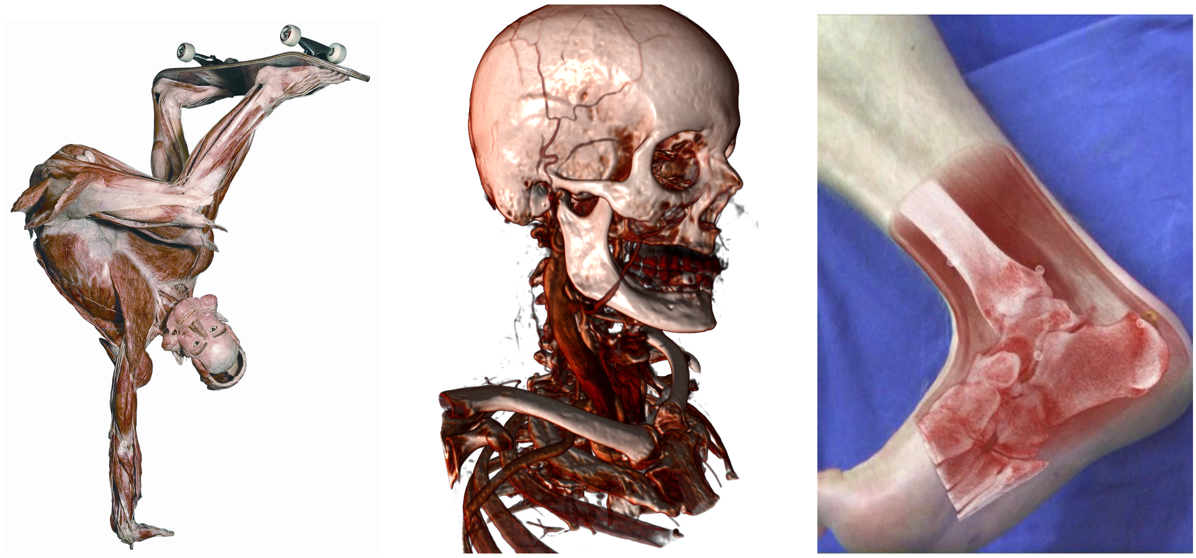
Today's standard computers reduce the presentation of anatomy again to a 2D monitor, however, real-time interaction with 3D rendered anatomy models and animation over time (4D) interactively enables additional views. Today's commercially available 3D stereo hardware, for instance the Nvidia Geforce 3D Vision [6] increase the spatial perception of objects. They appear to be located in 3D space, even though the point of view is still spatially limited. Today's anatomy atlases are commonly sold together with virtual 3D anatomy models and software to interact with this data to augmented printed instructional media. Beside desktop based atlases, the increasing bandwidth of private internet connections have inspired different companies to also offer online versions, for example visiblebody.com, innerbody.com and e-anatomy.org.

As an alternative to the instructional purpose, rendering of patient specific anatomy from data sets acquired with modern medical imaging modalities is capable of supporting the clinical workflow. Visualization of such data in Computer Aided Surgery (CAS) is able to support diagnostics and instrument guidance. While previously, surgeons had to keep

[4]"Die Posen ermöglichen es den Besuchern, das Plastinat besser in Bezug zum eigenen Körper zu bringen" [230]

[5]"... von der Erkenntnis, dass ich ein ästhetisches Element einführen musste, um das Publikum nicht zu schockieren, sondern seine Vorstellungskraft anzuregen" [230]

[6]Nvidia Corporation: http://www.nvidia.com



(a) Plastination (b) Direct volume rendering (c) AR in-vivo study

Figure 1.2: Recent media used to describe human anatomy. Image 1.2(a) courtesy of Gunther von Hagens, Institut für Plastination, Heidelberg, www.koerperwelten.com; Images 1.2(b), 1.2(c) courtesy of Oliver Kutter and Jörg Traub, Chair for Computer Aided Medical Procedures & Augmented Reality, TU München, Germany

their learned mental model of the anatomic situation around the deep seated operation site in mind, modern imaging modalities such as CT or MRI are able to scan the body within a few moments to show the current pathologic situation. In addition, such illustrated information can be registered with the patient and surgical instruments to facilitate the performance and sometimes enable operations that have not been possible before. In order to efficiently handle the huge amount of scan data from the clinical workflow, computer scientists invest increasing efforts to develop methods to effectively extract the important information. Various approaches for the visualization of volumetric, medical imaging data such as *Computed Tomography (CT)*, *Magnetic Resonance Imaging (MRI)* and *Ultrasound* have been proposed in the field of computer graphics [127, 59, 43, 39] and evaluated for purpose in the medical workflow [173, 145]. The need for techniques to explore medical volume data maintaining the ability to understand spatial relations of visualized objects within the volume data is lively discussed. Burns et al. [43] presented an importance-driven rendering method to highlight focus regions in visualized medical imaging data. Bruckner et al. [39] introduce "context preserving volume rendering model" [39] employing *ghosting*, which is a subdiscipline of the computer graphics domain *focus & context visualization*. Their method uses "a function of shading intensity, gradient magnitude, distance to the eye point" [39] to define transparency of occluding layers. Krueger et al. [126] introduced *ClearView*, an "interactive framework for texture-based volume ray-casting" [126]. Their method uses transparency combined with different shading effects to support the perception of spatial information within a region of interest inside the visualized volume data. Their visualization method uses a highlighted ring painted on the skin surface around the

Scenes observed at Körperwelten: Eine Herzenssache (Augsburg, July 2009)

A little girl, about five years old, is standing in front of a vitrine exhibiting a rocking horse ridden by a child cadaver. The rocking horse is the cadaver of a foal. The first question of the little girl is whether the rocking horse is real. The mother confirms. The second questions is whether her rocking horse at home is also real.

A middle-aged woman is standing in front of a vitrine showing a plastinated healthy lung collocated with a completely black smoker's lung. Jokes from her partner disclose that the woman smokes, however, it seams that she is not able to pay attention to what her partner is telling and her eyes keep staring like magnetized at the smoker's lung.

region of interest. Diepstraten et al. [53] propose a small set of guidelines of how to apply transparency to the automatic generation of illustrative industrial visualizations to communicate interior structure of objects. According to the rules "faces of transparent context objects never shine through other transparent faces. Opaque objects which are occluded by two transparent objects do not shine through. Transparency falls off close to the edges of transparent context objects and increases with the distance to edges" [53]. Despite providing a clear view onto the focus, contour rendering has the drawback that details of the transparent objects are lost and that it ignores material and surface information of the context objects [53]. Furthermore it only provides "two transparency states: fully opaque or fully non-opaque; semi-transparency cannot be visualized" [53]. Interrante et al. [102] determined "ridge and valley lines" providing visual cues to "communicate the shape and position of the transparent skin" and occluded anatomical objects. Another interesting approach is presented by Levoy et al. [132]. They map a solid, simple shaped texture on an outer semi-transparent surface model occluding an inner opaque surface model to "improve the ability to distinguish the shape of the surfaces and their relationship to each other" [132].

The reviewed attempts and media forms, i.e. cadaver dissection, anatomical drawings, ceroplasty, plastination and computer based rendering, tend to follow at least a subset of the following principles:

- Preservation of Topology: Anatomic components are presented systematically and topologically within their anatomic context to better create an overall map.

- Illustration of Vitality: The human body is designed not only as a physical structure but also as a living system to prevent the observer from thinking of death when looking at the subject.

- "Nosce te ipsum" (know thyself): Medial approaches are equipped with artistic elements to better transfer knowledge for giving health care and to gain awareness of the fragility of one's own body.

- Interdisciplinarity: Interdisciplinary teams consisting of anatomists, physicians and artists composed their individual abilities to develop the most advantageous form of presenting anatomic information for instructional purposes.

- Share and Distribute Knowledge: The researchers always intended to find ways to make their novel medium accessible not only to a small elite but also to the public.

The ability to visually scan the body of ones fellow human beings or to see through physical barriers is a human dream that has inspired different communities. Science fiction authors Jerry Siegel's and Joe Shuster's equipped their superhero Superman with the ability to see through real objects. Today's artists use novel rendering techniques to effectively design the contextual view into the body, e.g. Alexander Tsiaras [224]. Others like the photographer Katherine Du Tiel project within the scope of her photo series *Inside/Outside* the interior anatomy onto the skin of her human photo motifs. The idea of projecting anatomy onto the skin surface is also used for instructional issues in order to teach palpation. A series of recent books suggests to augment the learning effect by drawing hidden structures onto the skin to train diagnostic tasks [244, 176, 177].

Recently the application of backscatter X-ray raised public emotions. It is supposed to support security checks in public transport systems. On one hand people are fascinated of the ability of discarding the clothes in order to detect hidden weapons. On the other hand emotional and ethical reasons prevent this invention to be widely accepted and applied. Systems have been successfully tested at different airports, however, for example Germany forbid using the system for security checks due to violation of privacy and protection of human dignity [7].

Based on advances primarily in the disciplines of computer graphics, computer vision and display and camera technology scientists began to realize the dream of creating the view through the skin layer onto deep seated regions. Section 1.2 will review previous work published by the *Medical Augmented Reality* community. The review will strongly focus on head worn AR Systems with a closer look at targeted applications, their interactive components and visualization methods.

1.2 Augmented Reality Enabled View into the Human Body

Augmented Reality [12, 13, 151] considered as a new media form is a technology that has to incorporate man power and knowledge from multiple disciplines. The team composition is strongly depending on the target application. Strong efforts have to be undertaken to create developer platforms that combine the expertise of collaborating research domains. These developer platforms are both, virtual and physical. Software frameworks such as ARToolKit by Kato and Billinghurst [113], Bazar by Lepetit et al. [130] or CAMPAR by Sielhorst et al. [203] serve as the functional base and provide interfaces for sensor and display hardware, software for object tracking, visualization and registration. The physical platform has been

[7] status quo: 22.01.2010

described by Christopher Stapleton, IST, UCF, US as *real world labs*. Especially when the technology has to be applied to other highly specialized disciplines, for instance medicine, intensive multidisciplinary collaboration is fundamental for the successful development of systems that exactly fulfill the needs of their future application. Cleary states in his survey about research in the domain of medical robotics, which shares many problems with medical AR, "the completion of a medical robotics project requires a partnership between engineers and clinicians that is not easy to establish" [48]. Prof. Navab, head of the chair for Computer Aided Medical Procedures & Augmented Reality, TU München, Germany and Prof. Mutschler, head of the Trauma Surgery Department, Klinikum Innenstadt, LMU, Munich, Germany established together a *real world lab* that is called today the *NARVIS lab* [8]. The name has been derived from the preceding *NARVIS* project [9] that has been funded by the *Bayerische Forschungsstiftung (BFS)*[10] and successfully completed in August 2007. Similar to earlier innovative, interdisciplinary workshops (see section 1.1) that have been installed to create media driven views on human anatomy, the NARVIS serves as a physical platform for strong and fruitful collaboration between surgeons, computer scientists and engineers and for seamless incorporation of young talents and their creativity at the intersection point between education and profession.

The following section completes the literature review of approaches that enable the view into the human body and adds Augmented Reality to the media portfolio.

1.2.1 In-Situ Visualization for Medical Applications

User interfaces (UI) are crucial to take full advantage of any method, algorithm or functionality that is supposed to facilitate real world tasks. The objective of Engineering Psychology is the reduction of user's stress and frustration by incorporating knowledge about cognition and perception when designing UIs. Standards have been defined to ensure high quality Ergonomics of Human System Interaction (ISO 9241), however, the reality has shown that many of today's software products still suffer from shortcoming usability. Marleen Brinks [37] reports that 70% of IT specialists have experienced emotional or physical aggression against their computers. Victims of physical attacks are commonly the mouse (31%) and the monitor or computer case (15%). Only 24 % claimed that they have never felt any aggression.

The patient's benefit from minimally invasive surgery, which is one targeted application domain of this thesis, has been proven in many surgical domains. However, methodology of workspace setup and instrumentation considered as the user interface to perform surgery has strongly been adopted from open surgeries. The insufficient adaptation of the surgeon's workspace to the new operative interface has resulted in significant drawbacks for the surgeons. Park et al. [163] reports that 86.9 % of surgeons performing laparoscopy on a daily bases report physical symptoms and discomfort, which are caused by insophisticated interface design of instrumentation and display position and technology.

User interfaces in the medical domain, in particular when being applied intraoperatively,

[8]NARVIS lab: http://campar.in.tum.de/Chair/NarvisLab
[9]NARVIS project: http://www.narvis.org/
[10]Bayerische Forschungsstiftung (BFS): http://www.forschungsstiftung.de/

demand special requirements, for example precise control, ergonomy, robustness against malfunction, breakdown and faulty operation, sterility, minimal complexity and strong orientation to the workflow of target procedures. Peters [167], [168] provides an overview on approaches for image guided surgery and covers the full spectrum of systems. One of his major conclusions states that the "greatest challenge to the successful adoption of image-guided surgery systems relates not to technology, but to the user interface." Also Cleary *et al.*[48] identifies the user interface design, in this case of surgical robotic systems, as one of the key factors for usability and acceptance of IGS systems.

The "lost direct visual connection to the operative field", limited degrees of freedom of movement regarding instrument guidance and visual axes being "at odds with working or motor axes" are major factors causing discomfort for the surgeon in laparoscopy [163]. However, many other IGS systems share these factors, whenever imaging data is presented on remotely positioned monitors, which results in the division of the surgeon's visual and haptic workspace. With respect to neurosurgery, Friets et al. stated in 1989 "part of the gestalt of neurosurgery is to take the two-dimensional information and reconstruct it in ones's mind as a three-dimensional object. Then, depending on the surgical approach, this three dimensional object is mentally rotated and sliced as the neurosurgeon mentally tries to reconstruct what should be seen during the surgical procedure" [78].

Augmented Reality has been identified by researches referenced in the following sections 1.2.2.1 and 1.2.2.2 as a new genre of user interface to optimize man-machine interaction for particular tasks in the workflow of particular procedures that may benefit from image guided surgery (IGS). Figures 1.3 show two of our targeted minimally invasive procedures that intensively use volumetric x-ray data. Both procedures require the surgeons to guide instruments with the help of imagery, CT in the case of vertebroplasty and DynaCT in the case of drilling in spine surgery, that is presented on a monitor.

> **"... part of the gestalt of neurosurgery is to take the two-dimensional information and reconstruct it in ones's mind as a three-dimensional object. Then, depending on the surgical approach, this three dimensional object is mentally rotated and sliced as the neurosurgeon mentally tries to reconstruct what should be soon during the surgical procedure" [78].**

The motivation for medical in-situ visualization [189] is to avoid the complex task of mentally mapping information from 2D prints or monitor projections [142, 163] onto the three dimensional patient [78]. It is capable of combining all necessary information for task completion within only one workspace. Beside imaging data this may involve vital parameters, navigation data, planning data and patient specific additional information. Regarding intraoperative navigation, surgeons see the rendered anatomy from patient's specific imaging data, the patient and their working hands in their primary field of view directed to the operation site. Ergonomic constraints such as those described by Park et al. may be reduced or even solved when combining visual, working and motor axes and avoiding remotely located monitors [163].

We propose to distinguish two levels of AR enabling user interfaces. The first level addresses the physical hardware interface such as a see-through head mounted display [15,

In-situ visualization

In-situ visualization in medicine is a term describing the presentation of patient specific data superimposed onto the patient's body. Sauer et al. introduced the term for medical augmented reality describing the criteria that "...anatomical structures are being displayed at the location where they actually are" [189].

194, 31], a camera-monitor setup [88] or mobile phone approach [10]. An overview of AR enabling, physical hardware interfaces related to display technology in the medical domain is provided by Sielhorst et al. [202]. The second level addresses the AR scene as the user interface to correctly perceive and directly control virtual and real objects. The second level has been targeted by the present thesis.

(a) Navigated pedicle screw implantation

(b) Vertebroplasty

Figure 1.3: The surgeons' workspace is separated to the monitor showing navigational information and the patient.

With respect to the AR scene considered as a UI, Milgram et al. have defined in 1994 the well known *Reality-Virtuality (RV) Continuum* [151]. Here Augmented Reality (AR) has been classified as a subdomain of Mixed Reality (MR). According to the amount of virtuality/reality of a scene, the continuum lines up the *Virtual Environment*, *Augmented Virtuality*, *Augmented Reality* and the *Real Environment* (see Fig. 1.4). Azuma specified in 1997 AR systems as systems that combine real and virtual, that are interactive in real time and which register objects of the AR scene in 3D space [12]. Beside the augmentation

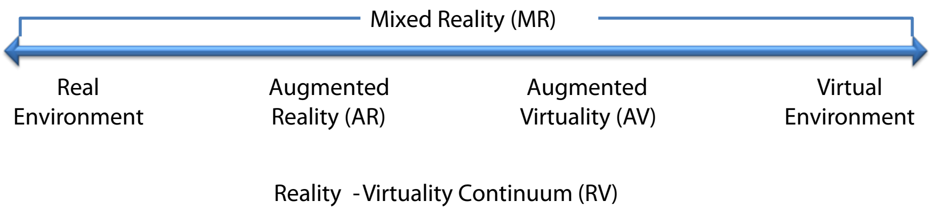


Figure 1.4: Reality-Virtuality Continuum introduced by Milgram in 1994 [151]

of visual information, which is the primarily addressed sense in the AR community, also further human senses have been augmented such as palpation or smelling. Peters mentions in his survey on image guided surgery that augmented reality ”refers to the situation where real-world images (video, endoscopic, ultrasound, or even electrophysiological) are integrated with the visualized model” [168]. The perceptive range is further extended to ultrasound that can not be perceived with human senses.

When applying MR to medical training and learning scenarios, Milgram's RV Continuum does not cover all perceptive sources that immerse a user into an MR/AR environment. Stapleton proposed an extension of the techno-centric view of Milgram's RV Continuum. Beside composition of the real and virtual world also the cognitive world has to be taken into account, of which gradually adds the user's imagination to perceive an AR environment. In order to explain the impact of imagination, one may consider the example of role plays that have become a standard media to teach and train teams in dealing with time-critical and highly stressful and emotion inducing tasks. Instructional designers develop extreme scenarios to prepare teams of firefighters, soldiers or clinicians for emergency cases that might happen during their professional life. The scenarios incorporate actors and patient puppets, burning houses or realistic surgery rooms. The major component of an effectively designed scenario is the story that produces stress, time pressure, hecticness, panic or further unpredictable behavior of the protagonists. For example, protagonists are assigned with fictive names and personal background and get different rules within the team to experience different perspectives. The workflow, for example the surgical one, can be interrupted by changing parameters, for instance unpredicted complications during operation, infrastructure breakdown (no standard electricity and light), noise, collapsing patient or surgical staff, burning smells, patient overload, etc. The cognitive input lets the trainees individually perceive the current situation in a training scenario. It causes individual reactions that are difficult to be predicted. However, this corresponds to real life situations and has to be taken into account of, when preparing teams and their members for real tasks of their professional life. Stapleton's continuum shown in figure 1.5 [206, 101] addresses the importance of imagination to be part of the composed environment into which the user has to be immersed.

1.2.2 Literature Review

The medical AR scene can be presented with various hardware setups such as ”AR windows” [197] and comparable semi-transparent panels [32] rigidly installed above the patient, endoscopic cameras [64], different types of see-through HMDs or camera or monitor based systems [134, 89, 88, 146, 169]. Regarding the first level of AR UIs, Milgram et al. [151] classify different display technologies serving as the physical UIs to present AR worlds. Sielhorst et al. [202] provide an comprehensive review of different technologies enabling Augmented Reality to be used in the clinical workflow. The following survey addresses primarily head worn AR systems as the physical user interface/display technology (first UI level) that allow the observation of the augmented patient from the natural point of view, the surgeon's eyes. The review will strongly focus on

- the interactive components of the systems, which is related to the user interface as part of the AR scene (second UI level),
- proposed targeted applications and the clinical benefit of AR,

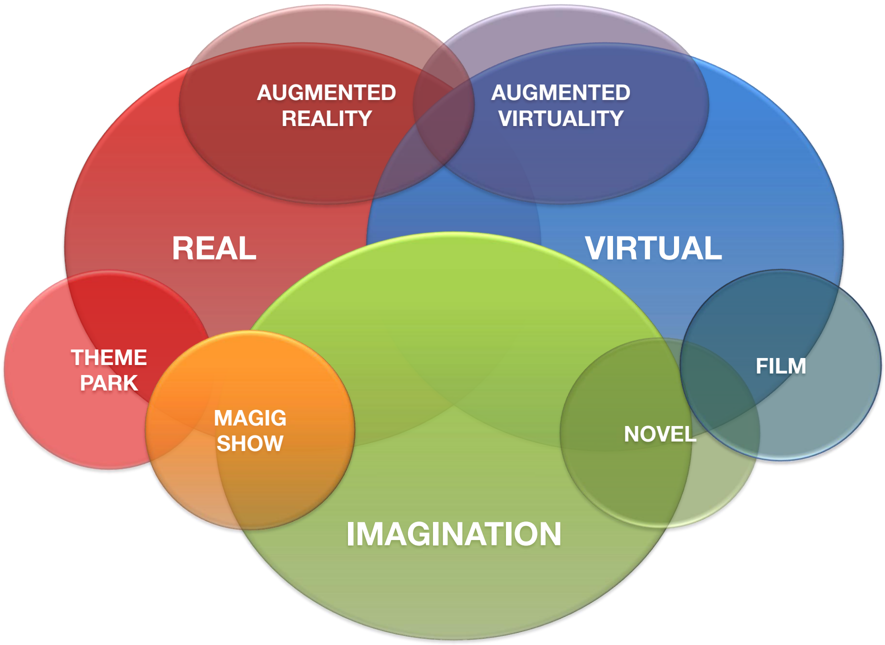


Figure 1.5: Stapleton's Mixed Reality Continuum incorporates imagination as a third extreme beside reality and virtuality [206, 101]. Image courtesy of Simiosys Llc, and University of Central Florida (UCF), USA

- visualization methods to communicate the target anatomy and navigation information,
- test beds and phantoms to simulate the patient and evaluate the AR systems,
- and approaches to improve depth perception of objects within the AR scene, which is related to the image composition of reality and virtuality.

The evolution of referenced AR systems is chronologically described while each paragraph addresses the work of one research group.

1.2.2.1 See-Through Head Mounted Displays

Even though this review is supposed to highlight AR systems being applied in the medical domain, one has to honor Sutherland's work of 1968 who introduced the first optical see-through head worn display [211]. His system was the result of fundamental research exploiting extensively the computation hardware of his time. In order to render simple wire frame models (3000 lines at 30 fps), he had to deal with issues such as clipping methods to render only in the user's field of view or with the missing depth buffer to handle the "hidden line problem", i.e. virtual occlusion. Beside ambiguous interpretations of shapes and positions of wire frame models being registered within the augmented world,

(a) Simulation of working environments outside the hospital.

(b) A fully controllable simulation environment for training intraoperative scenarios.

Figure 1.6: An instructional environment to train clinical teams at Institut für Notfallmedizin und Medizinmanagement (INM), Klinikum der Universität München. Projection of different environments onto the walls around the ambulance enhance the immersion of learners into their trainings scenario.

Sutherland reports that the "view of the objects was limited to certain directions". For this reason, some of his subjects misinterpreted the shape of an exposed cyclohexane molecule.

According to the knowledge of the author, Bajura et al. [15] are the first reporting on the medical application of a video see-through HMD. Their first version of the system comprised a stereo HMD, however, only one camera captures the reality and only the left eye is exposed with cameras images augmented with virtual ultrasound images. The right display exclusively shows synthetic imagery. In combination with US, the authors mention cardiac diagnosis, needle aspiration biopsy of a suspected breast tumor, impending surgical incision in neurosurgery and ultrasound based examination of pregnants as potential target applications. The latter has been realized and demonstrated in a first clinical trail. One of the "remaining technical problems" stated was the misleading perception of US images that "do not appear to be inside the subject" [15]. The authors introduce a "shaded polygonal pit" providing occlusion cues among virtual objects (see Fig. 1.7). However, everything real in between the user's viewpoint and the pit is consequentially hidden, such as the surgeon's hand and instruments, which reduces the usability of the augmented scene. Interestingly, the authors shortly discuss the aspect that "human visual system is not accustomed to seeing structure within opaque objects" [15]. For this reason their research cannot be "guided by a gold standard" [15]. State et al. [207] conducted two further experiments with pregnant subjects in order to analyze "How well could we see the fetus if we had sufficient compute power". The authors address the volume reconstruction of US sweeps and distinguish the importance of real time presentation to make a medical AR scene useful. State uses the tracked US probe to digitize reference points on the patient's abdomen for the registration of a virtual "pit" serving as a window into the body (see Fig. 1.7). In 1996 Fuchs et al. [79] from the same group propose applying the HMD- based AR system to ultrasound-guided needle biopsy of the human breast. They mention also sampling of lesions within the liver, kidneys, pancreas and adrenals as well as localizing "foreign body fragments resulting from accident and trauma emergencies". They mention that their system can help surgeons to assess the geometric relationships between US image and needle as well as between the US image and the patient's body. The anatomy shown in US imagery is usually scaled when being presented on the monitor. With the AR system, however, the size of the imagery corresponds to the scanned anatomy. Regarding the clinical workflow, the authors postulate that the AR guided needle biopsy can reduce the time of the procedure, reduce training time for physicians to learn how to perform biopsies, increase the accuracy of needle targeting and reduce trauma to the patient. The authors further report on their strategies how to smoothly incorporate the system into clinical workflow. The authors selected biopsy as the first target procedure, since patients are "generally healthy and able to cooperate with he use of the new equipment". The procedure can also be performed outside a hospital setting, e.g. at the developer lab space, since the "breast contains no major blood vessels or vital organs that might become accidentally damaged if the equipment were to unexpectedly malfunction" [79]. For the patient study, the authors asked the surgeon to test the system during the planning procedure to experience the augmented view. AR vision has been used to partially insert the needle towards the target lesion until "the physician was forced to abandon the AR guidance and continue the intervention with conventional ultrasound guidance" [184]. During the procedure "the HMD was worn by a separate observer" enabling a view "over-the-shoulder" [208] to identify remaining problems of the system. The authors further propose to track the needle in addition to the US probe and the HMD in order to show its virtual counterpart with further virtual objects inside the body. The needle can be extended with a projected trajectory to support "aiming". They also propose a "target enhancement marker", that can be interactively positioned to label detected lesions within space (see Fig.1.7(c)). The US slices are visualized in a way so that they fade away as soon as their position in space does not correspond anymore to the probe position. Motion of the US probe therefore produces a volumetric appearance of the US scans. In the same year (1996) State et al. [208] provided a more

(a) US images [15] (b) US reconstruction of a fetus [207] (c) US image and augmented needle [79]

Figure 1.7: A "virtual hole" or "synthetic hole" to see inside the subject. Image courtesy of Department of Computer Science, University of North Carolina at Chapel Hill, USA

technique oriented report on new advances of the system. State points out the importance of stereo AR views that enable depth perception to let the system become acceptable to the physicians. Care has to be taken, to correctly align the convergence angle of display and camera (convergence angle 4°, 64 mm interpupillary distance, 12.5 mm focal length for a estimated working distance of 50 cm). Mismatching angles "would cause eye strain". State also proposes a method to partially correct the misleading depth perception due to unnaturally occluding virtual US images superimposed onto the real skin. Non-visible virtual surface models of the skin surface and the US probe are rendered to the depth buffer for depth masking the AR scene. Penetrating US images now "disappear below the skin of the patient, except within the pit, where they remain visible" [208]. In 1998 the same group presented a new generation of the HMD-based system that uses a laparoscope camera to capture the internal anatomy [80]. Structured light is projected into the field of view of the laparoscope camera to restore depth information from fragments in the 2D video images. The resulting video textured 3D map is then presented through the previously mentioned synthetic hole [15, 207] to provide the physician with a view through the exterior anatomy of the patient onto the deep seated anatomy video recorded by the laparoscope camera. In contrast to a monitor based presentation of laparoscope video images, the views of the anatomy and used endoscopic instruments are aligned with the natural viewpoint of the observer to improve hand-eye coordination. The system also incorporates occlusion handling as introduced in [208].

In 2002 Rosenthal et al. [184] presented the results of a phantom study for US guided needle biopsy using an advanced version of the HMD-based system described in [15, 79, 208]. Comparing the standard method with the AR-driven method reveals a significant ($p<0.02$) smaller mean deviation when targeting the lesion with the AR mode. The visualization of navigation information has been further improved. A projection of the tracked needle onto the US image plane as well as the projected trajectory of the needle is shown.

In 1995 the group around Kancherla [112] and Wright [246] propose an optical see-through HMD based AR system to be used as an instructional tool to teach radiographic positioning. The first target application addresses the positioning of elbow joints to produce useful radio graphs for diagnosis. They speculate that students "will be better prepared for clinical work" when learning "through multiple senses" since they get "immediate feedback about the accuracy of the positions they create" [112]. In addition, students do not expose themselves to radiation when practicing and repetition of a learning unit is not limited. Rolland et al. [181] from the same group further described their progress in simulating joint movement in real-time and registering real and virtual anatomy. The application domain is extended to generally instruct the anatomy, physiology of dynamic bone structures. In addition, the system can be used to "study joint changes related to lifestyle choices and pathologies" and to explain patients "medical diagnoses, procedures, treatments, prognoses, and consequences of lifestyle choices such as obesity and high-impact exercise". AR as an instructional tool can support educational trends such as "active learning, individualized learning, and an environment that simulates real-life situations" since it requires the learner to actively use multiple senses to study the dynamic anatomy projected on a living and responding subject, for example a patient or a fellow student. Authors highlight that AR provides a "holistic sensory approach" presenting the anatomy in its natural context. In addition, the user can simultaneously view and palpate the anatomic region of interest. In [14, 182], they present first results of their *Virtual Reality Dynamic Anatomy (VRDA) Tool* that superimposes virtual knee joints onto a leg phantom. A future application of the system includes the visualization of ligaments and muscles with respect to the bones [9]. Rolland having mainly worked with optical see-through HMDs and Fuchs having used video see-through HMDs compare in [182] both types of head worn AR systems. In their review on previous work on HMDs being applied in the medical field, Fuchs adds catheterization as a possible clinical application. Regarding the composition of real and virtual objects Rolland and Fuchs state "the main goal of augmented reality systems is to merge virtual objects into the view of the real scene so that the users's visual system suspends disbelief into perceiving the virtual objects as part of the real environment" [182]. In 2003 Rolland et al. [183] introduce a training tool for performing

> **"..., the main goal of augmented reality systems is to merge virtual objects into the view of the real scene so that the users's visual system suspends disbelief into perceiving the virtual objects as part of the real environment" [182].**

endotracheal intubation based on a head-mounted projective display (HMPD). Their setup incorporates a Human Patient Simulator[11] that is equipped with retro-reflective material to reflect the anatomy that is projected onto the manikin. The anatomy is segmented from the male Visible Human. Their system is supposed to combine palpation and augmented vision to enhance the learning experience. Santhanam et al. [187] present a physics-based and physiologiy-based approach for modeling real-time deformations of lung models. The previous static lung model superimposed onto the Human Patient Simulator was replaced by an animated counterpart aiming at enhancing the learning experience. Authors suggest further clinical procedures to be trained with their system, which are the "investigation of patient's general breathing patterns and discomfort", "endoscope needle insertion and radiation oncology treatments", "clinical guidance" and "preplanned procedures such as lung transplants" [187]. Cakmakci and Rolland [44] further review previous work focusing on optical engineering and human perception aspects. Both publications [182, 44] are must-have-read-and-understood papers when designing new see-through HMDs.

Birkfellner et al. equipped an operating binocular with AR capabilities [31, 70] to be applied for the computer aided insertion of endsteal implants in oral and maxillofacial surgery. Birkfellner et al. suggest to use "surgical instruments such as operating microscopes" to realize AR views "since clinical acceptance is easier to achieve." Their strategy to smoothly integrate AR technology into the surgical environment considers AR as an "add-on to a normal navigation system", which does not negatively affect or obscure the surgeon's standard information sources. Regarding the visualization, their scene is augmented with virtual counterparts of the implants and the drill. In [71] Figl et al. propose to reactively render the virtual objects in a wire frame mode instead of solid polygonal models when the distance of the drill is closer to the target region than a predefined threshold. A skull phantom has been equipped with inserted metal spheres serving as the navigation targets

[11] Medical Education Technologies, Inc (METI)

to compare the performance of two different tracking systems. The authors of [71, 29] further highlight factors that have to be addressed to clinically establish AR. The system has to provide solutions for a ”wide variety of clinical specialties beyond the current classic applications.” This can only be achieved by installing a strong collaborative network to share expertise knowledge. The development of a high-quality user-interface has to transfer complex technology into an ”embedded system in the sense that it is no longer a dominant additional high-tech apparatus [...] but a simple, easy-to-use piece of equipment providing an aid to the surgeon in clinical routine use.” Wanschitz et al. evaluated the system with

The success of clinical AR depends on the user-interface, which has to become an ”embedded system in the sense that it is no longer a dominant additional high-tech apparatus [...] but a simple, easy-to-use piece of equipment providing an aid to the surgeon in clinical routine use” [71].

an in-vitro experiment. They positioned interforaminal implants in dry human mandibles and measured the positional and angular deviation between the planned and achieved implant poses. Birkfellner et al. present an experiment to assess the impact of stereo vision and additional navigational information for accurate instrument guidance [30]. For this reason, they attached metal target spheres inside a skull that was then filled with jelly to obscure direct vision of the targets. Subjects had to guide a needle probe through the jelly to the spheres using either a monoscopic vision or stereo vision. In addition, they render the target spots again as wire frame models as proposed in [71] to provide additional hints for the guidance task. Their result confirms that stereo AR vision can improve navigation. Efforts on advanced visualization methods for navigation pay off.

In 1995 Wagner et al. [234] started a series of in-vivo experiments with a monocular video see-through HMD for image-guided stereotactic navigation in tumor surgery of the head. Here, the bioptic excision of a tumor located in the upper jaw has been performed. Their augmented reality or ”composite reality” [236] system bases on the so called Virtual Patient System (VPS) of the company ARTMA Biomedical Inc., Vienna, Austria. According to the authors, ”the ultimate objective of any computer-assisted surgical device should be to support the physician during the actual surgical procedure in the least obstructive way” [234]. Their advances in the field of AR bases on foregoing research regarding the use of interventional video tomography (IVT) [223] for image guided surgery. Authors claim that IVT may make ”the use of 3D digitizing probes for the patient image coordinate transformation” obsolete. An IVT data set consists of video images, reconstructed surface topography from video images and motion data from tracking devices. After data acquisition, the IVT data set can be used to plan a navigated procedure, for example the trajectory of an instrument. In addition, further imaging data such as CT or MRI can be registered with the IVT data. During surgery the IVT video data is replaced with live video streams from a head mounted camera. Instrument guidance is supported by a series of superimposed squares indicating the planed trajectory (see Fig. 1.8(a)). The target region is shown as a two colored spot. In [235], further clinical cases are reported. For osteotomy, the authors support guidance of the saw by drawing red lines for the palatal arteries that have to be avoided and a blue line for the preplanned path of the

(a) Intraoperative AR scene: virtual trajectory superimposed on life video [234]

(b) Virtual extrapolation of needle and color feedback for targeting [192]

(c) Radius proportional to distance [229]

Figure 1.8: Augmented information for navigation in different video see-through HMDs. Image 1.8(a) courtesy of University-Clinic for Maxillofacial Surgery, Vienna, Austria; Images 1.8(b), 1.8(c) courtesy of Siemens Corporate Research (SCR), Princeton, USA.

saw. Also the main axes of the instrument is virtually superimposed as a line. According to the authors, this provides "navigational information and virtual imaging of nonvisual structures to the operating surgeon in an unencumbered, less restricting way, without interrupting the flow of the surgery." [236] In addition the authors distinguish that "it is of paramount importance that the surgeon can always quantify system inaccuracy by direct visualization" [236], which has not been realized in the proposed system. The

"... it is of paramount importance that the surgeon can always quantify system inaccuracy by direct visualization" [236].

authors conclude that their method "either reduced or did not increase operation time in general, but it did increase preoperative preparation time". Authors claim that the usage of their system allows "surgery to become less invasive". In addition, "surgery using virtual image guidance was believed to provide additional safety compared with surgery performed without it" [236]. Although the important step into the operating environment has been accomplished to start with the clinical evaluation of the proposed system, there is only a small number of reported clinical studies. Further studies have to be performed to confirm their first results. Freysinger et al. [77] discuss the usage of the system in combination with a teleconference system. They further address practical issues such as the intraoperative setup of the system, preparing the infrastructure and clinical staff, synchronization of the systems and backup versions. However, the system presented in 2002 [77] does not necessarily involve an HMD to present the AR scene.

Bockholt et al. [35] present a monocular video-see through HMD based AR system to be used in endoscopic and robotic surgery to "bridge the gap between the external planning session and the intra-operative case". Planning information shall be superimposed onto the patient's anatomy. Authors present several examples of how to apply the system, for instance augmented pedicle screws to be implanted into vertebrae or superimposed 2D slice and 3D volume rendering of the head. In order to keep hands free, they propose interaction with the AR scene using voice and pedals.

In 2000, Sauer et al. [194] propose a stereo video see-through HMD equipped with a third camera for head tracking. The system has been used for most parts of the work presented with this thesis and is described in detail in section 2.1. The authors technically describe the system and highlight the importance of exposing the very best version of the system to the potential end user to get reasonable feedback regarding promising applications. They also discuss their choice of a video see-through device and the problem of "confused spatial perception" when the depth cue occlusion is interfering with stereo vision. Maurer et al. [144] target planning and navigation in image-guided cranial neurosurgery. Authors report first experiences with the system regarding the perception of virtual anatomy superimposed onto a head phantom. The head phantom based on real CT data is registered with further pathological CT and MRI data. Video data is augmented with segmented virtual brain structures from such data. Authors mention that ideally one would create a "head phantom manufactured from the patient's MR or CT image" showing the pathology of interest. They state that the relative position and depth of virtual wire frame models and texture-mapped point clouds can be better perceived than polygonal, shaded 3D models. In particular depth perception suffers when closed structures in the video image, e.g. the ear or the eye, interfere with virtual structures. In addition to the anatomy, also a virtual counterpart of the needle has been rendered.

In order to start with clinical studies, their system was further modified to be used in a neurosurgical iMRI operating room. The authors introduce the term "in-situ visualization" which means that "anatomical structures are being displayed at the location where they actually are" [189]. The authors further analyze the effects of the system setup regarding the user's perception of the scene. They find that users can adapt to the offset between the optical centers of the cameras and their own eyes. The synchronized presentation of virtuality and reality is determined as one of the "most critical feature" [189]. Regarding the visualization of virtual objects, authors further suggest to "show segmented structures not as solids, but as wire frames, not with thick lines but with thin lines, not opaque but semitransparent, not finely structured but coarsely structured where details are not important" [189].

In [190] the system is extended to augment US slices similar to the system proposed by Bajura et al. [15]. Sauer introduces further interactive features to control entities of the AR scene. The user can target a structure of interest in the US image with his line of sight serving as a "pointer" [190] and position a wire frame sphere serving as a 3D marker by pressing the food pedal. When positioning the probe in a way so that the vertical line of the US image is close to the target structure in the image, the user can initiate an algorithm that automatically detects the structure. The user can further freeze several US images and show them simultaneously in-situ. An interactively positioned "virtual whole" [15], "which the user can drop in place at the current transducer location" [190] is proposed to enhance depth perception. Sauer et al. further adapted the system to the requirements of the surgical environment in order to smoothly transfer the HMD to the neurosurgical iMRI operating room [191]. These needs include the direct access of imaging data from the imaging device as well as a sterile setup of tracking targets. In addition, they performed three experiments to investigate the appropriate visualization of target anatomy and instruments. As phantoms, they prepared a cantaloupe with vitamin E capsules, a melon with vitamin E capsules and radiographic markers and a box filled with partially contrasted (Gadolinium) mashed potatoes. The authors distinguish the importance of also augmenting the instrument to be navigated to deep seated regions. When the instrument, in their case a biopsy needle, is virtually augmented and its path is extrapolated the user can "aim the needle towards a target" object [191]. In addition, authors tested different methods for visualizing the MRI data, i.e. as orthogonal slices and/or 3D wire frame models of segmented structures. In [192] they report an experiment with a needle placement phantom done by over a hundred subjects. Subjects successfully guided a needle through an invisible space to a target (Ø 6mm) without previous training time. In addition to the previously proposed visualization of the augmentation and extrapolation of the needle using a wire frame model, the targets are rendered with shaded 3D Tori and change their color when the trajectory intersects with the target area (see Fig. 1.8(b)). This visual feedback can support intuitive guidance of the needle. The authors suggest to use AR guidance "when the user encounters complex anatomy, where vital structures like nerves or blood vessels have to be avoided while the needle is advanced towards a target like a tumor." [192] In 2003 Sauer et al. [117] report the continuation of their work on

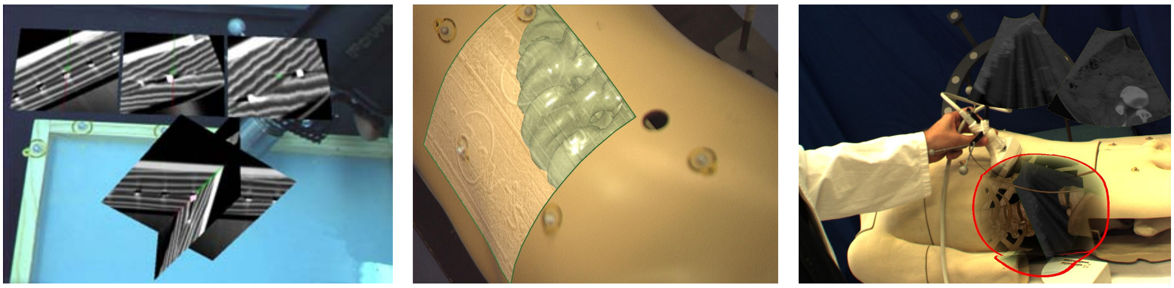
> **AR guidance may be "especially helpful when the user encounters complex anatomy, where vital structures like nerves or blood vessels have to be avoided while the needle is advanced towards a target like a tumor." [192]**

US augmentation and introduce an AR aided clinical procedure to interactively acquire, reconstruct and visualize a 3D US volume for needle biopsy. Gaze based interaction is used to position a virtual bounding box serving as a reference frame to entirely scan the region of interest and to acquire a high resolution volume without artifacts. Through a phantom study they show that also novices can easily accomplish the biopsy by following the proposed procedure. The system has been extended to also use CT data to support needle placement. Using an interventional 3D abdominal phantom that contains simulated liver lesions, different error sources have been analyzed that may affect the guidance of the virtually visible needle to wire frame models of the lesion within CT data. These errors were categorized as the "virtual placement error" indicating how well the user followed virtual navigation information, the "real placement error" indicating the final accuracy of targeting and the "AR system error" representing the accuracy of superimposing the virtual needle onto the real one [193]. Authors conclude that even tough the bending of the needle "introduces uncontrollable errors" the system can provide sufficient accuracy for the clinical task. The authors suggest that their system can be used as an alternative or a complement to CT flouro guidance and hope to reduce radiation. In [229, 233] several phantom trails in an interventional setup as well as a first animal study for MRI guided needle placement are presented. They extended their guidances visualization by a concentric ring around the target spot. The radius of the ring changes proportional with the distance of the needle tip to the target spot "to gain a better depth perception along the line-of-sight" [229]. In addition to the target spots shown as solidly shaded tori, a DICOM slice containing the target region has been coded with color and added to the scene. In contrast to their phantom studies showing an average error of 1mm, in the animal study they missed the target spot by 1cm, which may also be caused by the breathing motion of the pig. Authors note that "respiratory motion presents a challenge that has to be appropriately dealt with to make the AR guidance clinically practical and useful" [229]. In 2006 [51], the same group reported a follow-up user study to verify the results on CT guided needle placement [193]. They propose the same visualization of needle and target lesion and added cross sectional slices of the CT data to the AR scene. A "white grid superimposed on the outer perimeter of the interventional phantom provides additional depth cues" [51].

Based on the stereo, video see-through HMD called RAMP system of Sauer et al. [194], the group of Nassir Navab to which the author belongs began to develop further solutions for different medical applications. The research was focused on the used of in-situ visualization for medical training and simulation and for navigation in particular in trauma surgery [97]. In 2004, Sielhorst et al. [205, 204, 201] use AR visualization in conjunction with a gynecologist simulator system to train physicians to deal with patients with complicated deliveries. The augmented simulator comprises haptic and auditory feedback in combination with in-situ visualization of the baby and the mother's pelvis and offers "a look inside to gain insight" [204]. When using the AR system, "the physician now concentrates on

In-situ visualization applied to training and simulation can provide "a look inside to gain insight" [204]

the vaginal delivery rather than the remote computer screen" [204]. The authors further propose to reformulate the visualized performance of an expert while the student is training with the simulator. For this reason, the tracking data of the performing expert is recorded and then synchronized with the performance of the apprentice. This allows also for an after action review (AAR) to let the trainee evaluate his own performance. In [201] the authors point out the importance of immediate comments from an expert while supervising the performing trainee. The author mention the need for better input devices to interact with the AR scene since "a mouse is not an adequate pointing device for augmented reality" [204]. Supported by a series of cadaver studies, Traub et al. [222] and Heining et al. [97] proposed to apply the AR system to hip repositioning and intramedullary and medullary nail locking of distal nails. Even though surgeons who performed the experiments favored the AR system over conventional navigation systems, the positioning tasks using AR guidance failed. According to the authors, the inappropriate visualization of navigation information caused the failure of the system. They state "that a volume rendered or CT axes aligned slice rendering based image overlay is not sufficient for execution of navigated tasks" [219]. Consequentially, Traub et al. [219, 221] developed and evaluated advanced visualization modes. According to the authors, their most promising navigation interface combines three orthogonal slice views lined up close to the field of action within the AR scene into the so-called "instrument aligned orthoslice view". The latter component shows two orthogonal slices registered with a surgical drill whereas the intersection line of the slices corresponds to the drill axes (see Fig. 1.9(a)). For controlling drill depth and orientation, the surgeon has to combine information from several sources. Side views on "the instrument aligned orthoslice view" (normals of slices are orthogonal to the line of sight when looking along the drill axes) communicate drill depth and the line-up of orthogonal slice views are used to control drill direction. Sielhorst et al. [200, 199] evaluate



(a) Advanced visualization modes for AR aided surgical drilling [219].

(b) Interactively positionable virtual window for improved depth perception [200]

(c) US simulation incorporating After Action Reviews (AAR) [33]

Figure 1.9: Visualization methods for image guided surgery and medical training of the group around Nassir Navab. Images 1.9(a), 1.9(c) courtesy of Jörg Traub and Tobias Blum, Chair for Computer Aided Medical Procedures & Augmented Reality, TU München, Germany

different visualization methods that had been proposed in the literature regarding their impact for correct depth perception when visualizing anatomy in-situ. For the evaluation, a virtual spinal column had been superimposed onto a thorax phantom to simulate minimally invasive spine surgery. The spinal column has been presented opaquely, transparent, as a wire frame model and behind a virtual window registered with the phantom's skin similar to the "virtual hole" introduced by [15] (see Fig. 1.9(b)). The virtual window could be positioned interactively using the intersection point of user's line of sight with the skin surface as a reference. The virtual window and its comparison with earlier approaches to improve depth perception in medical AR scenes [200] are results of the author's diploma thesis [19]. The outcome of this preceding diploma thesis encouraged the group around Nassir Navab to advance the approaches of medical in-situ visualization and motivated the development of the new methods for contextual in-situ visualization presented in chapter 3. Blum et al. [33] integrated US simulation into the system. US images are generated from the analysis of the intensity values from CT data of a real human body. They use a phantom [24] that was built from CT data of the Visible Korean Human Project [164]. Further details on the phantom are given in section 2.2.2. Beside the US simulation, the proposed AR system incorporates the training feature After Action Review (AAR) that was previously introduced for a birth simulator [204, 201]. Following the FAST protocol, which is a standardized US diagnoses protocol, the trainee guides the US probe along the path proposed by an expert. Both, the expert and the trainee path, can be synchronized and reviewed.

For further details on projects of the group around Prof. Dr. Navab, the reader is referred to the PHD theses of Traub [219] and Sielhorst [199].

1.2.2.2 Augmented Magnifying Optics and AR Windows

According to Rolland [182] a stereo operating microscope "can be said to be equivalent to an optical see-through system operating on a microscopic scale". However, operating microscopes are usually mounted on a swivel arm, which reduces the freedom of movement of the observer and consequentially the range of views of the AR scene. In addition, the AR views typically show only a small (microscopic) portion of the patient's anatomy of interest. The target region is usually not presented within the context of exterior anatomy, for instance the patient's skin. In [119] King et al. report their work on a stereo optical see-through AR system that is based on a standard operating microscope. In 2000 they presented the results of a set of studies on volunteers, phantoms and also seven patients [118]. They conclude that the development of adequate visualization methods is one of the key tasks to make their system useful and acceptable in the surgical environment. Initially the group identified the stereo view on the AR scene and the shading of virtual objects as crucial to provide enough visual cues for correct depth perception. However, along with further trails and experiments they found out that the simple superimposition of virtual objects behind the transparent real tissue cannot provide the expected result for a satisfying depth perception of all objects in the AR scene. They state that using visualization techniques taking the advantage of the depth cue motion parallax is not helpful with their system called *MAGI* since "the operating microscope is usually left in the same position for significant periods of time" [118]. King notes that the shaded surface based data representation looks realistic, however they rather proposed virtual structures mapped with "locally unique spatial frequency together with careful control of illumination" to create the helpful cues for correct depth perception. It is mentioned that a wireframe representation of such data can approximate such texture mapping. Johnson [108] further investigated the problems of the optical see-through operating microscope causing misleading depth perception. He mentions that the simple superimposition of semi-transparent virtual structures onto real objects "diminishes the surgeon's ability to use the stereo information effectively and get an accurate sense of depth". It is speculated that the disparity of a stereo AR system is not strong enough "to overwrite what the observer knows to be true" namely that the observer has not the physical ability to see into the patient. Johnson concludes [108] that one may achieve "less confusing or ambiguous" compositions of real and virtual objects, when manipulating "physical surface" exploiting the features of "a video based system". As a further implicit problem of the operating microscope affecting the perception is the limited FOV. This prevents the observer from using visual but also proprioceptive cues to perceive "the object of interest in relation to the ground, the horizon or other familiar structures" that help to locate the object of interest in a scenery. In 2004 the same group reported on their "'Clinical Experience and Perception in Stereo Augmented Reality Surgical Navigation"'[56]. At that time the MAGI system was tested with 17 clinical cases wherein in three operations, namely vestibular schwannoma, ethmoid meningioma and adenocarcienoma, the clinical outcome had been improved. They also presented the result of an experiment evaluating the perception of an object with and without the presence of a "physical surface" that simulated the real occluding tissue in the foreground of the medical AR scene. The result of that experiment showed that the estimation error increased and the target object was perceived more distant when the surface was present.

In 1986 Hatch et al. [180] introduced an "reference-display system" allowing for superimposition of radiographic images onto the real anatomy. The reference-display is built upon an operating microscope for stereotaxic procedures in neurosurgery. The authors propose a three dimensional ultrasonic digitizer system to track the patient relative to the reference-display. They determine the approximate error of the system as 1mm, which has been corrected in [180] to 2mm (AVG), however, ranging from 0.7 to 6mm. Authors propose "easily visualized contours" of the tumor extracted from imaging data to not diminish "the quality of the conventional operative field image" [180]. Their clinical evaluation of the system since 1983 has been driven by the following strategy. Operative cases have been selected for their clinical studies when "CT imaged structures" are "identifiable at surgery such that an estimate of the system accuracy would be possible". Second, only cases were taken into account "in which the operative procedure would not be dependent upon the system for its desired and safe completion" [78]. Even though repositioning of the operating microscope and consequential new registration of real and virtual entities to up to one minute, the "greatest clinical utility of the system appears to be in providing accurate and reliable guidance to the surgeon in his operative approach to small lesions" [78]. They mentioned that the "determination of the extent of tumor resection is a more recent but equally attractive application" [78].

Paul proposes to augment the AR scene presented through a surgical microscope with "2-D contours from the intersection of preoperative 3-D multimodal images with the surgical microscope focal plane" [165].

Beside augmented magnifying optics such as operating microscopes or operating binoculars and see-through head mounted displays there is a third category of display technology that allows for in-situ visualization from the natural point of view. AR windows positioned in between the surgeons head and the patient have been introduced for example by [32, 141, 133, 73, 69, 68]. On one hand, AR windows avoid a cumbersome device to worn on the user's head. On the other hand perspectives are limited since the window has to be placed within the field of view. In addition, the physical device in the field of view may restrict the freedom of motion [202]. The author refers the reader to the survey "Advanced Medical Displays: A Literature Review of Augmented Reality" of Sielhorst et al. [202] for further information. This document focuses on head-worn displays.

CHAPTER

TWO

DESIGN AND DEVELOPMENT OF CONTEXTUAL IN-SITU VISUALIZATION

The work presented with this dissertation is based on video see-through head worn displays being the primary component of an AR enabling 3D user interface. Intensive preparatory research on hardware, tracking [191, 193], registration, visualization and synchronization [199] builds the basis for the development of the visualization and interaction methods introduced in chapters 3 and 4. Section 2.1 starts with the introduction of the complete AR system setup in section 2.1. Section 2.2.2 introduces a new genre of manikins adapted to the needs for developing visualization methods and navigation systems to be used for AR assisted surgery. Correct and intuitive perception of the AR scene is crucial to let the AR technology become acceptable and useful. Section 2.3 gives an overview of human visual perception serving as a base knowledge to develop sophisticated visualization methods for medical AR systems.

2.1 The Used AR System

The AR system that was used for the development of visualization and interaction methods is based on optical tracking and a video see-through head mounted display (HMD). In this setup, tracking is limited to infrared (IR) camera systems and divided into two system types. The first one is an outside-in tracking system that locates experimental phantoms or later the patient as well as surgical instruments. The second type is an inside-out tracking system that tracks the HMD worn by the observer. The major reason for this division is the required accuracy in medical applications when registering virtual objects of the AR scene with real ones. Fig. 2.1 gives a schematic overview of the AR system to be used for intra operative applications.

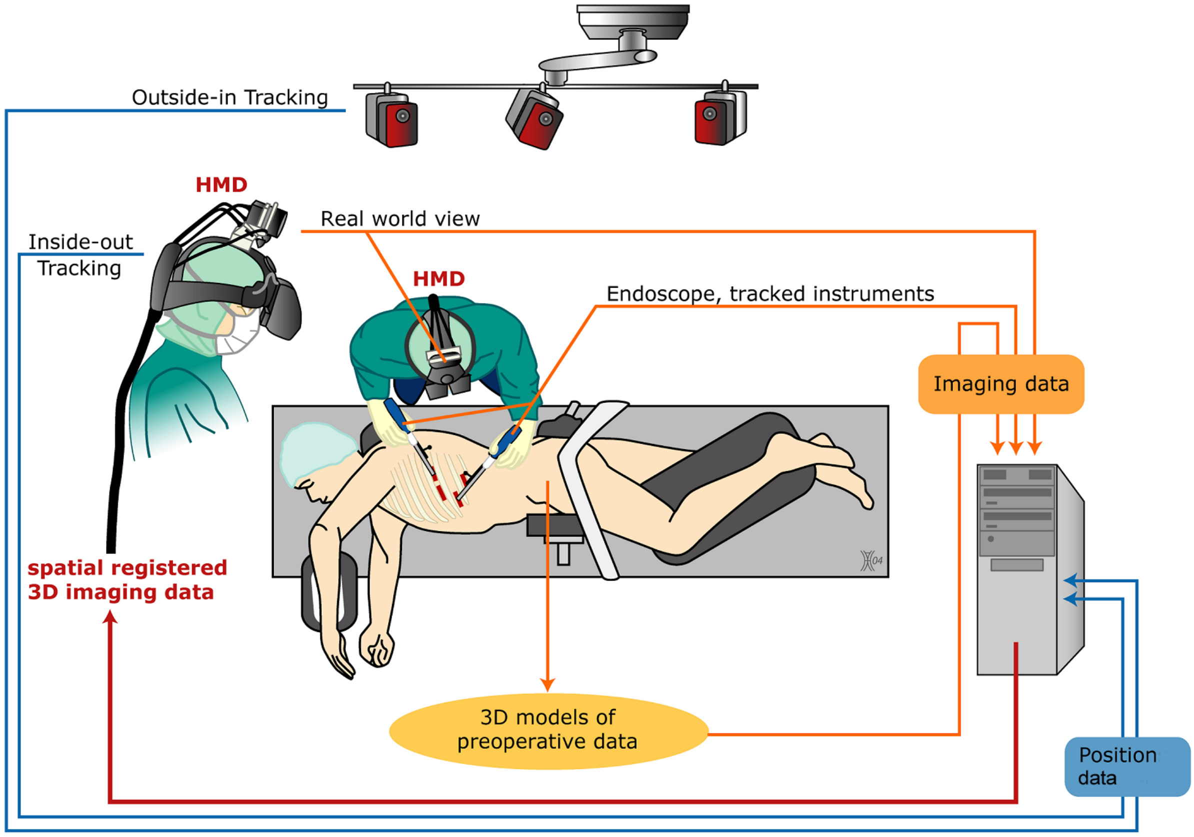


Figure 2.1: Schematic overview of the AR system including an external, optical tracking system and a stereoscopic, video see-through head mounted display with a single camera inside-out tracking system.

2.1.1 Outside-In Tracking

For robust and accurate tracking of virtually extended objects within the surgeon's augmented workspace, an external optical outside-in tracking system [1] is used that consists of four infrared cameras fixed to the ceiling covering a large working area ($3 * 3 * 2\ m^3$). The term *outside-in tracking* describes the setup of an external, optical tracking system, wherein tracking cameras are installed rigidly to cover a larger tracking volume. The tracking targets are equipped with marker sets that can be optically detected and tracked by the tracking cameras. Such markers can either passively or actively send out the signal to the cameras. Active markers are usually infrared light emitting elements (LEDs), while the cameras are synchronized with the light emission. In the case of the tracking configuration being used for the AR system, the tracking cameras *ARTtrack2* are equipped with an infrared flash that illuminates retro-reflective marker sets, i.e. passive markers, attached to the tracked objects. The reflection is then detected by the camera sensors and their optical center is computed to determine their position in space. To recover the six degrees of freedom of a rigid body, the external optical tracking system requires at

[1] A.R.T GmbH, Weilheim, Germany

least three rigidly attached markers. More than three markers helps to deal with partial occlusion of the tracking target due to user interaction and results in more robust and accurate tracking. In most medical applications, however, one would prefer a minimal number of markers since space is usually scarce and expensive. The flash and camera sensors can be synchronized in order to avoid active sensors being flashed by another camera positioned in front of the first one.

ARTtrack2 cameras can be attached as hanging or standing to any construction with one screw coupling. *A.R.T* also provides a T-junction to install two cameras abreast, which then can be attached to a tripod or any other construction. According to the author's experience the wide range of installing options was extremely helpful to achieve the optimal arrangement of cameras in the lab, in the trauma room, or other locations. As well as *ARTtrack2*, which was installed in our setups as either two or four camera sets, *TrackPack* has also been used. Table 2.1 provides the specification for both tracking systems. After an initial calibration procedure the outside-in tracking system provides

Table 2.1: Comparison of the two outside-in tracking systems being used for development and experiments.

	ARTtrack2	**TrackPack**
max working distance	up to 4 meters	2.5 m (for 12mm markers)
max frame rate	60 Hz	up to 60Hz
FOV (h,v) - focal length	72.8°, 58.2° - 3.5 mm	93.5°, 77.2° - 2.6 mm
	57.9°, 45.3° - 4.5 mm	72.8°, 58.2° - 3.5 mm
	42.9°, 33.0° - 6.0 mm	
	33.9°, 25.8° - 8.0 mm	
	22.2°, 16.8° - 12.0 mm	
suggested working volume		2.0 m x 1.5 m x 1.5 m
Delay		< 20ms
max # 6DOF targets	20 (fast) or 2 (accurate)	up to 4 "6DOF"
Flash energy	8 steps	8 steps
Data output	100 MBits	Firewire A (6 pin)
weight	0.96 kg	0.41 kg
dimension	109 mm x 78 mm x 120 mm	76.9 mm x 60.0 mm x 77.8 mm
IR LED flash	880nm	850nm
Noise	built-in fan	Noiseless (no fan)

accurate tracking as long as the tracking cameras remain at their position. In the optimal case, the system is capable of tracking the targets with high accuracy due to the benefit of optical tracking. The error sources that potentially diminish the accuracy of the systems have been investigated by Sielhorst [199].

2.1.2 Video See-Through Head Mounted Display and Inside-Out Tracking

The core component of the AR system is a video see-through head mounted display (HMD) allowing for stereo vision. Two instances of the system have been used for development and experiments, which are the *RAMP system* developed by Sauer et al. [191, 193] shown in Fig. 2.2(a) and an advanced system developed at the *Chair for Computer Aided Medical Procedures & Augmented Reality, TU München* [199]. The latter HMD system will be called from now on *CAMP system* (see Fig. 2.2(b)). Table 2.2 provides a comparative overview of both HMDs. In order to allow the user to see through the head worn display,

Table 2.2: Comparison of the two video see-through head mounted displays being used for development and experiments.

	RAMP System	**CAMP System**
HMD	1024x768 px, 2.34 arcmin/px	NVIS NVisor SX
HMD resolution	1024x768 px, 2.34 arcmin/px	1280x1024 px, 2.25 arcmin/px
HMD contrast ratio	40:1	200:1
HMD color range		24-bit
HMD FOV (h,v)	40°, 30°	48°, 36°
Color camera		PTGrey Flea
Color camera resolution	640x480 px (interl., analog)	1024x768 px (progr., digital)
Color camera update rate	30 Hz	30 Hz
Lens	M-Mount 8mm	C/CS Mount 8mm
Tracking camera		PTGrey Flea
Tracking camera resolution		640x480 px
System Weight	$\tilde{1}$ kg	$\tilde{1}$ kg
Max. system framerate	30 Hz	30 Hz
System Latency	$\tilde{1}00$ ms	100 ms

two color cameras rigidly attached to the HMD simulate the eye's view. Even though the optical center is shifted about 6 cm (depending on the individual user's anatomy) above the real position of the user's eyes, no negative user feedback during experiments has been recorded. From the author's experience, subjects rather complain about the weight of the HMD in particular when the HMD is requested to be worn for long lasting experiments.

For superior registration quality, both HMDs are equipped with an inside-out tracking system, which is synchronized with the outside-in tracking system described before. Inside-out tracking allows for a high rotational precision [99] when tracking the HMD (see Fig. 2.2). If markers were attached to the HMD and tracked by the outside-in tracking system, small head rotation having a small effect on the CCD (Charge-coupled Device) chips of the tracking cameras would result in a comparably larger error with increasing distance of the tracking target. The inside-out tracking system uses a single camera, which is mounted on top of the HMD and oriented along the line of sight of the observer. The camera is equipped with a fish eye lens to cover a big field of view for more robust tracking. It is sensitive to infrared light only. Infrared flash LEDs surround the camera to illuminate the retro-reflective passive markers described in section 2.1.1. The camera detects and tracks a special tracking target serving as a reference frame that consists of a rigidly attached set of nine round, flat markers [191, 193, 198, 199] (see Fig. 3.2). Since the reference frame is positioned close to the real objects to be superimposed by virtual ones, for instance the patient augmented with his/her virtual anatomy, small effects on the CCD chip of the inside-out tracking camera due to head reorientation can be detected and used to realign the augmentation. The reference frame enables the transition between the inside-out and the outside-in tracking systems since both tracking systems use the same coordinate system as the reference frame.

Inside-out tracking provides high rotational accuracy, however, for high precision in depth, along the user's line of sight, one would prefer the outside-in tracking system. As a preliminary result the combination of inside-out and outside-in tracking has shown increased acceptance and usability of the system. With the standard configuration of the system the outside-in tracking locates robustly the tracking targets, however, virtual objects are present in the AR scene when the inside-out tracking locates the reference frame, i.e. it stays in the field of view of the head mounted tracking camera. This configuration limits the freedom of movement of the observer and causes flickering of the virtual part of the scene when hands or instruments occlude the tracking camera's view on the reference frame. For a user study described in section 5.3 infrared markers were attached to the HMD enabling outside-in tracking to be used as a backup system for head pose estimation. The backup system, however, only gets active when the inside-out tracking system fails. A simple red/green dot on the display informs the user about current accuracy level of the system, i.e. either rotational or depth accuracy. With this configuration, users can even walk around the body and get extreme perspectives on the augmentation without tiring flickering.

In general there are two types of see-through HMDs used in the AR community. Both, optical and video see-through HMDs, have their advantages and disadvantages as discussed by State et al. [208], Sauer et al. [194], Rolland and Fuchs [182] and Sielhorst et al. [202]. Regarding the targeted medical AR environment, flexible image composition and high accuracy are, however, the major arguments for video-see through technology.

In optical see-through HMDs virtual objects are projected onto a partial transparent layer placed in front of the user's eyes. The real environment is directly perceived by the user. However, with today's computer hardware the virtual part of the AR scene cannot be rendered with the same resolution and color depth as reality. This results in a visible discrepancy between reality and virtually, which let virtual objects look even more artificial. With video see-through HMDs the image composition of virtual and real objects is based on data from the same image space. i.e. using the same resolution and color range. Even though the hardware limits the quality of the final AR image, this technique allows for a much smoother transition from reality to virtuality. In contrast to optical see-through technology being able to present only half transparent virtual objects, in video see-through HMDs, both real and virtual objects are digitally accessible, which allows for much more flexible methods of image composition. The accuracy of registering virtual objects with real ones is strongly affected by the nature of optical see-through systems. First, virtual data can only be presented with a certain delay, which strongly depends on the power of the hardware used and complexity of virtual objects. This causes floating virtual objects to frequently be not precisely registered with real ones. In addition, the time lag would lead to undesirable and for the user fatiguing effects like "perceivable jitter or swimming" [193]. Second, a rigid eye-to-HMD transformation cannot be guaranteed, which is fundamental for reaching the required registration accuracy in medical applications. In video see-through HMDs reality and virtuality can be synchronized in order to guarantee accurate registered objects at any point in time. Even-though there is a delay regarding the user's motion and what he actually sees, the accuracy has a much higher priority than reactivity on fast motion in the targeted application domain. When considering the camera as the eye, the eye-to-HMD transformation stays rigid after a one-time calibration procedure [199]. Further evaluation and description about different types of HMDs can be obtained at [207, 151, 150].

(a) RAMP system (b) CAMP system

Figure 2.2: The video see-through head mounted displays consist of two color cameras simulating the eye's view and an infrared camera for inside-out tracking.

2.1.3 Combination of Components

The AR software framework CAMPAR provides methods for synchronization, registration, image processing and visualization of medical volume data such as CT or MRI using different rendering techniques [203]. The framework is based on a plugin structure that can be iteratively extended. The AR systems uses a PC based computer to render 3D graphics, to compute and include tracking data, to synchronize and combine imagery of virtual and real entities. The specification is Intel Xeon(TM), CPU 3,20 GHz, 1,80 GB RAM, NVIDIA Quadro FX 3400/4400. Computer graphic objects are implemented in C++ using OpenGL [2], Cg [3] and GLSL [4]. All augmentations on targets, which are tracked by the optical outside-in tracking system, have to be positioned respectively to

[2]Open Graphics Library: http://www.opengl.org

[3]C for Graphics is a high-level shading language developed by NVIDIA: http://developer.nvidia.com/page/cg_main.html

[4]OpenGL Shading Language: http://www.opengl.org/documentation/glsl/

the reference frame of the inside-out tracking system. The following equation calculates the transformation H^{t}_{ref} from the reference frame to an exemplary target (H^{to}_{from}):

$$H^{t}_{ref} = H^{t}_{ext} * (H^{ref}_{ext})^{-1} \tag{2.1}$$

Transformations H^{t}_{ext} and H^{ref}_{ext} are provided by the external, optical outside-in tracking system. The former describes the transformation respective to the origin of the tracking system to a target, the latter one is the transformation from the origin of the tracking system to the marker frame for inside-out tracking.

The transformation for in-situ visualization can be described by H^{CT}_{ref}.

$$H^{CT}_{ref} = H^{CT}_{phan} * H^{phan}_{ext} * (H^{ref}_{ext})^{-1} \tag{2.2}$$

The optical outside-in tracking system provides the transformations H^{phan}_{ext} and H^{ref}_{ext}. We attached CT markers [5] coated with retro reflective material to the phantoms used in our experiments. Compared to the usual content of CT images showing the patient's anatomy of interest, these markers have a distinguishably high intensity value. For this reason, these markers can be segmented automatically from the volumetric imaging data unless there is further material such as metal screws with similar intensity values within the targeted image space. Correspondence of segmented positions and tracking data results in a registration matrix H^{CT}_{phan} that aligns the data volume with the tracked object. Regarding a real surgical situation, data of an intraoperative CT scan with an optically tracked C-arm device can be registered with the patient and visualized as proposed by Feuerstein et al.[65].

2.2 Test Beds for Try Runs

For the development and evaluation of sophisticated visualization and interaction methods we determined the importance of realistic test environments. Potential problems that may occur in the real future operating environment of the AR system have to be detected in an early developing stage. The objective is to adapt the system preferably to the real conditions than to lab conditions. For this reason, a new genre of manikins has been developed which is strongly adapted to the needs of the development of AR applications in the field of trauma surgery. Beside the realistic patient model an interdisciplinary lab space has been created to attract a sample of end-users, in this case medical students and trauma surgeons for frequent evaluation of the AR system.

Section 2.2.1 gives a short literature review of proposed testbeds for medical AR systems. The novel manikin prepared within the scope of this project is described in section 2.2.2.

2.2.1 Test Beds for Medical Augmented Reality Systems

Realistic test beds for developing and evaluating medical AR systems for interoperative usage are important for progressive research. Experimental arrangements such as cadaver,

[5]Beekley Corporation: http://www.beekley.com

animal or in-vivo studies are expensive to develop and difficult to be justified to showcase and evaluate AR technology. Custom-made phantoms are often presented by the community to simulate a minor anatomic part or conditions. However, this either can produce artificial problems that are not present in real conditions or disguise important issues. Both leads consequently to a non optimal system design and configuration.

It is extremely difficult, maybe even impossible, to access or build a phantom that corresponds to the physical properties of the human anatomy. Different companies like Gaumard, Laerdal or Meti offer manikins that simulate the patient. These manikins are part of many medical curricula to teaching medical emergency teams but also more structured workflows such as birth simulation and anesthesia. However, there is no solution yet on the market that provides a mannequin coming with real medical imaging data sets that are adapted to the needs of AR applications.

Many follow the strategy of using CT or MRI imaging data of unrealistic, custom-made phantoms to design a navigation or visualization system. Quite often, the produced phantoms and corresponding imaging data are capable of evaluating the performance of a certain task, such as accuracy and speed of instrument guidance. However, such imaging data excludes information that can be used in real conditions or includes information that misleads the user and results into conclusions that are not valid in real medical scenarios.

For instance, Sauer et al. [191] use a cantaloupe and a box filled up with mashed potatoes for the evaluation of AR guided needle biopsy. Birkfellner et al. [30] present a phantom skull with jelly that "was covered with polypropylene dust in order to make the jelly opaque" to simulate navigated neurosurgery. Traub et al. [221] built a wooden box for navigated surgical drilling that is equipped with metal spots serving as target spots and filled up with silicon. Sielhorst et al. [200] took a CT scan from a thorax phantom with nothing except a plastic spinal column inside the body to evaluate the quality of depth perception of different visualizations modes. In 2007, our group built a phantom for pedicle screw implantation in spine surgery consisting of replaceable vertebrae [21]. The CT scan of vertebrae surrounded by silicone and peas was then used for the evaluation of an AR navigation system. For further studies, a thorax phantom of the company Aesculap [6] has been employed to develop and evaluate visualization [200, 27] and interaction methods [26, 20]. With the evaluation of a virtual mirror [26, 20] for laparoscope augmentation a complex blood vessel structure has been simulated with wooden branches [25]. Cadaver and animal studies come close to the real conditions in the OR. However, only few researchers having access to interdisciplinary facilities report on such lab and evaluation environments [233, 66, 27] and only few groups report on evaluations on real patients [120, 30, 89]. Unfortunately, such ideal evaluations are only possible and justified at final stages of AR product testing. There is therefore a considerable need for appropriate phantoms for the development and testing of radical AR solutions.

2.2.2 The Visible Korean Human Phantom (VKHP)

The Visible Korean Human Phantom (VKHP) is a new genre of phantoms that provides a realistic environment to develop and evaluate future AR visualization and surgical

[6] B. Braun Melsungen AG: www.aesculap.de

navigation techniques. In this regard, the creation of the VKHP follows the statement of Park et al. [164] that "the VKH has exciting potential applications in the fields of virtual surgery, virtual endoscopy, and virtual cardiopulmonary resuscitation".

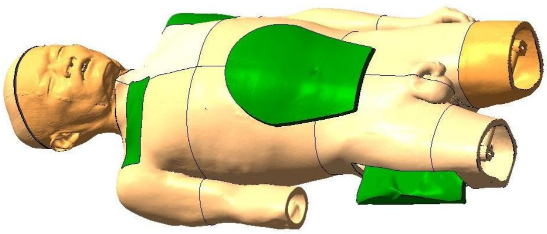
Beside the *Visible Human Project* (Caucasian) (VHP) and the *Chinese Visible Human* (CVH) [250], the *Visible Korean Human* (VKH) [164] project provides full body imaging data consisting of MRI, CT and a photographic data set.

The VKHP is created from the CT data set of the VKH and five removable windows were fabricated into the phantom's skin (see Fig. 2.3(a), 2.3(b)). This includes a 30x10 cm approach at the back of the phantom for the simulation of dorsal instrumented spine operations, for instance vertebroplasty, kyphoplasty or pedicular screw implantation and fixation as shown in figures 2.3(a) and 2.3(b). In addition, windows were integrated at the proximal shoulder and the proximal femur for internal osteosyntheses procedures, for instance plate and screw fixation. At the right clavicular region a 10x3 cm approach was build in to allow intramedullary clavicular fixation as well as intrathoracal drain application. Such operation sites can be used for training surgeons in standard procedures within an almost realistic setting and developing new solutions for AR based image guided surgery in different surgical disciplines.

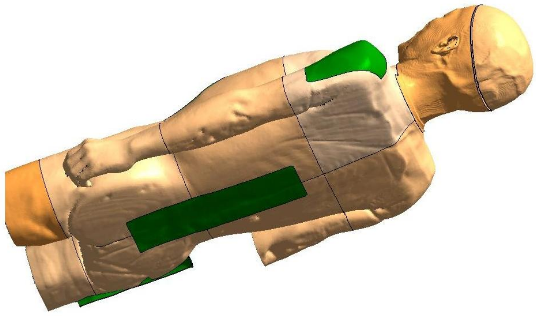
The VKHP consists of the skin and bone structure from the CT scan, which is suitable for providing a testbed for different procedures in navigated hip, shoulder, spine and neurosurgery. First, CT data of the VKH was prepared using the software 3-Matic and Mimics of Materialise [7] to create surface models of the anatomy to be printed in 3D. The required surface model format is STL [8] to be applicable for rapid prototype 3D printers. The virtual phantom was divided into components of maximum size of 700mm x 380mm x 590mm for the printing machine. In general, there are machines on the market that provide bigger print volumes. However, the components were prepared to be removed and plugged together for being able to adjust the VKHP according to a particular application, for sharing parts of it among different lab spaces and to be able to replace single parts that may be destroyed during evaluations.

Our true-scale *Visible Korean Human Phantom* (VKHP) (see Fig. 2.4) ranging from the head to the hip was printed from the prepared STL data (see Fig. 2.3(a), 2.3(b)) with additive rapid manufacturing technique, *selective laser sintering*, by the company FIT [9]. The VKHP has been printed in layers (150 μm steps). Depending on the quality of the STL surface data a theoretic level of detail of 0.7mm can be achieved. To enhance the realism of the skin, the phantom was coated with skin colored powder, which also avoids non realistic specular highlights on the raw material. For high quality registration of the virtual CT and MRI data with the VKHP, we attached Beekly [10] CT markers to the skin of the VKHP. The CT markers are coated with an infrared reflective foil in order to make them visible for our infrared camera tracking system as described in section 2.1.1. The original VKH data set was acquired without the mentioned CT spots or comparable artificial landmarks that can be used for registration. For this reason, a high resolution scan

[7] Materialise Group, Leuven, Belgium
[8] Standard Triangulation Language
[9] FIT Fruth Innovative Technologien GmbH, Parsberg, Germany
[10] Beekly Inc.



(a) VKH data prepared in STL format.



(b) A set of windows (green) to operation sites have been prepared.

Figure 2.3: The Visible Korean Human Phantom (VKHP) prepared with different operation sites for future experiments related to different procedures in navigated hip, shoulder, spine and neurosurgery.

from the two separated anatomic regions of the VKHP mentioned above was taken. After the scan, two rigid registration transformations are computed. The first transformation is computed using the point correspondences between the CT spots segmented from the CT scan of the VKHP and the same spots detected by the tracking cameras. The second registration transformation between the original VKH CT data and the new VKHP CT data is determined using non-linear intensity based registration. Mutual information is selected as a similarity measure and a best neighbor optimizer was used [91]. For practical reasons, the VKH data set was divided into two separate tracked anatomic regions, the thorax/hip region and the head region. This allows us to share the VKHP for parallel work. For this reason the registration procedure has been conducted for both regions separately.

The interior of the VKHP is equipped with the bone structure printed from the VKH CT data set. This allows us for simulating different procedures related to spine, hip and shoulder surgery. However, it does not provide a testbed for soft tissue applications such as tumor resection in different organic regions. There are rapid prototyping 3D printers on the market, which are capable of printing all kind of materials with varying physical properties. Since the VKH data includes a set of segmented data volumes showing the main organs such as heart, colon, lung and liver a second generation of the VKHP can incorporate also soft tissue organs that can also be removed when not needed or replaced when being damaged due to experiments.

In order to encourage other research groups to use the VKHP, the STL data, which is ready to be printed, can be obtained by contacting the author of this thesis.

2.3 Perception in AR Environments

One of the objectives of this work is to find more sophisticated methods to compose real and virtual objects of an AR scene. First, the observer has to correctly perceive the positions, order and shapes of objects in the field of view. Second, when exploring the

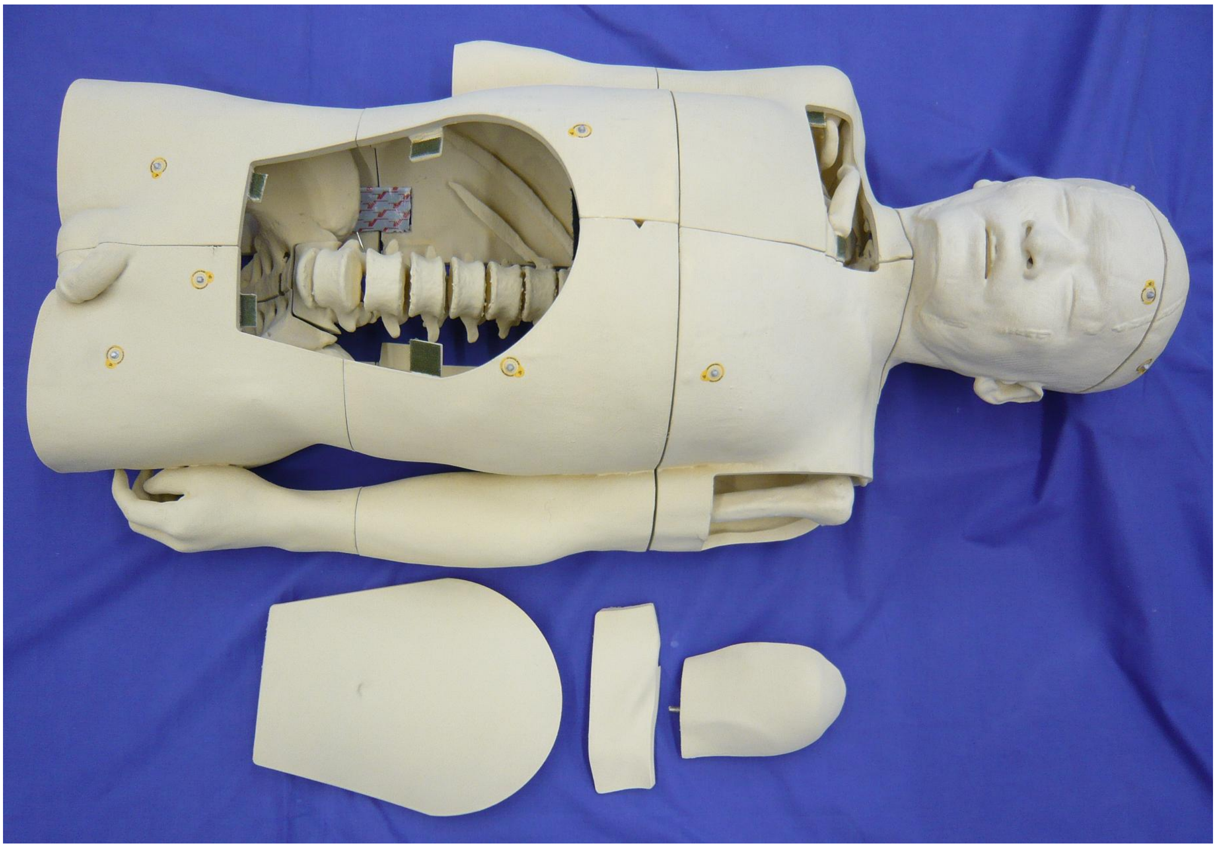


Figure 2.4: The printed Visible Korean Human Phantom

AR world, interaction has to provide visual feedback to enhance the perceptive quality. This chapter starts with the problem statement in section 2.3.1. Then the stages of the human visual perception pipeline will be analyzed in section 2.3.2. Section 2.3.3 gives an overview of depth cues that are used to perceive distances to and between objects within our field of view. The chapter finishes with section 2.3.4 that summarizes visual cues to perceive shapes of three dimensional objects.

2.3.1 Misleading Depth Perception in AR Environments

A popular application of AR systems is to use the augmented view in order to explore the visually hidden interior of real objects such as the internal components of printing machines or underground pipeline systems. However, all technical approaches having been proposed by the AR system to achieve this see-through effect share one problem that prevents the technology from being fully accepted and useful in real life applications. This undesired visual effect is stated by the AR community as the problem of misleading depth perception of virtual and real objects within AR worlds, which becomes particularly obvious, when virtual objects are physically positioned behind real ones [82, 13, 200, 110, 111, 209, 128, 27, 212, 11, 131, 157, 209, 191, 66, 15, 79, 239].

The widely used approach of simply superimposing deep seated virtual entities onto real ones fails to enable correct perception of relative and absolute positions of all involved objects. When the observer moves within the AR scene, a virtual object rigidly fixed to a position inside a real object would even seem to float above the covering object. In most cases, the observer actually knows about the real position of objects due to the nature of objects in the viewed scene. For example the spinal column shown in figure 2.5(b) can be easily judged to be located below the skin and the perception can be corrected. However, intuitive understanding of the scene topology gets lost. In particular, this becomes a problem when the user starts to interact with the AR scene, for example when he has to guide a surgical instrument to the spinal column.

Figures 2.5 show two medical examples when the body is simply augmented with its interior anatomy. The virtual spinal column shown in figure 2.5(b) as well as the subject's virtual brain shown in figure 2.5(a), both extracted from CT data, appear to be located outside the body even though they are spatially correctly registered with the real anatomy.

(a) Augmented parts of the brain from CT data (b) Augmented spinal column from CT data

Figure 2.5: Virtual anatomy, for instance the brain or the spinal column, is presented superimposed on the patient's head and on a thorax phantom, both seem to be located in front of the patient.

2.3.2 Stages of Human Visual Perception

According to Zimbardo [170], vision is the most complex, highest-developed human sense. It enables orientation and supports the equilibrium. Without vision, i.e. if we close our eyes, we loose aplomb immediately [170]. Due to its importance, vision is also the most investigated sense.

The process of human vision can be subclassified into three stages; the *sensation*, the *organization* and the *identification* or *classification* [170]. Fig. 2.5(b) shows an exemplary medical AR scene that will be used to identify the reason for the perceptive conflict within process of human perception.

1. Sensation: Light is reflected from objects of the 3D world, the so-called distal stimuli, within the observer's field of view and enters the eye. The light is then projected onto the retina and becomes the two-dimensional so-called proximal stimuli. In this stage, the proximal stimuli is nothing more than a huge amount of independent sensor information measured by the photo receptors of the retina, which encode the physical energy to neural information. Regarding the AR scene shown in figure 2.5(b), the cameras attached to the HMD build an intermediate step to sensate the distal stimulus. A preliminary "proximal stimulus" of the reality is the result of the projection of light onto the CCD chip of the cameras. At the same time the frame buffer of the computer receives the distal stimuli of the virtual world, for example the rendered spinal column, and the pixels capture the proximal stimulus. The proximal stimulus from the camera is superimposed by one of the frame buffers and creates a new augmented proximal stimulus rendered again to the frame buffer, which is presented on the head mounted displays. The images on the light emitting displays showing the augmented distal stimulus are then projected onto the user's retina and become the final augmented proximal stimuli.

2. Organization: This step of the perception process provides the answer to the question: *How does the viewed object look like?*
   Processes of the stage organization have to reconstruct the 3D distal stimuli from information of the 2D proximal stimuli. The human visual system now generates a so-called *percept* from the distal stimulus. The percept is a sufficient descriptor of the exterior environment and includes information about simple, sensory attributes like hues, edges of viewed objects, lines, homogeneous areas, texture information, colors and color transitions. Organization involves processes to assemble homogeneous areas by scanning for color and texture similarities, to group such areas according to figural or color properties, to judge areas and objects to be in the foreground or background and to perceive depth. If we consider the examplary AR scene, the observer views among other things two homogeneous areas, i.e. the percepts, representing the spinal column and the body of the patient. In this stage, the observer still does not identify the percepts as the spinal column and the body. However, he is able to define a depth order primarily using the depth cues occlusion and motion parallax.

3. Identification and classification: This step of the perception process provides the answers to the questions: *Which object is it?* and *Which function has the object?*
   Now the percepts of the distal stimuli get a signification by comparing the percepts with representations in the observer's memory. In reference to the example, the surgeon identifies among other viewed objects the two homogeneous areas as the spinal column and the patient. The visual system notices that the previous judgment of the depth order was wrong. The user's knowledge about anatomy corrects depth perception.

After having run through the stages of the perception process, the surgeon possesses all information about the distal stimulus engendered by the viewed physical and virtual objects. However, the stages *organization* and *identification* come into conflict. When the spinal column and the body in the background are identified, the observer has to use additional mental workload to correct the previously perceived order in the stage *organization*. Because of this conflict the intuitiveness of observing the AR scene gets lost. In particular, hand-eye coordination, for example reaching the operation site with an instrument or the surgeon's hand, becomes difficult.

The problem shall be modeled with the addition of a forth stage, *order correction*, which includes the process of correcting depth perception. The task of the present work

Figure 2.6: Perceptual conflict between stage *organization* and *identification* requires a forth stage *order correction*.

is to find ways to optimize the process of visual perception without being augmented by a stage *order correction*. The virtual and real proximal stimuli projected to the frame buffer and the CCD chip have to be intelligently composed. The resulting composition, the augmented distal stimuli presented on the monitor, has to enable the processes of the stage *organization* to correctly perceive objects within the AR scene without getting in conflict with stage *identification*.

2.3.3 Visual Sources for Depth Perception

In general, the perception of depth is an intuitive and natural exercise for most of the people. They do not spend much mental workload on measuring the distance to the car in the front to avoid a collision. They simply rely on what they see and have learned to interact with their spatial environment. However, artists have shown, that the human perception can be easily duped or at least confused. Popular examples are the Ames Room [18] or the perspective studies of MC Escher [62].

Visual depth perception concerns the estimation of relative and absolute distances among objects and absolute distances between the object and the observer. When processing the *proximal stimuli* during the stage *organization*, a wide range of depth cues deliver spatial information to put observed objects into order and relation or even estimate distances. This section will analyze the applicability of the natural depth cues to the capabilities of the used HMD.

Regarding the vision through the HMD, the estimation of absolute distances between the observer and objects within the field of view is extremely difficult. The optical hardware would have to exactly simulate the wearer's individual human optics. This is, however, not possible yet. Within the scope of user studies, we observed that in particular subjects wearing the HMD for the very first time, first try to scan their environment with their hands, e.g. feel the distance to the phantom body by touching it or moving their hands in front of the camera. Even though they see an object, e.g. an instrument to be used for the experimental task, lying on the table in front of them, they have difficulties to reach it with their hands. Mainly, the relative offset of the mounted cameras to the natural eye position and geometry and the optical zoom due to the camera lenses hinder the perception of absolute distance of objects. We, however, also observed that subjects get quickly familiar with the optical configuration they are exposed with and learn to synchronize the visually perceived information with their body motion within a short time.

The effectiveness of certain, natural depth cues depends on the depth of the observer's field of view. Cutting [49] defines three spatial domains to categorize depth cues according to the distance of objects of interest within observer's working environment.

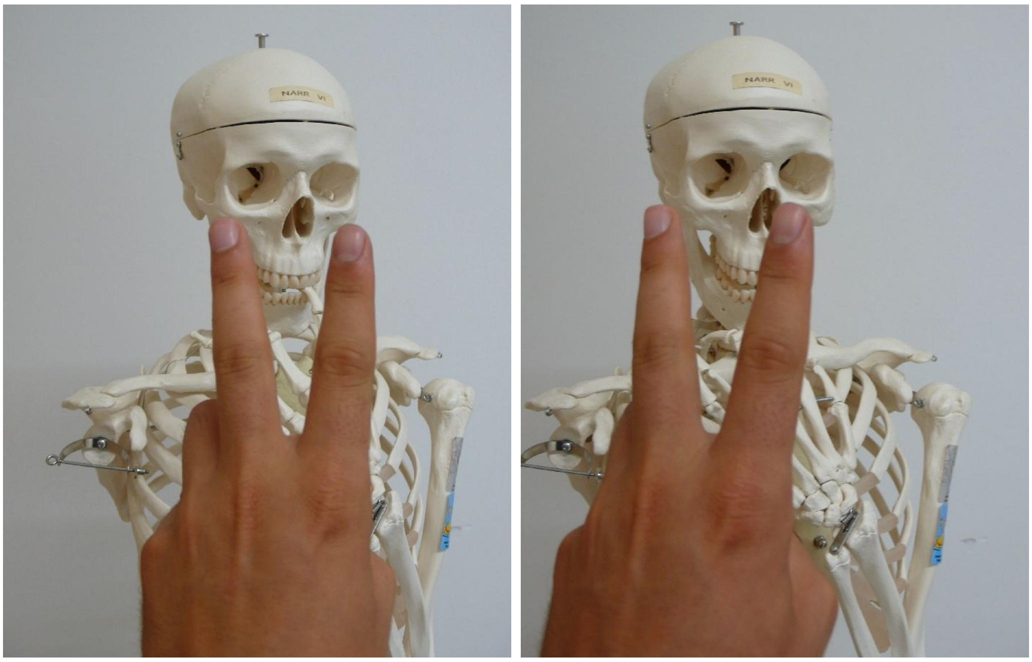
- Personal space(up to 1-2 m): e.g. Performing surgery.
- Action space(1-2m to 10-30 m): e.g. Throwing a basketball.
- Vista space (10-30 to ... m): e.g. Playing soccer.

According to the intra operative application of the AR system, the surgeon's workspace corresponds to the category *personal space*. For this reason, sections 2.3.3.1, 2.3.3.1 and 2.3.3.3 strongly focus on the most effective depth cues within this space. The four strongest depth cues are *occlusion*, *binocular disparity* or *stereopsis*, *motion perspective* or *motion parallax* and *relative size*. Sections 2.3.3.4 addresses *chromostereopsis* as an additional visual cue inducing an illusion of depth. Mirror reflection, transparency effects and shadow can provide depth information, however, require the perceiver to identify percepts in order to derive attributes such as light effects or material properties. Depth information being available in the stage *identification and classification* will be discussed in sections 2.3.3.5, 2.3.3.6 and 2.3.3.7.

2.3.3.1 Oculomotoric Criteria for Depth

The oculomotoric criteria contributes to depth perception due to the fact that most people view their environment with two eyes. The horizontal distance between human eyes has been determined as around 6 cm. Because of this, each eye receives a slightly different distal stimuli, which is projected onto the retina and stimulates slightly different receptors. The brain then combines ocular, sensory information, the so-called *binocular source*, and generates depth information for objects within the field of view. The resulting depth cue is called *stereopsis* or interchangeable *binocular disparity*. Figure 2.7 shows two slightly different images projected onto the retinas of both eyes (in this case on the photo diodes of a CCD chip in a digital camera). The binocular disparity increases when the distance between two objects in the field of view, for instance the fingers and the skull, increases. In a self-experiment the reader closes the right eye and views a close object with his left eye. Then his fingers forming the victory sign are positioned in between his eyes and the object. The resulting scene is the distal stimuli for the open left eye, which is shown in the left image. When the observer now closes the left eye and opens the right one, while fingers and object stay at their position, he will view an image that slightly differs from the left one.

The depth cue *convergence* demands the usage of ocular musculature. If an object is positioned near to the observer and focused with both eyes, the lines of sight of the eyes



(a) left eye (b) rigt eye

Figure 2.7: View on the scene with the left and the right eye.

converge and cross at the focal point. Information about convergence is gained from the movement of the eyes measured by the receptors of their musculature. These receptors transfer the stimuli to the brain, which can estimate by experience values of the distance to focused object. If the focal point is located at distances greater than three meters, the lines of sight of both eyes are almost parallel. The movement of musculature gets too little and convergence becomes irrelevant for depth perception.

In the case of the used AR system, the cameras' lines of sight converge at a distance of around 60 cm. The predefined orientation of the cameras takes over the job of convergence when the focused object is exactly positioned at the focal point of the cameras. This means, the point of interest in the AR scene is projected to the center of both displays. For closer or more distant objects, convergence is performed by the observer viewing corresponding distal stimuli projected at different positions on the displays of the HMD. The distal stimulus moves within the predefined field of view of the cameras. However, convergence being applied to direct vision of a real scene would result in a simultaneous movement of the field of view and the stimulus. For this reason, perceptive experiences related to convergence can not be seamlessly transferred from natural vision to video see-through vision. Stereo photography, which is almost as old as photography itself suggests parallel orientation of cameras having a distance of around 30% of the natural eye distance. The optimal camera configuration for video see-through HMDs in different workspaces has to be estimated in order to achieve optimal stereo vision. Our group started to develop a system combining optical eye tracking and reorientable color cameras mounted on the HMD to focus objects within the field of view. Even though, the sensor information from ocular motion can be accurately tracked and intelligently used to control the camera pose, perceptual information can be also strongly biased by other cues such as the optical model of the lens.

2.3.3.2 Motion Induced Depth Cues

The following information sources for depth perception require relative movement between the observer and the observed scene. A popular example to explain motion induced depth cues shows an observer sitting in a moving train and watching the landscape outside the window. The relative movement of the projection of stationary objects caused by the observer's movement is called *motion parallax* [49]. The observer sitting in the train perceives distant objects, e.g. a mountain peak at the horizon, as passing by slower than closer objects such as the railway poles. *Motion perspective* [85, 86]) provides the similar effects on percepts as *motion parallax*. In contrast to *motion parallax*, the whole field of viewed objects moves and the observer stays at his or her position. The reader may imagine an observer watching a group of balloons in the sky.

Another depth cue being quite similar to *motion parallax* is *covering and uncovering planes*. When the observer moves two planes respectively, for example two house walls, these walls seem to move relatively to each other. The wall A in the front covers or uncovers the wall B, which is located behind A. This cue gets ineffective as soon as the observer moves perpendicular to the planes.

Although, the mobility of the observer in our targeted AR scenario is restricted due to further surgical staff and equipment of the OR. The surgeon is still able to move his head to enable motion induced depth cues. Regarding the exemplary AR scenes shown in figures 2.5, slight motion even reduces the perceptive quality since the spinal column or the skull seem to move relative to the skin even though anatomy and skin are fixed relatively to each other. An earlier approach of the author proposed a virtual window that is attached to the body's skin surface to reveal the view into the body [22]. An evaluation with 20 surgeons has shown that the virtual window, which covers and uncovers the augmented spinal column due to the observer's motion, can help to improve depth perception [200].

2.3.3.3 Monocular, Pictorial Criteria for Depth

The depth cues described above are motion induced or are based on stereo vision. Monocular cues also provide strong depth information for objects within a scene viewed with only one eye.

According to Cutting [49], the depth cue "*occlusion* exceeds all other information" for depth perception and "the effectiveness of *occlusion* does not attenuate with distance". However, *interposition* does not provide information about absolute distances between the objects and to the observer. It only allows to perceive whether one object is closer to the observer than another object or not. Exactly this depth cue confuses the vision in the targeted medical AR scene as shown in figure 2.5, where the spinal column partially occludes the thorax.

The source of information *relative size* concerns the relation between perceived size and distance to physically similar sized objects. The objects, positioned in variable distances to the observer, create different sized percepts. Their size decreases with distance. If the observer does not have any previous knowledge about the natural size of the viewed object at least two similar objects have to be present in order to let *relative size* become effective. If the objects are familiar to the observer, for example he identifies percepts as birds in the sky, one object is sufficient to estimate absolute distance to the observer. In this case, the terms *familiar size* or *assumed size* are used in the literature [60], [61]. *Familiar size* can cause illusions about distances. For example, if the viewed objects are trees of a certain species and some of them are younger than others, they don't have the same size but look similar [28]. In addition to the surgeon's knowledge about the actual position of the spinal column inside the body, *familiar or relative size* may help to correct depth perception (see Fig. 2.5). However, according to Cuttings hierarchy showing the effectiveness of depth cues within different spaces [49], the cue *occlusion* is much stronger than *relative or familiar size.*Occlusion rather blocks the ability to judge the dimensional relation of the body and spinal column.

The following depth cues *texture perspective*, *linear perspective*, *relative height* and *aerial perspective* have only few effects on the perception of objects within personal space.

A group of similar objects forming a pattern enable the depth cue *texture perspective* or *relative density*. This cue concerns the projected retinal density of a cluster of objects or textures, "whose placement is stochastically regular, as they recede into the distant" [49]. The ground plane shown in figure 2.14(d) shows the effect of *texture perspective*.

Linear perspective is a depth cue that relates size to distance. Parallel lines, for example rails approaching the horizon, seem to converge somewhere at the horizon line. *Linear perspective* is a "systematic combination of several other sources of information" [49]. It primarily combines cues from *relative size* and *texture perspective*. *Linear perspective* requires distance and has almost no effect within the *personal space*.

Relative height results from the position of objects relative to the horizon line. Objects positioned higher above the horizon line seem to be located nearer than objects positioned lower or close to the horizon line. If we compare objects beneath the horizon line, we perceive the opposite effect. The effectiveness of *relative height* or *height in visual field* to estimate depth decreases with distance and is not helpful in *personal space*.

Small particles, moisture or/and pollution, in the air let distant objects appear blurred and diffused and their color seems to be less intensive. Objects, which are positioned closer to the observer, have intensive colors, seem to be more detailed and rich in contrast. Painters use this effect to create pictorial effects for depth. For example, when watching the landscape from a lookout point, distant objects like peaks of mountains "appear blue and almost of the same hue as the atmosphere itself" [213] in contrast to lusciously colored trees located in the foreground. Effectiveness of *aerial perspective* increases with distance. However, with too large distance, the objects become indistinct or nearly invisible and the depth cue becomes ineffective. Within the *personal space*, *aerial perspective* hardly contributes to depth perception. Gerbino et al. [84] and Metelli [148] mention, that the concept of *aerial perspective* is justifiably applied to the perception of transparency. They study the visual perception of the depth order of two or more colored transparent sheets. Instead of the air, the sheets become the medium that generates information for depth perception. The perception of transparency will be discussed in section 2.3.3.5.

2.3.3.4 Chromostereopsis and Color Composition

Sophisticated color composition can strongly support the perception and usability of a rendered scene [238, 138]. One aspect of color composition is the ability of the human visual system to correlate distances with colors. The phenomenon *chromostereopsis* or sometimes also called *color stereoscopy* [41, 58, 6, 231] allows the perceiver to assign relative depth positions to uniformly colored objects. Einthoven explains this effect with the chromatic difference of magnification, "for example, blue rays are refracted more than red rays by the ocular media, their foci not only lie at different levels (chromatic aberrations) but make different angles with the optical axis, and will thus stimulate disparate points. It follows that individuals with temporally eccentric pupils see red in front of blue, while with nasally eccentric pupils the relief is reversed" [58]. The difference of the index of refraction between blue and red is around two dioptres and therefore causes a comparably extreme effect with a black colored background (see Fig. 2.8). Covering partially the nasal (or the temporal) part of the pupils can enhance (or weaken or even reverse) the depth effect [124]. Thompson et al. report that the perceived depth "may often be the combination of a luminance-based depth effect and a color-based depth effect". This is supported by their finding that some people rather see the blue disks in the foreground and they observed that the "depth effect is reversed when the brightness of the display is changed" [161]. Reversion of the effect can also be affected by switching to bright backgrounds. Winn et al. [245] investigated the reversion and change of depth level by controlled color changing of the background and the objects in the foreground.

In order to enhance *chromostereopsis*, special goggles have been developed that artificially increase the angles of refraction and enable the usage of a wider range of colors to induce depth information, for example the goggles produced by the company Chromatek [11]. Even though, the observer has the feeling of being able to allocate depth information, the cue looses effectiveness as soon as motion comes into account. When moving sidewards relative to figure 2.8, the depth cue *motion parallax* stays disabled, since relative distance between red and blue disks does not change.

Color layout has been determined in computer graphics and other communication fields as extremely important to shape "the perception, interpretation, and memories of everything seen" [138]. For example, creating a color contrast among virtual and real objects can enhance the visibility of important information in the scene. An extreme example is the so-called *pop outs* that intensively attract the observers attention [17]. Regarding depth perception, *chromostereopsis* can help to perceive the topology of a complex scene [237]. In the medical domain, for example Alan et al. [47] use *chromostereopsis* for illustrative visualization of complex vasculature.

In the targeted AR scene, the color of reality is predefined, however, the colorization of virtual objects is flexible. One can imagine drawing interior anatomy rather bluish in order to create the contrast to the warm, "reddish" skin color as shown in figures 2.3(a). However, care has to be taken since observing extreme blue and red shades next to each other can result in a composition that is extremely discomfortable to view.

[11] American Paper Optics Inc.: http://www.chromatek.com

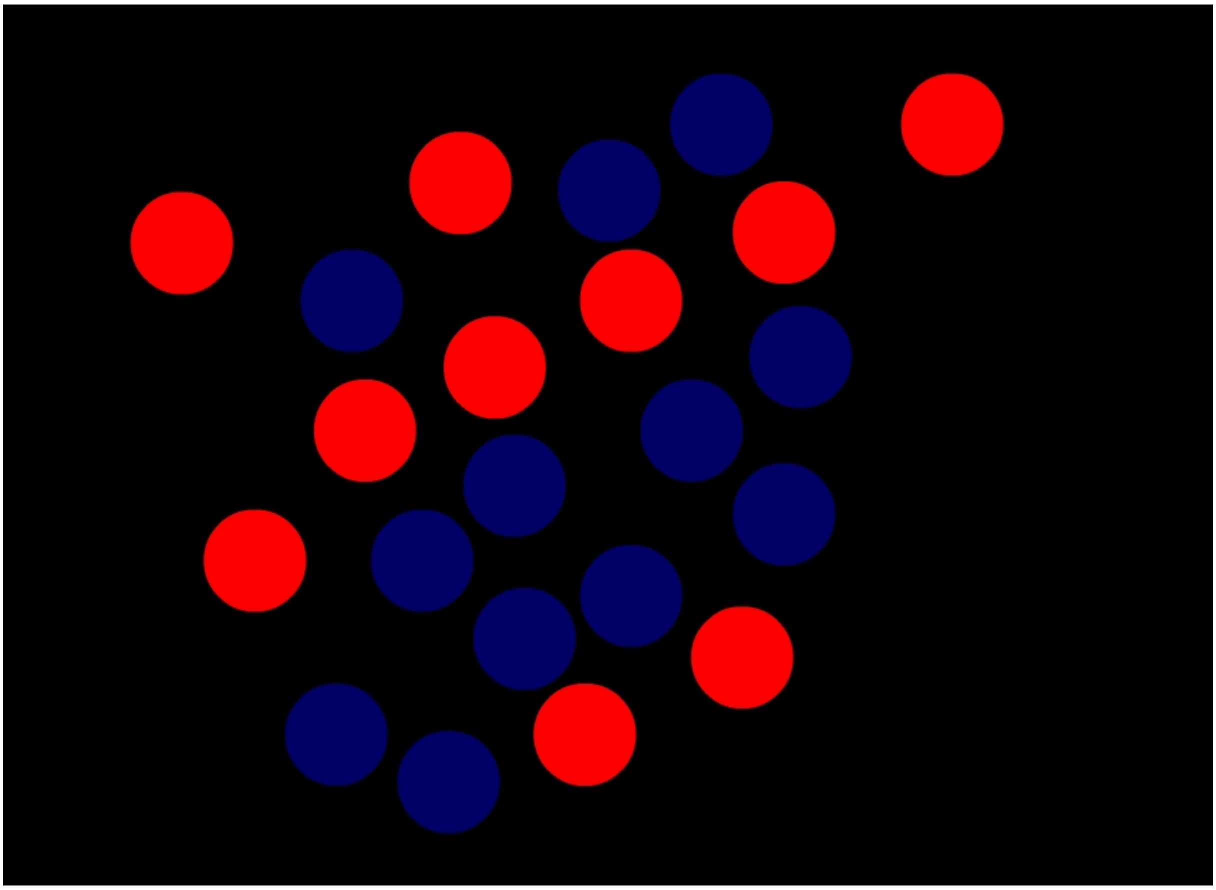
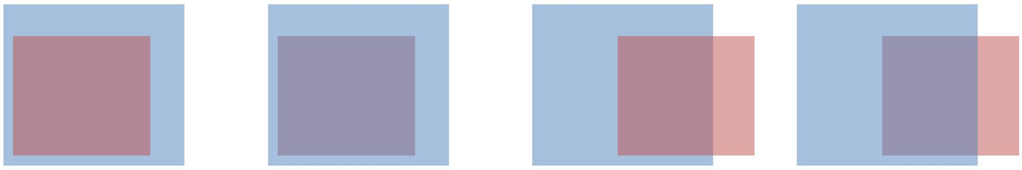


Figure 2.8: Chromostereopsis

2.3.3.5 Transparency

The term *perceptual transparency* (PT) is related to the perception of two objects being located in two different depth layers, while one object can be seen through the other one [123]. Even though computer graphics and print media are capable of effectively simulating *physical transparency*, the correlation of material properties of an object that does not completely absorb or reflect electromagnetic radiation to PT is still not entirely investigated. *Physical transparency* is not essential to actually perceive transparency [148].

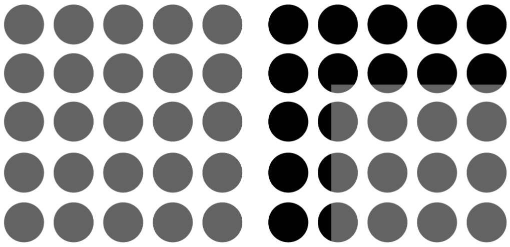
When an object B is opaque or transparent and placed partially behind a transparent object A, and A and B share the global background C, a proximal stimulus pattern consisting of four surfaces is generated and organized in the stage organization. Grouping the occluded and non-occluded part of B as one object requires experiences about color composition to derive depth information [196]. If B is completely behind A, a proximal stimulus pattern consisting of only three surfaces is generated, which is grouped. If the observer does not know about the color of the occluded object B, he cannot be sure whether A is behind B or vice versa. Figure 2.9 shows two transparent objects in different constellations relative to each other. The existence of a global context in which transparent objects are embedded is essential for PT. When viewing for example an isolated surface consisting of several transparent layers without showing the background, the transparency



(a) (b) (c) (d)

Figure 2.9: Only if the colors are known, experiences in color composition can tell, which layer is in the foreground.

is not obvious anymore (see Fig. 2.10) [98, 81]. However, the stereo presentation of interior



(a) (b)

Figure 2.10: The proximal stimuli has to provide a global context pattern in order to perceive the transparency of a surface.

objects behind a transparent layer seems to weaken depth perception. Johnson et al. [107] introduce an AR system that renders virtual, medical imagery under a transparent surface, for example the skin of the patient. They mention that "the presence of transparent surfaces diminishes the surgeon's ability to use the stereo information effectively and get an accurate sense of depth".

Wang et al. [238] recently presented a study that investigates the best color composition when rendering volumetric medical imaging data such as CT. One important aspect of their study is the parametrization of semi-transparent layers, in their case the transparent skin revealing the view onto interior bone structure. The outcome is a set of rules how to achieve the best composition providing a good balance of visibility, depth perception and color harmony. They suggest to mix colors of transparent superimposed layers rather locally than in a global way. When both superimposed layers are rendered with an individual alpha value, the opacity of the object in the foreground should show a higher weight. They further suggest to use cold colors such as green, blue or cyan in front and warm colors such as red, yellow or magenta in the back, which would be contrary to the findings of *chromostereopsis*. However, their explanation for the "unexpected" effect is the overlapping arrangement of colors in their renderings while *chromostereopsis* requires the objects to be located side by side. The group around Wang also found that higher lightness of foreground objects is beneficial to perceive the depth order. For blending objects with the equation 2.3, Wang et al. [238] proposed to use hues for background and foreground objects that lay on the opposite sites of a color wheel. The color wheel has been introduced by Itten [104] and is intended to be used as a tool for the definition of harmonic color schemata [153].

$$C = C_1 * a_1 + C_2 * a_2 * (1 - a_1) \tag{2.3}$$

The choice of opposite colors helps to avoid the introduction of new colors and renders blended areas rather with neutral color or a less intensive representation of original foreground color. All other pairs on the color wheel, however, produce new hues when being blended, which would rather confuse the definition of depth order of objects.

Regarding the AR scene, the color of real objects should not be changed, e.g. changing the skin color to blue. This would counteract the efforts of carefully and smoothly augmenting the natural environment of an observer.

Transparency in combination with shading effects can have a big impact on the perception of depth layers. In our everyday life, we permanently have to deal with the transparent objects and estimate their depth in space. For example when diving into water accurate timing is required to avoid a belly flopper. Water is an almost transparent medium. When diving into the water it can be difficult to use the edge of the pool or the lake shore as a reference to the transparent water surface. Our field of view is focused on the location of impact and the proximal stimuli of the water surface is not embedded into a global context pattern. However, small waves cause light reflections on the water surface and form a "naturally" structured texture enabling the diver to perceive the distance to the water surface and to timely prepare for the impact. For this reason, the water is usually sprinkled during high diving contests. A similar effect can be observed when looking through curved glass.

The depth perception of objects in the background completely covered by transparent objects in the foreground can be improved by adequate lighting conditions, texturing and material attribution. The natural reflection of incoming light on the water surface produces specular highlights. Such highlights change the color and even occlude parts of the interior as shown in figure 2.11(d). In fact, this effect is enhanced by the primary depth cue in personal space, which is *occlusion*. When the surface is not entirely flat and follows a certain texture pattern, this effect can be even improved. The ability of the perception system to group similar percepts [170] may serve as a motivation to use semi-transparent, structured textures for depth perception.

Figure 2.11(a) shows two transparent, superimposed layers. When enabling shading, the layers can be identified as 3D bodies (see Fig. 2.11(b) and 2.11(c)). However, it is still difficult to perceive depth order of the two objects. Adjusting the material parameters of the surrounding object to let it look like glass enables the perceiver to define the depth order due to the occluding specular highlight (see Fig. 2.11(d)). The manipulation of geometric surface properties, here the bumpiness of the glass surface, using a texture [12] can enhance the spatial appearance of the covering object and finely graduate the occlusion of the deep seated object (see Fig. 2.11(e)). In addition, the texture can be used to colorize the covering object. This enables the depth cue *occlusion* due to material parametrization and illumination but also due to color fragments of the texture structure (see Fig. 2.11(f)).

[12] This effect is known as bump mapping in computer graphics.

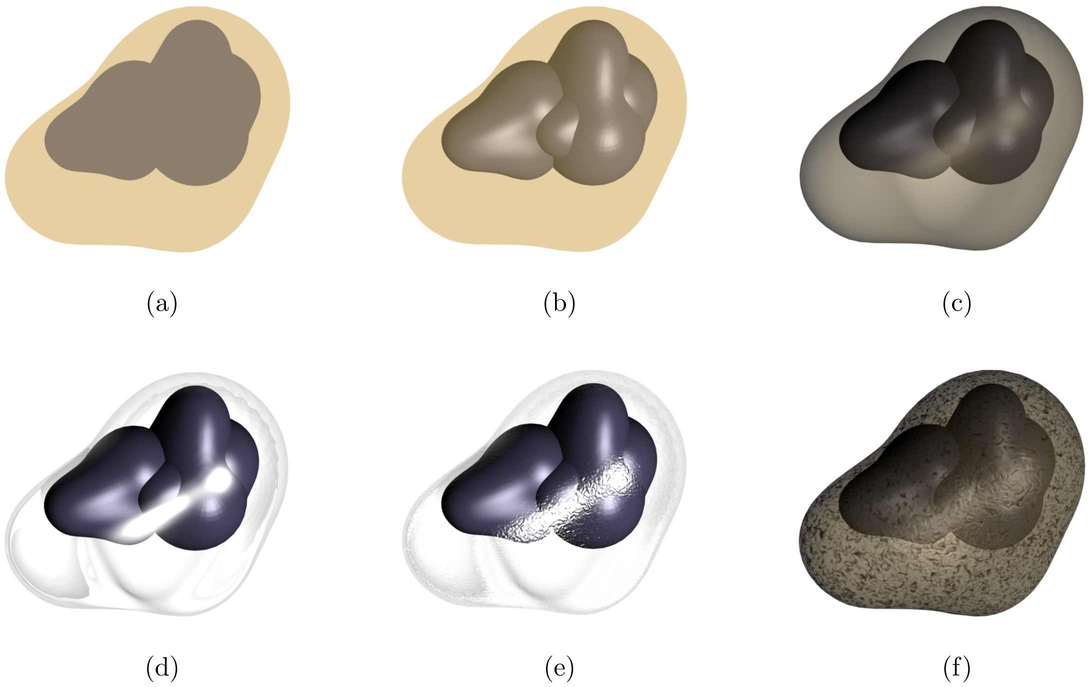


Figure 2.11: Perception of superimposed transparent layers.

Regarding the figures 2.11(d) , 2.11(e) and 2.11(f), in the stage organization, the perceiver deals with three percepts; the semi-transparent area covering the bluish area, and a white area (the specular highlight) occluding the bluish area. The only unambiguous conclusion about depth order of percepts is the white area being in front of the bluish area. In the stage *identification* the material of the covering object is classified as glass or a glass-alike material and the white object as a light effect due to the reflective properties of glass material. The perceiver is then able to conclude that the percept corresponding to the specular highlight is located on the surface of the covering object and has the same position in depth (see Fig. 2.11(d)).

Regarding the design of the AR based view into the body, the most obvious approach to handle the problem of misleading depth perception may be to simply render the skin transparent. In the exemplary scene shown in Fig. 2.5, the virtual anatomy is completely enclosed by the real skin. One visualization mode of an earlier evaluation study about depth perception in medical AR scenes [200] rendered the skin transparent. Even though the resulting composition provides only ambiguous information about depth, the results of the study attest a perceptual progress compared to simple superimposition.

Regarding figure 2.5 the real skin has to be perceived as the covering layer having natural material properties. These properties are, however, changed during an interventional procedure. The patient is cleaned at the beginning and the skin becomes wet and produces probably more specular highlights that when it gets dry. Removal of hair, blood, disinfection, liquids may change the appearance of the skin as well. The objective is to reduce the falsification of the natural appearance of real entities in the AR scene, which could result in misinterpretation of the anatomic situation and negatively affect the intervention itself. We rather try to use geometric parameters (see approach described in section 3.1) of the skin and changes of the skin color (see approach described in section 3.2), while preserving the original skin color to create texture patterns that partially occlude the virtual objects in the background.

Following the work of Wang et al. [238], future work on the AR system could include an automatic routine to adapt the visualization individually before each clinical application. This routine could incorporate the detection of the skin color in the video image and the adaption of the transfer function for rendering of the CT data.

2.3.3.6 Shadow

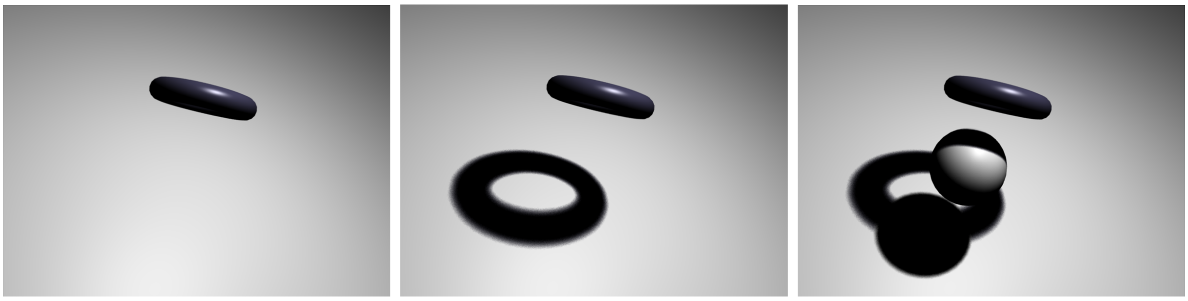
Shadow is a function of one or several light sources and illuminated objects. The illuminated objects are generally called *occluders* and project shadows onto the *receivers* [94]. It provides information about the position and size of the occluder (see Fig. 2.12 and 2.13(c)), the geometry of the occluder (see Fig. 2.13(b)) and the geometry of the receiver (see Fig. 2.14(c)), i.e. on which the shadow is projected [116, 139].

Shadow cast provides visual hints do perceive the height of an object in front of a background layer or depth order of various objects that do not occlude but throw shadow on each other as shown in figures 2.12. Kersten et al. investigated the perception of shadow

(a) (b) (c)

Figure 2.12: Shadow provides information about relative position of objects respective a background.

in particular its contribution for depth perception. His group shows that an observer is able to gain information about "the shape of the object, the shape of the background surface and the spatial arrangement of the object relative to the background" [139]. Figures 2.13 show that only due to the shadow, the blue object can be determined as a torus. When further objects are inserted into the scene, shadow extremely supports the perception of relative position of objects. In the case of the scene shown in figure 2.13(c), shadow provides the observer with the information that the sphere is located directly beneath the torus. From this information one can further derive the relation of sizes of both objects. Even though the static scenes shown in figures 2.13 already give an idea about the power of shadow for perceiving the topology of 3D scenes. Kersten et al. note that "shadows were perceptually most relevant for the recovery of spatial arrangement, especially when the shadow is in motion".



(a) (b) (c)

Figure 2.13: Shadow further allows for the perception of shapes, relative size and spatial order of objects.

Kersten et al. [116] arranged a psychological experiment introducing the so-called ball-in-box scene. Their "results support the hypothesis that the human visual system incorporates a stationary light source constraint in the perceptual processing of spatial layout of scenes" and "the information provided by the motion of an object's shadow overrides other strong sources of information and perceptual biases, such as the assumption of constant object size and a general viewpoint". The authors examine the perception of the 3D trajectory due to the shadow caused by the moving objects. The ball-in-box has been replicated in figure 2.14. Although the depth perception of the ground plane can be easily managed due to the depth cue *texture perspective*, it provides no relation to the balls position. Using the previously introduced depth cues *relative size* or *relative height*, one would argue to perceive both balls at the same depth within the visual field. Shadow reveals the true position of the balls and creates a connection to the ground plane. When the ground plane is textured, depth perception can be even quantified. In addition to the three shadow induced cues mentioned above, which help the observer to better understand the layout of a scene also spatial information of the light source can be roughly estimated. If the illuminated surface is not planar and the light source moves independently from observer motion above the scene, information about the position of the light can be estimated. For example when viewing a half moon, the observer can estimate the position of the sun.

With respect to the three stages of the perception pipeline, the perceiver has to identify percepts as shadow and put them into relation to the occluders causing the shadow projection. Even though, figural hints, in few cases the shadow might have similar shape as the occluders, will allow the perceiver to group pairs of shadow percepts and occluder percepts, it is assumed that fusing shadow percepts with their origin is performed in stage *identification*.

An extensive survey on real-time shadow algorithms in Computer Graphics for improving the realism of virtual worlds and better understanding their topology has been published by Hasenfratz et al. [94]. They discuss in particular the problem of creating real-time soft shadows. In virtual worlds designed for instructional objectives, shadows

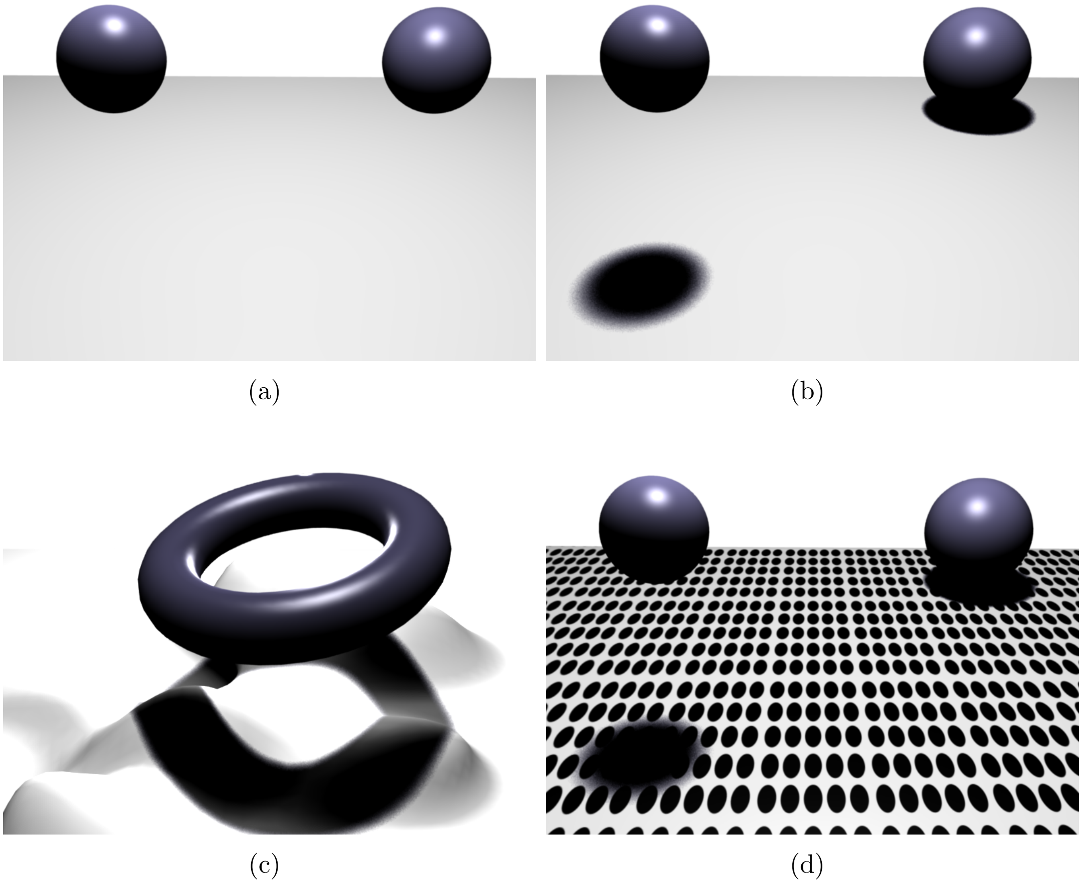


Figure 2.14: When similar objects with differing sizes are positioned at the same height of the visual field, shadow is capable of enabling perception of relative depth order among objects.

have also been proposed as a metaphor to better distinguish information that may be occluded [179]. Here, the shadow is used as an abstraction of a complex 3D shape, while the shadow metaphor is used as a natural link between the 3D and 2D representation of an object of interest, in this case human anatomy.

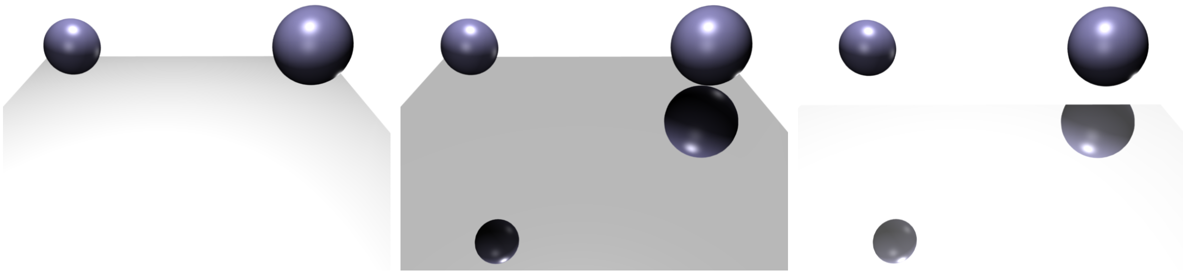
Also designers of AR worlds have determined the importance of virtual shadow being synchronized with real shadow to enhance the realism and better immerse the user into the augmented environment [92, 105].

2.3.3.7 Reflection

Reflection is a function of the perceiver's viewpoint, the position of the reflective object and the characteristics of the object to be reflected.

The most common reflective objects in our daily life are mirrors supporting us in navigation and observation tasks. Mirrors provide us with an additional perspective on regions of interest for fast decision making in traffic situations, to support hand-eye coordination for shaving the beard, or for exploring physically restricted areas like the oral cavity during dental treatment.

The combination of visual information from a mirror image and the observer's viewpoint results in an impressively enriching information about the order of the objects. Figure 2.15 shows two similar looking objects and a plane that can be changed to become reflective. The presented scene is in fact a variation of the earlier introduced ball-in-box scenario used to show the effectiveness of shadows for depth perception. The objects have the same color but not the same size. Due to the depth cue *relative size*, the right ball in figure 2.15(a) would be perceived closer than the left one. However, the reflective ground plane reveals the true position of the objects as shown in figure 2.15(b). Also the borders of the mirror plane that cut off the mirror image help to locate the objects (see Fig. 2.15(c)).



(a) (b) (c)

Figure 2.15: Modified virtual ball-in-box scenario to distinguish the importance of reflection for spatial perception. Without the mirror ground plane the bigger object seems to be located nearer due to the depth cue *relative size*. The presence of reflection can correct the depth order. Reflections provide the perceiver with additional perspectives to entirely explore an object or a structure of objects.

When motion comes into play, the interactively repositioned mirror provides even stronger depth cues due to visual feedback in the changing reflected image. *Motion parallax* or *motion perspective* can support the perception when the observer moves either his head or the mirror. Also *interposition* contributes to perception of depth when the visualization partially occludes the mirror plane. As an advantage over shadows, the projection of an object onto a reflective area allows for viewing physically hidden or occluded areas. An example from the real world is the dentist's mirror to explore areas inside the oral cavity of the patient which are not directly visible.

Further materials within our natural environment show reflections that are intuitively perceived such as water surfaces, glass or metal planes. It is much more difficult to derive exact spatial relation between the reflecting object and the reflected one than with mirrors, however, some rough cues are still available as shown in figure 2.16. Similar to the perception of transparency and shadow, the percepts of objects and their reflected counterpart cannot be related to each other in the stage of *organization*. For example, the scene presented in figure 2.15(b) shows six different percepts which are organized independently from the fact that two of them are reflections. First, one of the objects in the scene has to be identified as the mirror or any other object showing reflective material

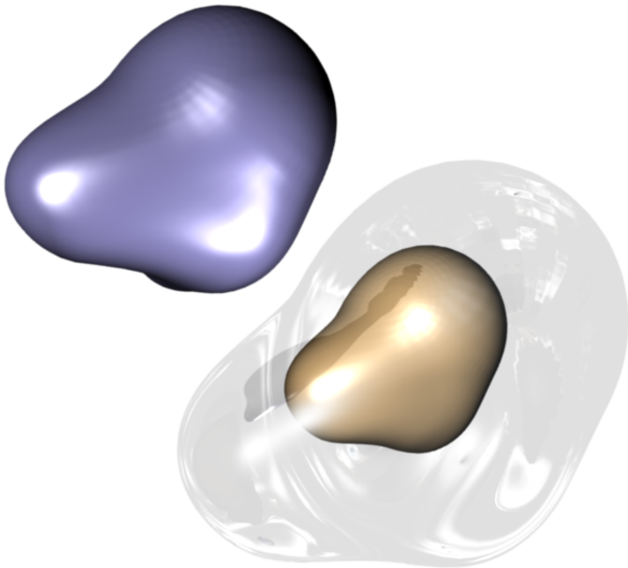


Figure 2.16: Reflective material in our natural environment provides cues for spatially relating objects to each other.

to further mentally restore the distal stimulus.

2.3.4 Perception of Shapes

Illumination and colorization are major attributes defining the quality of perceiving the shape and dimensions of objects within a 3D world.

2.3.4.1 Lighting and Shading

To simulate lighting of the real world, the standard light model in computer graphics uses three different types of light; ambient, diffuse and specular light.

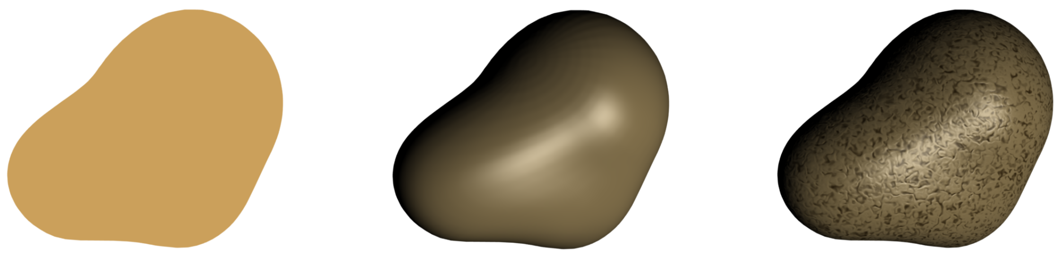
- Ambient light is the global illumination that has no positional light source or direction. For this reason, it does not exist in the real world. It is, however, used in computer graphics to approximate light that arrives uniformly from the sun and is scattered by so many other surfaces so that it uniformly illuminates an environment.

- Diffuse light has a defined point source, for example a lamp. The intensity of its reflection on surfaces depends on the orientation of the surface. However, this light component is reflected equally in all directions once it hits a surface.

- Specular light comes from a point source and produces highlights on smooth surfaces because it reflects the incoming light like a mirror in a particular direction defined by the surface shape. Intensity of these highlights depends on the material parameters of the illuminated surfaces and the specular ratio of a light source that illuminates the scene. For example, we get more intensive effects on a glass bottle than on a tennis ball. On a sunny day, sunlight contains more specular light than on a cloudy day, when light is mainly diffuse.

The color of a certain point on the surface of an object depends on the normal of this point respective to the orientation and position of the light source. The visual effect is called ”shape from shading” [38] and provides the observer with information about the shape of the object as shown in figure 2.17. Extremal values are areas on the surface, which light cannot reach because the object casts a shadow on itself and areas, where the gradient of the point on the surface is parallel to the direction of the light and highlights can be seen.

In computer graphics a technique called bump-mapping has been introduced that employs gradient textures to generate finely structured, textured surfaces. The gradient textures are used to apply a lighting model such as phong shading [171] to the textured surface. Using this technique can dramatically increase the realism of natural materials such as stone, skin or wood. The pattern of specular highlights created by a bump map can be applied to transparent surfaces to support the task of spatially sorting background and foreground objects as shown in figure 2.11(f). *Shape from texture* will be further addressed in the following section.



(a) no shading (b) phong shading (c) bump-mapping

Figure 2.17: Enhancing the realism with shading.

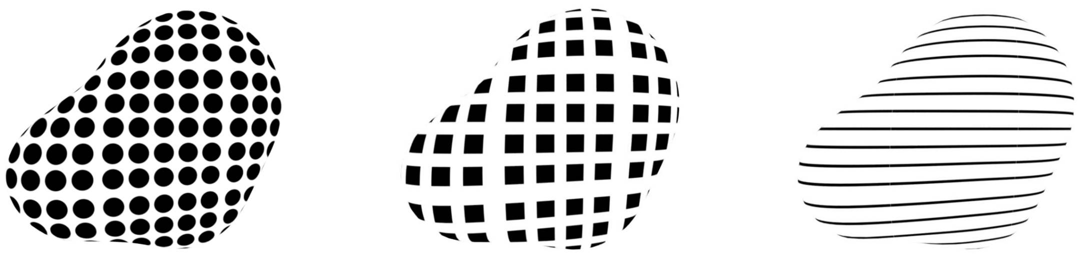
2.3.4.2 Texture Gradients

Shape from texture has been established as a method of computer vision to reconstruct a 3D object from its 2D projection by using information from its uniformly patterned color [7, 83]. Similar to *shape from shading* human perception is able to use such patterns as a visual cue to perceive the shape of textured objects [85]. Gibson et al. [85], one of the pioneers having proposed psychological theories regarding the perception of *shape from texture*, argued that the density gradient serves as the primary source for shape perception.

In computer graphics, the individual colorization of objects is usually realized with a technique known as texture mapping. Samples of a natural structure are mapped onto an object and enhance the realism of a virtual scene. In addition, textures can help to better perceive the shape of an object. Many of these disks are positioned like grid points on the texture and form a uniform pattern. Reorientation of simple shapes can be easily identified. In the case of disks, the eye interprets a spatial rotation, which is, however, ambiguous respective the direction of rotation, if just one disk is presented as shown in figure 2.18. A structured pattern of simple shapes create an unambiguous impression

Figure 2.18: Round disks as texture primitives

about the three-dimensional shape of the textured object (figure 2.19).

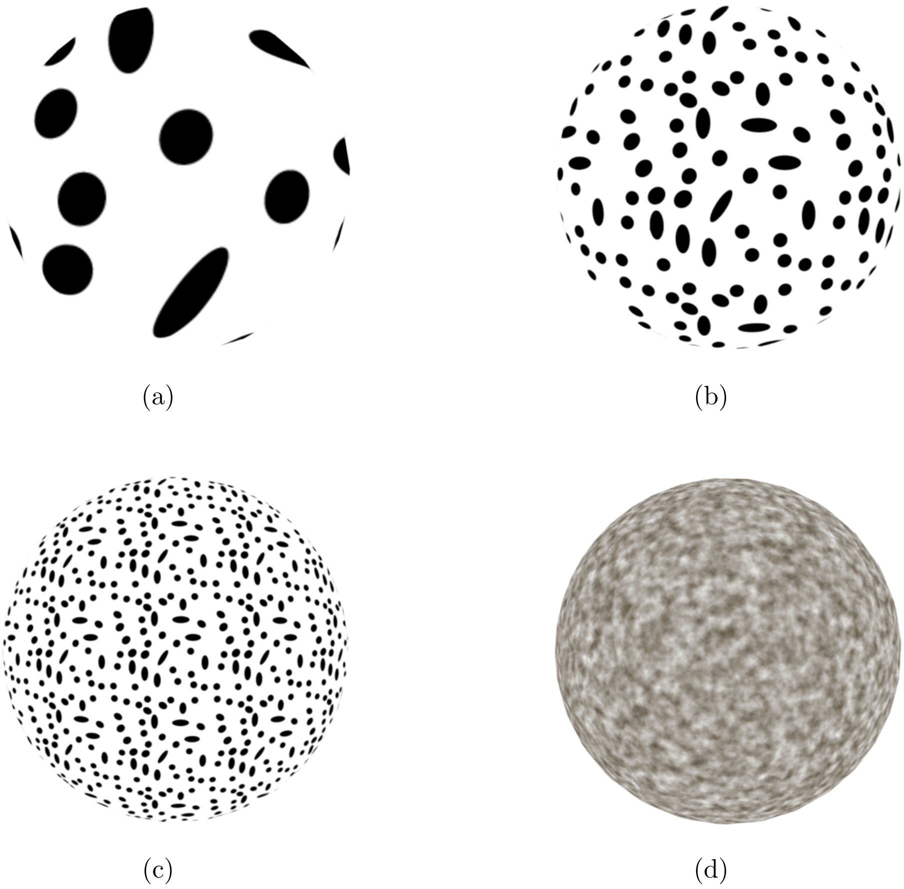


(a) (b) (c)

Figure 2.19: Shape perception from regular blob and conture textures.

Todd et al. [217] arranged a set of experiments to evaluate the perception of curvature of regular and irregular conture and blob textures. Participants view a simple textured shape consisting of two converging planes while the intersecting line is centered vertically in the projection image of the scene. The scene was presented on a monitor. Participants had to judge with varying field of views and optical slant of the planes the sign of curvature, either concave or convex. Interestingly, with irregular blob texels 76% correct sign of curvature was perceived. Todd further investigated the perception of more complex shapes such as "doubly curved surfaces " [216] that are depicted with anisotropic texture. They report on a high degree of accuracy to determine the shape of differently textured objects exposed to the particupants of their study. Spheres shown in 2.20 have a much simpler shape than the ones used in Todd's study [216]. However, when considereing the patient's skin surface as the target shape to be depicted with texture as visual cues the presence of concave curvature can be neglected. Figures 2.20 induce that with increasing density and simplicity of the object shape the ability of the human visual system to perceive the shape of the object increases.

Figure 2.20(d) employs a texture generated with the texture generating filter *Clouds* provided by *blender*. Having shading disabled, the anisotropic, color texture looking like organic material still gives hints on the shape of the object. The human body is depicted with similar texture due to skin folds, hair, pores, scratches or birth marks. For the proposed methods of *contextual in-situ visualization* presented in section 3 the irregular, anisotropic texture information from the skin, however, serves an essential visual cue for



(a) (b) (c) (d)

Figure 2.20: Shape perception from irregular, anisotropic blob textures.

the design of the context layer in order to create the view through the transparent skin surface into the body.

2.3.4.3 Proprioception/Kinesthesia

Proprioception, or often called interchangeably *kinesthesia*, is the perception of the position and orientation of one's body. This information is gathered from stimuli of sensory receptors, the so called proprioceptors, found in muscles, tendons and joint capsules [114]. Proprioceptors give feedback about tonicity and muscularity, motion of joints respective deflection, velocity and direction of motion and position of parts of our extremities.

Muscle memory and hand-eye coordination result from kinesthesia and can be trained. It enables equilibrium as well as fine motor skills. In the surgeon's personal space exact hand-eye coordination is essential to perform highly critical tasks requiring fine and precise movement. As mentioned before, subjects having tried the HMD for the very first time, first carefully explore their AR environment by carefully handling instruments and approaching a simulated patient. Their proprioception, however, can be quickly synchronized with visual information from the see-through AR systems. After a short warm up period subjects could easily interact with the system. The current AR system setup does not provide haptic feedback. However, there are several interaction induced visual cues to align proprioception and visual perception, for instance shadow and mirror reflection.

CHAPTER

THREE

CONTEXTUAL IN-SITU VISUALIZATION

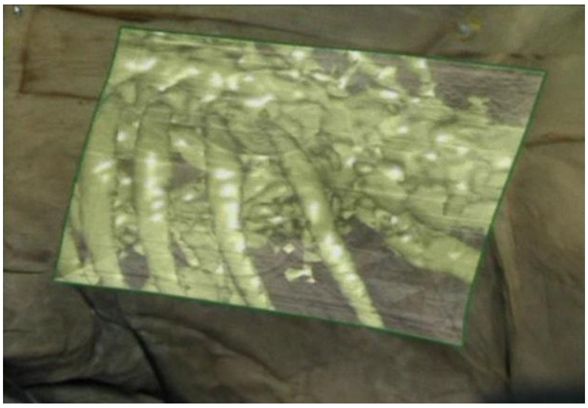
The following chapter presents two novel approaches to achieve sophisticated image composition of real and virtual anatomy of the patient. These approaches generate artificial depth cues to correctly and intuitively perceive depth, pose, and shape of all involved objects in an AR scene.

For this purpose, a new term has been derived from the expression in-situ visualization [189] (see also info box 1.2.1). The objective of *contextual in-situ visualization* is to present embedded virtual objects in context with real ones to solve the previously described problem of misleading depth perception.

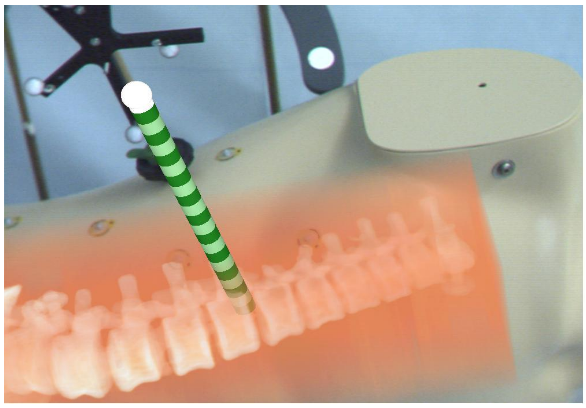
From the author's working experience, feedback from surgeons, students and many other users who experimented with the AR system described in section 2.1, one may derive the following guidelines generating sophisticated the in-situ visualization of medical imaging data and navigational information.

1. The visualization method has to provide a non-restricted view at the anatomical region of interest, which includes concealing obstructing anatomy and emphasizing the operation site.
2. Visualized imaging data has to be smoothly embedded to correctly and intuitively perceive distances among objects. The quality of the transition between virtuality and reality has to create the perfect illusion "to overwrite what the observer knows to be true" namely that the observer does not have the physical ability to see into the patient [108]. For this reason, the factor *aesthetic of the image composition* has played a major role to develop the methods. According to Norman "Products designed for more relaxed, pleasant occasions can enhance their usability through pleasant, aesthetic design" [159]. The end-user shall be prevented from thinking of technical issues that allow the view through the skin. He should simply accept and take the advantage of his augmented vision.
3. When using the AR technology intraoperatively, surgical instruments have to support the observer's perception when instruments are navigated within the patient's body. This can be achieved by virtually imitating real visual feedback from interaction such as changing light conditions.

Sielhorst et al. [200] described an evaluation comparing the perceptive efficiency of seven different render modes visualizing the spinal column in situ. Their results show that a virtual window overlaid on the patient's skin is capable of overcoming the lack of misleading depth perception. The virtual window can be positioned interactively by head motion of the user wearing the HMD, which is described in section 2.1. The development of a virtual window has been inspired by the "synthetic hole" introduced by Bajura et al. [15]. Such windows enable powerful depth cues such as occlusion and motion parallax to correct depth perception of the scene, however, they look rather attached to the skin than smoothly embedded. In contrast, *contextual in-situ visualization* is supposed to provide a much more intuitive and elegant method of data presentation. Here, visual depth cues are composed of natural attributes of the individual anatomy and the observer's point of view.



(a) Application of the virtual window [200, 22] to CT data of a cadaver.



(b) Virtual counterpart of the instrument is shown outside and inside the body of a thorax phantom.

Figure 3.1: When applying the virtual window to 3D visualization of CT data superimposed onto a cadaver as shown in Fig. 3.1(a), the unsatisfying image composition becomes obvious. Here the virtual window is not smoothly embedded into the scene. The virtuality of the window catches one's eye and one does not perceive the window and the skin surface as belonging to the same depth layer.

The basic idea of a new approach of how to compose virtuality and reality of the target scene is to compute a transparency map for the skin shown in the video using different information sources. In other words, one has to adaptively define the alpha values of pixels of the video images that are captured by the cameras of the video see-through head mounted display simulating the eyes's view. The composition of virtual anatomy behind the transparent skin serving as a context layer is inspired by "ghosting" a well known technique being applied to virtual worlds, which is a subdiscipline of the field of research focus & context visualization [126, 39, 40].

Within the scope of this doctoral thesis, two approaches of *contextual in-situ visualization* have been developed. The proposed methods can be applied for any video see-through technology being used for AR applications. In the medical field this includes for instance laparoscope or arthroscope cameras. The geometry based approach described in section 3.1 uses information from the geometry of the skin taken from acquired medical imaging data, in this case CT data, and the head pose of the wearer of the HMD in order to define the transparency of pixels.

The outer anatomy should have a rigid nature to effectively use the information from skin geometry that has to be present in the scan data. This is the case for example, in minimally invasive procedures neurosurgery or spine surgery (dorsal approaches). However, in many cases the geometry of the anatomy is modified between the scanning procedure and the operation. For example in open surgery the skin geometry completely changes. When the visualization of deep seated anatomy regions is required, the scan volume often does not cover enough skin to make the data based approach applicable. Alternative imaging modalities, for example data from nuclear probes [240], provide information of only small parts of the 3D anatomy, e.g. the vessel structure. In these cases the video based approach described in section 3.2 has to be applied to enhance the quality of embedding virtual anatomy into the patient.

The video based method follows the approaches described by Kalkofen et al. [111], Stoyanov et al. [209] and Avery et al. [11] using exclusively information from video data of video based AR systems to generate context information. The new approach augments previous methods with additional visual cues reacting on changing light and color conditions, deformation, and interaction with surgical instruments.

When comparing both methods, the geometry based one uses a much wider range of information sources to effectively design the context layer. For example the 3D information of the skin allows to install a virtual window on the patient that can remain at its position to enhance the depth cues motion parallax and occlusion when looking through the window into the body. However, the video based approach is better applicable in surgical scenarios in which the target anatomy is deformable or an open approach is chosen.

3.1 Geometry Based

This section presents a geometry based approach to achieve *contextual in-situ visualization* in medical environments. The transparency of the video images depends on the geometry of the patient's skin surface and the viewing geometry of the observer wearing the HMD. Integrated surgical instruments to be guided by the user within his field of view reactively influence the visualization. The effectiveness of the approach is demonstrated with a cadaver study (see Fig. 3.2(b)) and a thorax phantom, both visualizing the anatomical region around the spinal column, and an in-vivo study augmenting the head of a patient, with her skull (see Fig. 3.2(a)). The visualization method divides the skin surface into two domains **TransDom** and **OpaDom**.

- **TransDom**: The semi-transparent region of the skin defining the vision channel into the body.
- **OpaDom**: The opaque skin outside the vision channel.

The vision channel to the inside of the patient can be interactively defined by the motion of the observer's head who is wearing the HMD. The center of the vision channel follows the line of sight until the user logs its position and the vision channel becomes a virtual

(a) In-vivo study

(b) Ex-vivo study

Figure 3.2: Views on the AR setups that have been arranged to create the examples of *contextual in-situ visualization*.

window installed on the skin surface. While moving relative to the window the strongest depth cues in personal space [49] occlusion and motion parallax support the correction of depth perception.

While the virtual anatomy within **OpaDom** is completely occluded by the skin, within **TransDom** a adaptive transparency pattern partially occludes the deep seated anatomic structure generating little but effective context information. The selection of the remaining fragments on the skin is performed by the three functions described in sections 3.1.1.1, 3.1.1.2, 3.1.1.3. These functions (see Fig. 3.3) take geometrical properties of the skin surface and the observer's head pose as input parameters to compute the transparency map within **TransDom**. Three different parameters are taken into account to calculate the transparency (respectively opacity) of a pixel in the video image:

1. The curvature within the region around a fragment on the skin surface,
2. the angle between the fragments normal and the vector from the position of the fragment to the cameras of the HMD and
3. the distance of a fragment on the skin surface to the intersection point of the line of sight with the skin surface.

For the calculation of the transparency map storing a transparency value for each single pixel of the video image, the skin surface is extracted from volumetric imaging data and rendered off-screen. A fragment shader program then executes the mentioned three functions to compute transparency values that are stored to a designated texture called transparency map (render-to-texture approach). This map is used to blend the video image with the previously rendered virtual entities of the scene with simple alpha blending in this case with the OpenGL blending function:
glBlendFunc(GL_SRC_ALPHA, GL_ONE_MINUS_SRC_ALPHA).

3.1.1 Method

The following paragraphs describe in detail the three functions.

3.1.1.1 Curvature

The first parameter curvature (**CURV**) (see Fig. 3.3(a)) refers to the curvature around a particular 3D position of a fragment on the surface model of the skin [126]. A curvature value is determined for every vertex of the surface model of the skin and interpolated for every fragment during the rasterization process in the rendering pipeline. The curvature value for a vertex is calculated by comparing the normal vector $\vec{n}$ of the vertex with the normal vectors $\vec{n_i}$ of all neighboring vertices. N represents the set of normals of the neighboring vertices.

$$CURV = 1 - \left(1 - \frac{\sum_{n_i \in N} ||\vec{n} - \vec{n_i}||}{2 \cdot |N|}\right)^{\alpha} \tag{3.1}$$

The sum of length values resulting from $||\vec{n} - \vec{n_i}||$ are calculated of every neighboring normal $\vec{n_i}$, whereas one length value corresponds to the angle between two normals. High length values implicate a high curvature value within the region of the vertex position. The division of the sum by two-times $|N|$ clamps the curvature value to the range of $[0, 1]$. $\alpha \in \mathbb{R}_+$ can be modified interactively by the user to adjust the parameter **CURV**. Increasing α approximates the parameters **CURV** to 1, decreasing approximates it to 0. While increasing α, regions with lower curvature contribute more to the opacity of the pixel. Thanks to the parameter **CURV** regions of the skin providing high curvature become more opaque than regions with low curvature. For this reason wrinkled, bumpy and sinuous regions on the skin can still be seen, while flat parts become transparent. A high weighting of the parameter can be suitable for the visualization within the face regions as anatomy provides high curvature due to nose, eyes and mouth (see figures 3.3 and 3.4). The calculation of the parameter has to be done only once, since the direction of normals on the surface model of the skin does not change.

3.1.1.2 Angle of Incidence Factor

The second parameter angle of incidence factor (**ANGLE**)(see Fig. 3.3(b)) is defined as a function of the vector pointing from the position of the fragment to the eye $\vec{v}$ (view vector) and the normal $\vec{n}$ on the skin surface. It provides contour information, since it calculates the dot-product of the normal and the view vector.

$$ANGLE = 1 - (\vec{n} \cdot \vec{v})^{\beta} \tag{3.2}$$

For approximately parallel vectors $\vec{n}$ and $\vec{v}$, this parameter is close to 0 and results in low opacity. For almost perpendicular vectors $\vec{n}$ and $\vec{v}$ the parameter converges to 1 resulting in high opacity. Similar to the **CURV**, this parameter can be adjusted interactively by $\beta \in \mathbb{R}_+$. The parameter **ANGLE** is modified by the observer wearing the HMD and defining the camera positions and consequently the view vector for every fragment on the skin surface. For this reason patches of the skin facing directly to the observer appear more transparent than others. In contrast to **CURV** the parameter **ANGLE** is view-dependent.

3.1.1.3 Distance Falloff

The third parameter distance falloff (**DIST**) (see Fig. 3.3(c)) is a function of the distance between each surface fragment and the intersection point of the line of sight and the skin surface. The parameter can be calculated using any monotonically increasing falloff function. The transparency of skin patches decreases with respect to the distance to the intersection point. Skin patches having a greater distance to the intersection point than a chosen maximum distance are rendered completely opaque. Thus this parameter divides the skin in the **TransDom** and **OpaDom**, whereas the transparent region can be considered as the vision channel to the region of interest inside the body. The falloff function leads to a smooth border between the domains **TransDom** and **OpaDom**.

$$DIST = \left(saturate(\frac{distToViewIntersecPoint}{radiusOfTransparentRegion}) \right)^{\gamma} \tag{3.3}$$

Saturate is a function that clamps its input parameter to $[0, 1]$. The falloff is calculated by a division of the distance to the intersection point of the line of sight with the skin surface by the defined radius of the **TransDom**. The parameter **DIST** implements a restriction of the observer's vision channel to the important region of the virtual entities. The adjustable exponent $\gamma \in \mathbb{R}_+$ defines the smoothness of the transition between the domains **TransDom** and **OpaDom**. A smaller value for γ results in a smoother transition and a higher degree of opaqueness close to the intersection point of the line of sight with the skin surface. Vice versa, a high value for γ results in a sharp transition of the domains and a low degree of opaqueness close to the intersection point.

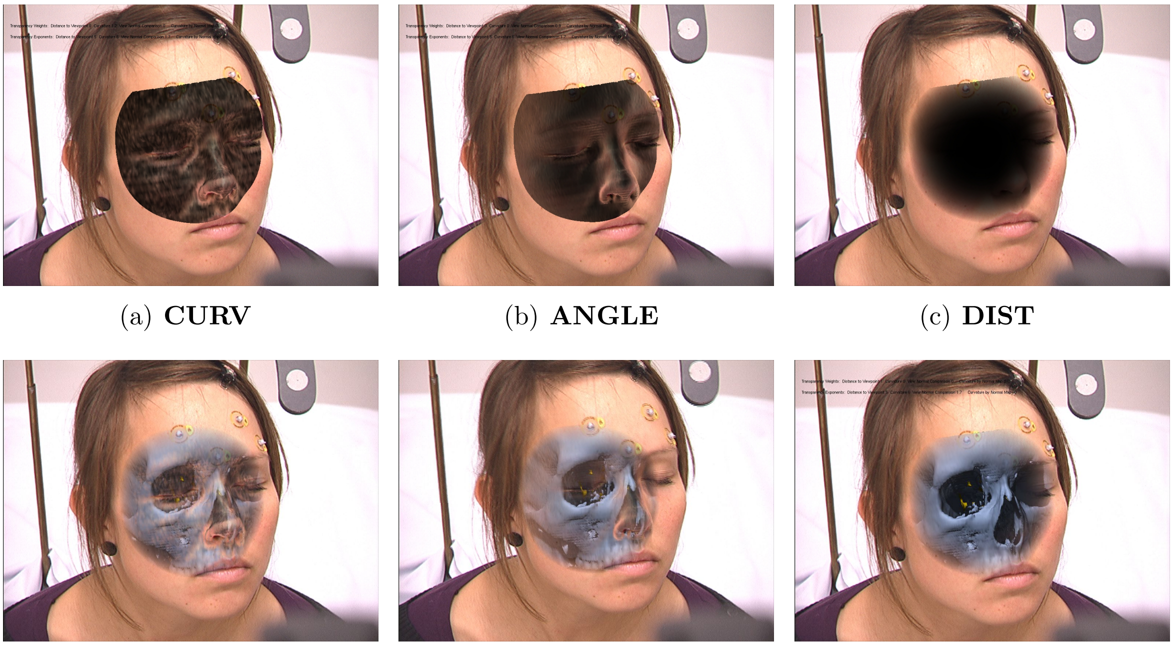
3.1.1.4 Final Transparency/Opacity Determination

In order to calculate the opacity value of a pixel on the video image, the maximum of the three parameters is determined. In addition to the respective exponents α, β and γ, the parameters can be interactively weighted by $w_1, w_2, w_3 \in \mathbb{R}^+$.

$$opacity = saturate(max(w_1 \cdot CURV, w_2 \cdot ANGLE, w_3 \cdot DIST)) \tag{3.4}$$

The resulting opacity value is then stored to the mentioned transparency map, which defines the alpha value of each pixel of the incoming video image. Finally the colors of the previously rendered virtual objects are blended with the manipulated video image.

In order to more effectively attract the user's attention to the region of interest and to increase the depth cues occlusion and motion parallax to correct depth perception, the border between **TransDom** and **OpaDom** can be enhanced with a colored line on the skin surface having user-defined thickness [126]. For this purpose the distance of a pixel to the intersection point of the line of sight and the skin surface, which was already determined for calculating the **DIST** parameter, is used. The border is rendered by setting the red fraction of colors of those fragments to 1 having a distance to the intersection point within the range $[(1 - thickness) \cdot maxdist, maxdist]$. *thickness* denotes a user-defined thickness value for the border and *maxdist* denotes the user-defined radius of **TransDom**.



(a) **CURV** (b) **ANGLE** (c) **DIST**

(d) **CURV** with anatomy (e) **ANGLE** with anatomy (f) **DIST** with anatomy

Figure 3.3: Figures 3.3(a), 3.3(b), 3.3(c) show the three parameters without any virtual objects in the black background. Figures 3.3(d), 3.3(e), 3.3(f) show the transparency map blended over a virtual skull.

3.1.1.5 Interactive View

Endoscopic instruments commonly enter the body through so-called trocars. Trocars are small pipes inserted into a small cut through covering tissue in order to protect the skin and other tissue from further lesion while moving the instrument. The proposed method integrates virtually extended endoscopic instruments that become visible only inside the body. This approach clearly distinguishes the position of the penetration through the skin surface. This is not the case, when a virtual counterpart of the real instrument is rendered inside as well as outside the body as shown in Fig. 3.1(b). In addition, the method employs the virtual shadow to provide visual feedback for user interaction and enable additional depth cues. For this reason, surgical instruments are virtually extended as soon as they penetrate into the patient's body.

The proposed method requires the instrument to be rendered twice after having rendered the skin surface. In the first rendering pass, drawing to the color buffer is disabled and stencil test is performed when the instrument is rendered. The stencil [1] bit is set for those fragments affected by the instrument that fail the depth test. In other words, pixels from the instrument laying beneath the skin surface. In the second rendering pass of the instrument, a second stencil test allows to store pixel values from the instrument to the color buffer, if the stencil bit is set on their respective screen position.

[1] The stencil buffer is an additional buffer beside the color buffer and depth buffer on modern computer graphics hardware.

(a) Composition of the three functions (b) Skin augmentation (c) Enhanced **ANGLE**

Figure 3.4: Figure 3.4(a) shows a sophisticated composition of all three functions. Figure 3.4(b) shows the augmentation of the polygonal skin surface extracted from CT data below the semi-transparent video texture. Figure 3.4(c) shows a perspective from above enhancing the **ANGLE** since the angle between view vectors and normals increases.

The transparency settings of the video image around the penetrating instrument can be adjusted using the previously calculated intersection point of the instrument and the skin surface. Within the video images, the region around the port remains opaque, so that the real instrument entering the patient is not occluded by the visualized anatomy. The transparency of pixels within this region depends on the distance to the estimated intersection point. For this reason, a fourth function **INST** is used for transparency calculation of the skin around the region of the trocar, which is decreasing linearly with respect to the distance to the intersection point:

$$INST = 1 - \frac{distanceToInstrIntersecPoint}{maxdist} \tag{3.5}$$

In addition, to phong shading for the illumination of the virtual anatomy, a soft shadow mapping algorithm was integrated to provide improved spatial information of virtual objects [226]. The endoscopic instrument or any other virtual object can cast shadow onto the anatomy and therefore causes visual feedback from user interaction.

3.1.1.6 Visualization Pipeline

This section summarizes all necessary rendering steps for the present visualization technique. Figure 3.5 provides a schematic overview of the functions and their parameters to compute the transparency map.

First, **CURV** is computed for each vertex. This step has to be done only once, since the curvature is independent from other parameters, for instance the line of sight of the observer. For the final calculation of the transparency map the following parameters are taken into account:

- The center point of transparent region **TransDom** which can be defined either by the intersection point of the line of sight from HMD pose data and the skin surface when the the vision channel is not fixed to a certain position or by a stored position on the skin surface.

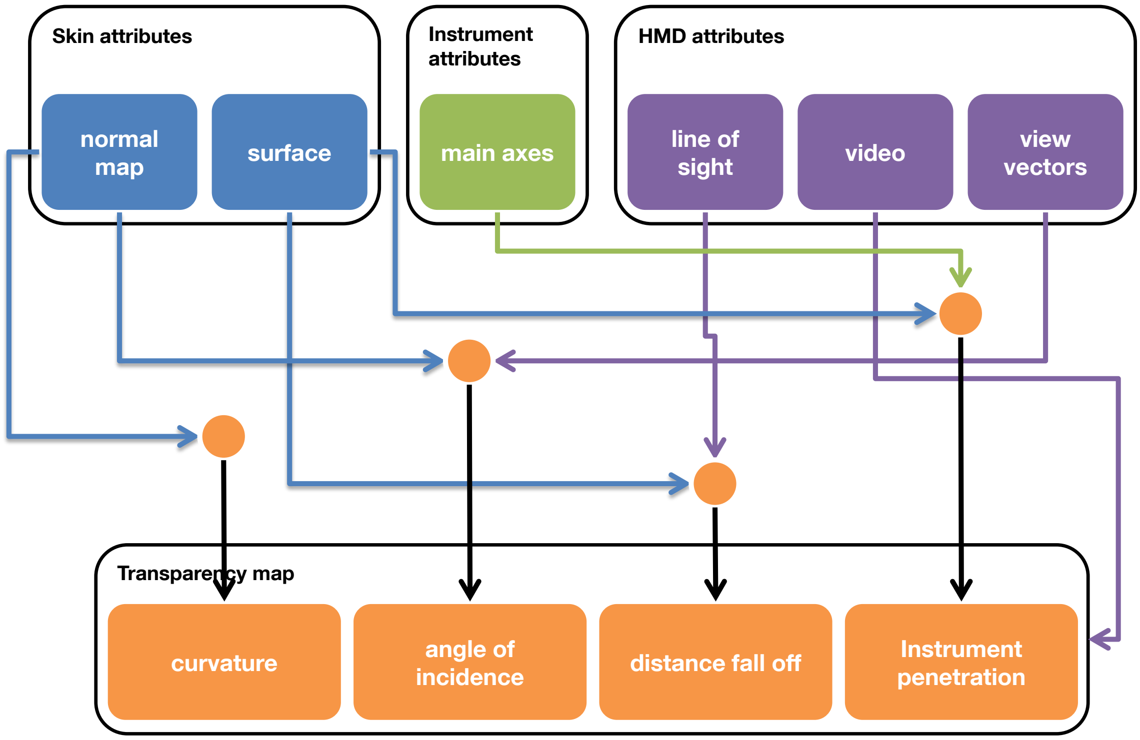


Figure 3.5: Dataflow diagramm to compute the transparency map

- The maximum distance to the center point of the transparent region, which defines the size of the vision channel (**TransDom**).
- The previously explained weighted results of the functions **DIST**, **CURV**, **ANGLE**.
- The intersection point of the skin surface and the instrument main axes, if an instrument is used and penetrated through the skin.
- The border size of the vision channel.

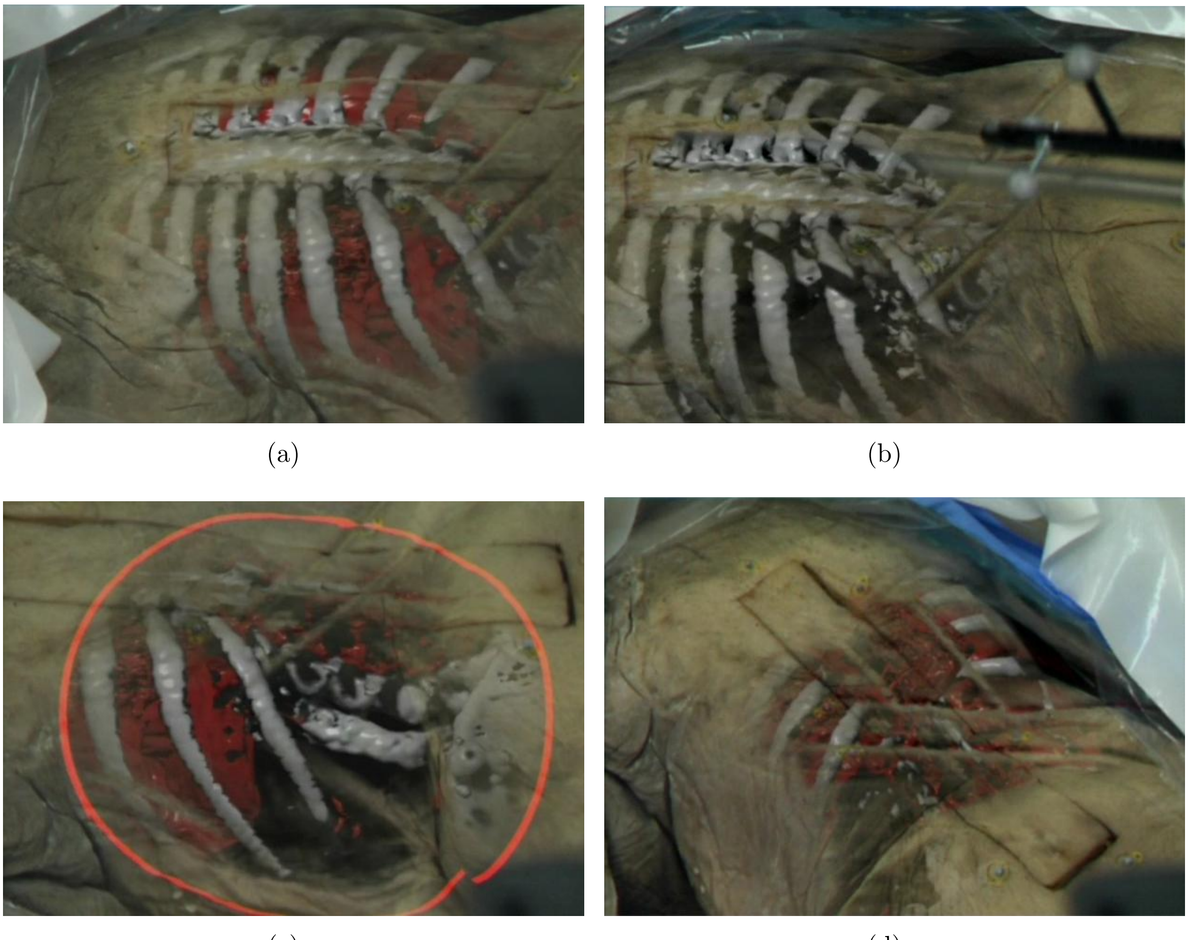
For rendering the instrument to the virtual scene a stencil test is performed by drawing the instrument and the skin surface. The instrument is only rendered to the frame buffer if the z-test fails, which means that only fragments of the instrument are considered that are behind the skin surface. Then the shadow map is calculated by rendering offline the instrument and virtual anatomy from the view geometry of the light source. Polygons are shaded using phong illumination and appliance of the calculated shadow map. Finally the alpha values of the video texture are manipulated by the transparency map and the resulting semi-transparent video texture is blended over the virtual scene.

Figure 3.3 shows all functions to compute the transparency map except the one depending on the instrument penetration.

3.1.2 Evaluation

The effectiveness of the presented visualization technique is demonstrated with the resulting images of a cadaver study, a thorax phantom and an in-vivo study augmenting the head of a patient.

A CT scan was taken of a cadaver starting from the pelvis and ending at cervical vertebrae. Before data acquisition, CT markers were attached to the skin around the region of the spinal column to automatically register the visualization to the body. Figures 3.6 show the results of the cadaver study. We did not insert a surgical instrument into the body, however, figure 3.6(b) shows the shadow cast onto visualized ribs caused by a user guided surgical instrument and a virtual light source positioned above the scene. Figure 3.6(a) shows an expedient combination of the three transparency parameters including the highlighted borderline to create a virtual window. Images of an augmented thorax



(a) (b) (c) (d)

Figure 3.6: Spinal column and ribs can be visualized with a suitable transparency map for the video images (see Fig. 3.6(a)). Shadow is cast by a surgical instrument onto the virtual ribs as shown in Fig. 3.6(b). The instrument was not inserted into the body during this study. In addition the lung is visualized in combination with the lumbar vertebrae in figure 3.6(c) and partially removed ribs using clipping volumes in figure 3.6(d).

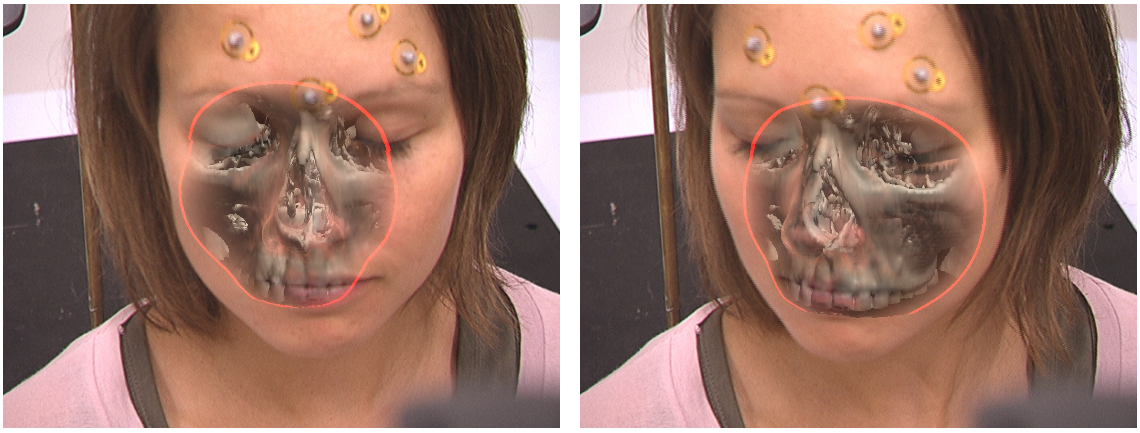
phantom showcase the integration of the surgical instrument for keyhole surgery. Originally the phantom consisted only of a spinal column installed inside the phantom. However, we extended the phantom's virtual anatomy by surface models segmented from the Visible Korean Human (VKH) (see section 2.2) data set. Virtual models are registered manually to the thorax phantom. This is sufficient for the evaluation of our visualization method. The region of the skin around the port for inserting the surgical instrument into the phantom gets opaque when the instrument is penetrating the body. Thanks to this effect, the view on the port is not restricted since it is not occluded by the virtual anatomy (see figures 3.9). The instrument and the anatomical surface models cast shadow on each other and thus provide helpful visual feedback while navigating the instrument to the operation site. Furthermore, the surgical instrument is virtually extended only inside the phantom.

We also had an opportunity to conduct a an in-vivo study. Susanne [2], had a ski accident and got a CT scan of parts of her head. She is fine again and agreed to participate in the study. The face structure provides high curvature due to nose, mouth and eyes and therefore is perfect to showcase the effectiveness of the present visualization method. At the day of the accident, no CT markers were attached before the CT scan. For this reason, visualized anatomy was registered manually with her head using natural landmarks. Fig. 3.7 shows different combinations of the three parameters to calculate the transparency map. The vision channel through the face surface follows the line of sight of the observer. Users can intuitively navigate the extended field of view to observe the patient and determine a reasonable visual port to the operation site (figures 3.7). When the region of interest is defined, the vision channel can be locked. It then turns to a virtual window overlaid on the skin surface. The window borders can be highlighted and broadened to enhance the visual cues occlusion and motion parallax for improved depth perception while the observer is moving around the virtual window. Also the patient can be moved while the virtual window remains at its previously defined position. We integrated a tracked spoon into the scene to interact with the in-situ visualization. In order to do this, we created an exact virtual counterpart of the spoon to visualize it also beneath the skin surface and to cast a shadow on visualized anatomy. We do not insert the spoon into the mouth because the data set did not include this region. The spoon was positioned under the skin surface by pressing it against the cheek, which makes the greenish virtual spoon visible (see figure 3.8(a)). Shadow is also cast onto the virtual skull while moving the spoon around the head space (see Fig. 3.8(b) and 3.8(c)).

The first generation of data based *contextual in-situ visualization* is based polygonal surface models to allow for real time visualization. However, surface models have to be segmented, triangulated and smoothed before the visualization can be performed. The quality of the transparency effect, in particular created by the parameters *curvature* and **ANGLE**, strongly depends on the accuracy level of the surface model of the skin. Rendering speed is regulated by the number of processed triangles of the anatomical surface models. For this reason high quality of data visualization suffers from low performance speed. Both factors are essential regarding the acceptance and ease of use for the AR system.

In 2006, one year before the first generation of *contextual in-situ visualization* has been

[2] name changed



(a) Fixed vision channel

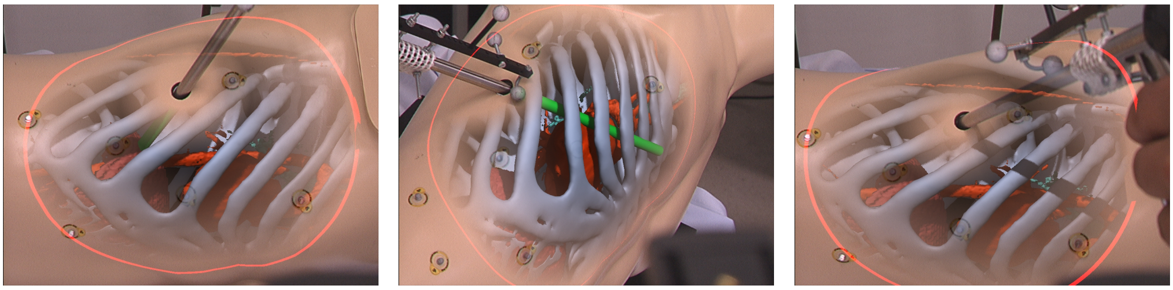
(b) Subject moves while the vision channel remains at its defined position.

Figure 3.7: The vision channel can be locked and turns to a virtual window fixed on the skin surface (see Fig. 3.7(a)). The patient can move relatively to the observer and the vision channel slides over the anatomy (see Fig. 3.7(b)).

proposed [27], Peter states in his report in image guided surgery "The increasing speed of ray-casting algorithms may alleviate this problem somewhat by providing interactive volume visualization without the need for a priori segmentation" [168]. Consequentially, in 2009 a second generation of geometry based *contextual in-situ visualization* has been proposed by Kutter et al. [128]. The GPU accelerated direct volume rendering uses raycasting to achieve the highest resolution for data visualization in real-time. The method follows the visualization principles of the first generation of *contextual in-situ visualization* except the integration of shadow mapping, that would require additional rendering passes. The advanced rendering pipeline incorporates further features such as color filtering to detect blueish surgical gloves in oder to discard and therefore occlude the virtual fragments of the rendered anatomy. The second generation of *contextual in-situ visualization* has

(a) (b) (c)

Figure 3.8: A tracked spoon that is virtually extended inside the body. To position the spoon under the skin surface we press the spoon against the cheek (Fig. 3.8(a)). Shadow is cast by a tracked spoon onto the virtual skull (Fig. 3.8(b) and 3.8(c))



(a) (b) (c)

Figure 3.9: When the instrument is inserted through the port the skin remains opaque within the region of penetration (Fig. 3.9(a) and 3.9(c)). Furthermore objects cast shadow on each other and therefore facilitate the navigation of the instrument (Fig. 3.9(b) and 3.9(c)). The instrument is virtually extended inside the thorax phantom (Fig. 3.9(b)).

been used in the experiment presented in section 5.3.

3.2 Video Based

Many surgeries do not use imaging data that is large enough to cover the skin geometry. Even if the skin geometry is available, one deals with only static context information captured at a certain point in time. For this reason, the skin serving as the major context layer cannot adapt to geometry deformation or color changes.

We decided to follow the approaches of Kalkofen et al. [111], Stoyanov et al. [209] and Avery et al. [11] using exclusively information from video data of video see-through devices to generate context information. We extended these approaches with additional visual cues reacting on changing light and color conditions, deformation, and interaction with surgical instruments. This section presents the rendering pipeline of a video based *contextual in-situ visualization* strongly optimized for surgical AR environments.

3.2.1 Method

Our visualization pipeline installs a transparent vision channel [172, 27] virtually onto the deepest context layer of the AR scene, which is in our case the skin, that is aligned with the line of sight of the HMD user. A transparency fading effect is applied at the border of the vision channel in order to generate a smooth transition from the original skin color of the video to the assigned transparency of the vision channel.

$$AF = ((1.0 - AL)/FR) * (CPVC - VCR + FR) + AL; \tag{3.6}$$

AF describes the alpha value within the region of transparency transition. It is calculated with the parameters of the initial global alpha value inside the vision channel (AL), the width of the concentric disk describing the region of transition (FR), the current position of the pixel within the vision channel being used for the calculation of the alpha value(CPVC), and the vision channel radius (VCR). The proposed rendering pipeline for our method of image composition involves four main components using information from tracking data and video images. The transparency generating the context information within the vision channel is parameterized with the following functions on the video data:

- **MATERIAL1**: A direction invariant Sobel filter for the detection of edges in video images caused by surgical instruments and trocars, surgeons' hands, the open operation site itself, blood drops, birthmarks, scars, cover sheeting, artificially introduced marks by a sterile pen onto the skin.
- **MATERIAL2**: Detection of specular highlights caused by bright metallic light reflecting instruments, humid skin surface, fluids, bright/specific colors in general.

In addition, two combined shadow effects enhance the spatial appearance and provide interactive visual feedback:

- **SHADOW1**: The combination of **MATERIAL1** and **MATERIAL2** is projected shadow-like onto virtual objects behind the skin layer using OpenGL supported projective texture mapping.
- **SHADOW2**: Shadow mapping [241] applied to all virtual objects.

Video data coming from a video capturing device like an HMD, a laparoscope or an arthroscope camera is processed by a GLSL [3] based shader. For modular computation and final composition, Frame Buffer Objects (FBO) for GPU accelerated off-screen rendering have been employed. Regarding **MATERIAL1** (see Fig. 3.10(a)), the transparency (AE) value of each pixel of the transparency map is computed by comparing the gray values (GV) of both the original image pixel (OP) and the extracted edge pixel (EP) and adjusting it with a transparency level (AL):

$$AE = OP.GV/EP.GV + AL \tag{3.7}$$

MATERIAL2 thresholds the gray scale version of the video data to detect brightly colored regions (GV), which remain opaque while darker regions are coded with a high transparency value:

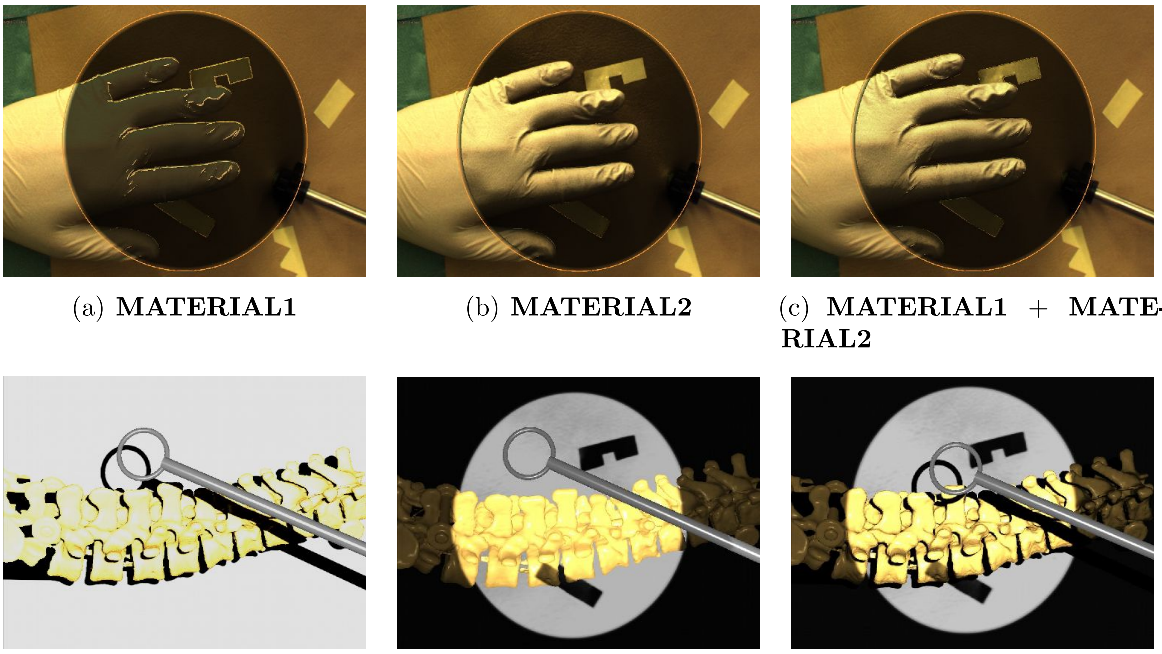
$$AS = AL + 1.0 - (1.0 - OP.GV)/(1.0 - ST) \tag{3.8}$$

The alpha value (AS) of a pixel of the video image uses the following parameters. OP is the original pixel value, AL is again the initial global alpha value inside the vision channel and ST determines the threshold for bright regions (see Fig. 3.10(b)). **MATERIAL1** and **MATERIAL2** are then combined as shown in Fig. 3.10(c).

The result is then forwarded to another shader program performing projective texture mapping [241] to feed **SHADOW1**. **SHADOW1** and **SHADOW2** apply *percentage closer filtering* (PCF) [175], which is supported by the used graphics hardware [4] to create

[3] OpenGL Shading Language

[4] Nvidia™Geforce 8800 Ultra

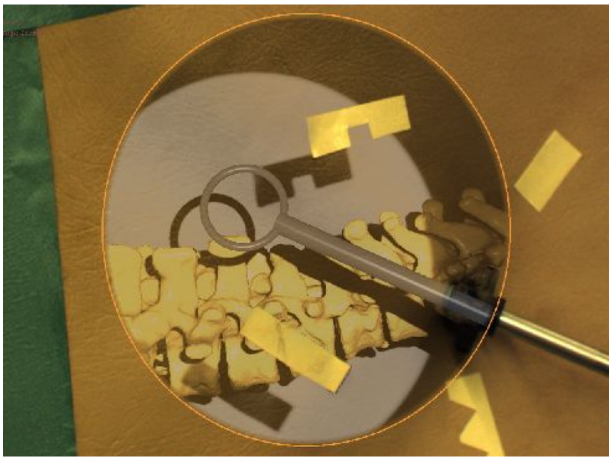


(a) **MATERIAL1** (b) **MATERIAL2** (c) **MATERIAL1 + MATERIAL2**

(d) **SHADOW2** (e) **SHADOW1** (f) **SHADOW1 + SHADOW2**

Figure 3.10: Intermediate steps of the pipeline described with screenshots of an exemplary AR scene.

smooth shadow and avoid aliasing artifacts (see Fig. 3.10(d)). The method of combining **SHADOW1** and **SHADOW2** follows a natural shadowing principle based on *diffusion* and *interference*. The blurriness of a shadow increases and the contrast decreases in proportion to an increase of the distance between a light source and a fixed object in space. According to our assumption that 3D objects are located behind the context layer, features filtered by **MATERIAL1** and **MATERIAL2** can be projected with **SHADOW1** onto these objects so that their shadows will be sharper and have more contrast. Blurriness can be adjusted with PCF kernels and both, **SHADOW1** and **SHADOW2**, can be weighted. However, blurring and weighting of **SHADOW1** is based on an approximation of the distance between the skin surface and virtual objects behind the skin since the 3D pose of the skin is not known. The light for calculating **SHADOW1** and **SHADOW2** is positioned relative to the HMD of the user in our case $20cm$ to the right.
At the end of the rendering pipeline, we deal with two textures called *Processed Video* and *Final Virtual Scene*. *Processed Video* represents the video data with a previously individually computed alpha value for each pixel within the vision channel. The other texture contains all virtual 3D information including the composition of **SHADOW1** and **SHADOW2**. *Processed Video* and *Final Virtual Scene* are then composed with OpenGL blending and rendered into the Frame Buffer resulting in the final AR scene (see Fig. 3.11(a)). Figures 3.12 and 3.10 provide an overview of all rendering steps and render targets.



(a) Final AR scene to be drawn to the frame buffer.

(b) One of 26 subjects.

Figure 3.11: Finally composed AR scene and experimental setup.

3.2.2 Evaluation

Figure 3.13(a) shows an endoscopic instrument penetrating a previously inserted trocar above a simulated operation site for key hole surgery. The skin around the trocar shows in this case only homogeneous color distribution providing almost no information for **MATERIAL1** and **MATERIAL2**. However, guidance of instruments causes deformation of the skin and influences the color brightness of the skin. This reactively changing light reflection can be detected by **MATERIAL2** to generate reactive transparency changes. In order to feed **MATERIAL1**, we attached small glue strips having a slightly different color than the skin (see Fig. 3.13(a)). For intra operative surgical planning in some cases sterile markers are used to draw rough sketches of the hidden anatomy onto the patient's skin. **MATERIAL1** and **MATERIAL2** can be configured to detect also the color of such sterile markers to enable additional context information. Alternatively light patterns can be projected onto the skin. Figure 3.13(c) shows a hand held flashlight causing a bright spot on the skin that is detected by **MATERIAL2**. This sterile light spot again causes partial opacity of the vision channel. In addition, its shadow is thrown onto the virtual anatomy and a virtual plane in the background of the scene. **SHADOW1** can produce shadow from any object that is shown in the video image. This includes the surgeon's hand (see Fig. 3.13(d)) and surgical instruments such as a scalpel (see Fig. 3.13(b)) or the endoscopic instrument with its attached marker tree for optical tracking. The effect of masking the scene with the user's hand has been previously shown with distinguishable blueish gloves having an exclusive color in the AR scene [128]. The current configuration of **MATERIAL1** can also deal with skin colored gloves that are mainly used in todays ORs (see Fig. 3.10(c)).

The generated graphical effects from **MATERIAL1** and **MATERIAL2** enable the strongest depth cue in personal space, which is occlusion [49]. Also the strong depth cue motion parallax can be generated to enhance perception when the transparency is calculated independent to the head motion of the observer. The methods works robust

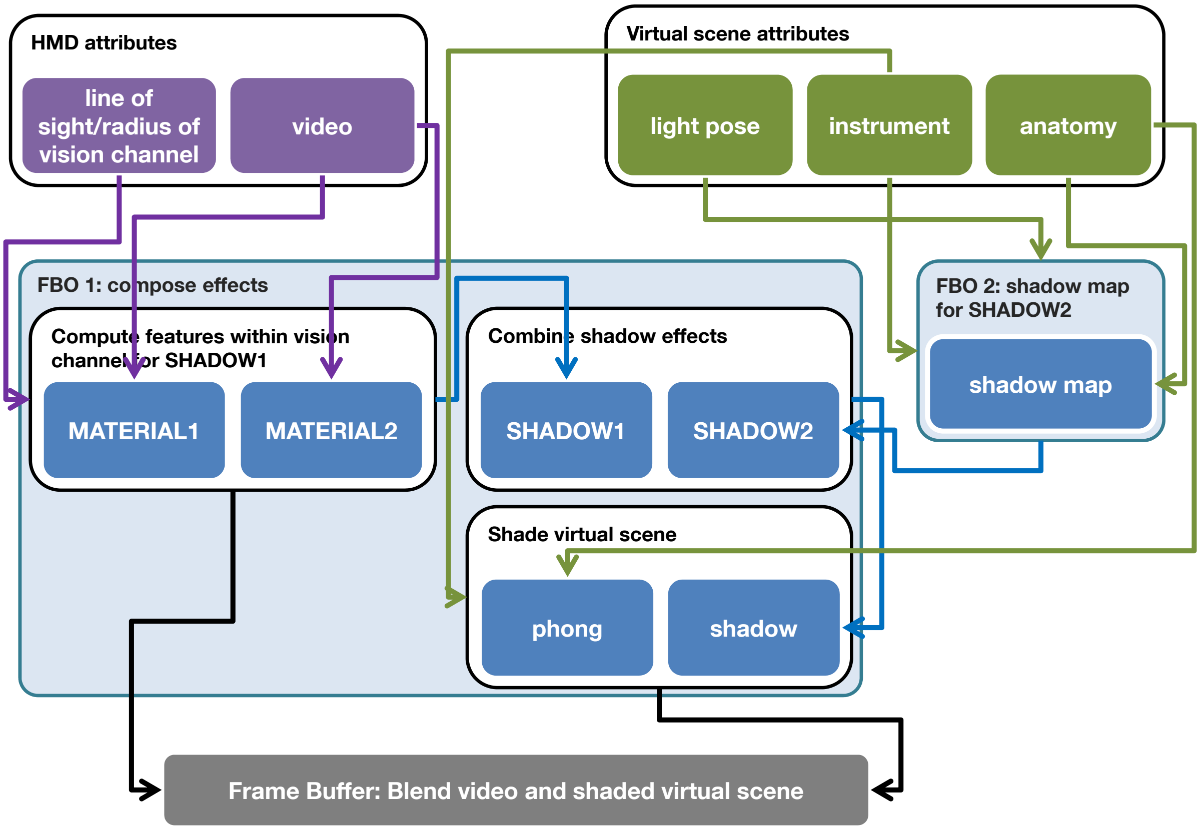
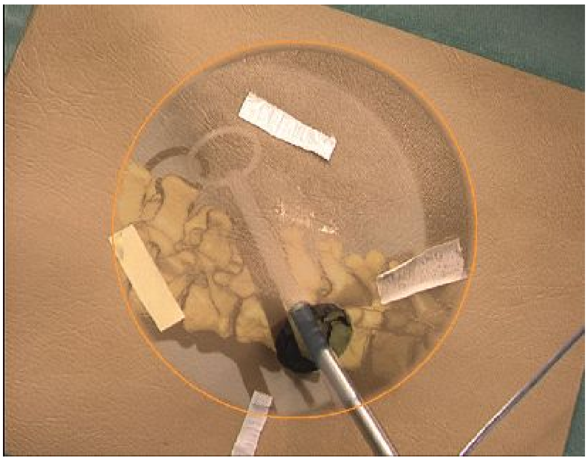


Figure 3.12: Rendering pipeline of video based *contextual in-situ visualization*

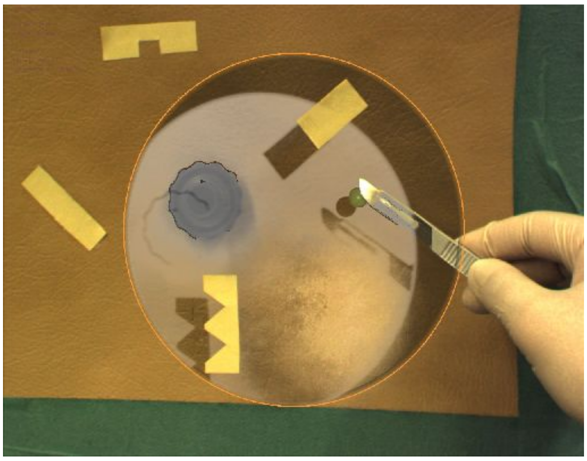
even though the illumination, color and geometry conditions completely change during the surgical procedure, e.g. the skin is cut or deformed, minimally invasive trocars are installed or an open operation is being performed. The effects **SHADOW1** and **SHADOW2** create shadow cast, which can improve the 3D layout of the AR scene. Experiments [116, 139] have shown that shadows can have a strong impact on the perception, in particular when motions come into account. Regarding the surgical task, shadow gives visual feedback that can be intuitively interpreted when the operator starts to interact with the AR scene.

3.3 Discussion

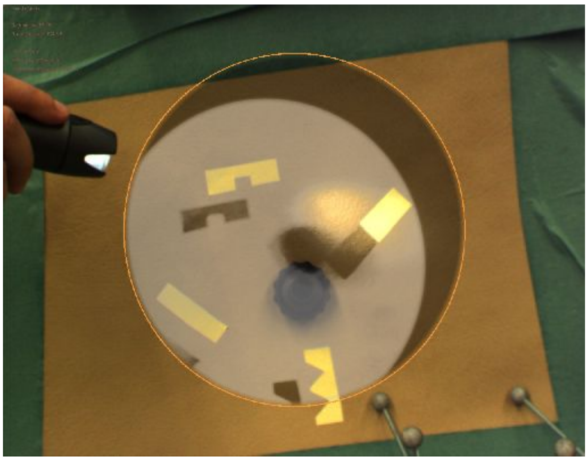
The proposed methods for *contextual in-situ visualization* can be further improved. In particular the combination of features from both approaches can result in a more flexible approach. In particular the features **MATERIAL1**, **MATERIAL2**, **SHADOW1** and **SHADOW2** can complement the geometry based approach. Video features such as edge patterns can augment the generation of the transparency map. For instance light changes or scars on the skin can be extracted to create further depth cues from natural features of



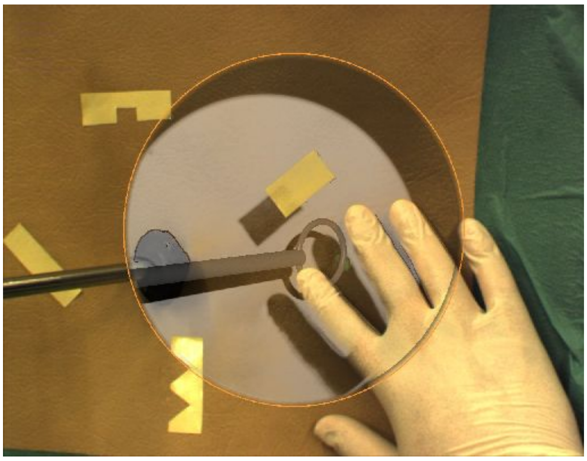
(a) Deformation of the skin



(b) Extracted instruments cast shadow



(c) Flashlight causes context information.



(d) Hand causes shadow feedback.

Figure 3.13: Different features of the proposed image composition.

the AR scene.

Regarding the video based approach, it would be interesting to dynamically adapt the parameters of **MATERIAL1** and **MATERIAL2** according to an online video analysis. One may think of an algorithm that creates a histogram of the incoming video images. Analysis of the histogram over time can result in a special transfer function, which can be used to define a sophisticated and self adjusting transparency map of the video image. On one hand this transparency map must not occlude important information and on the other hand it should provide enough texture pattern from fragments of the video to enable depth cues as proposed with both methods of *contextual in-situ visualization*.

Another idea involves the usage of light that is invisible for the human eyes to affect **MATERIAL2**. Fig. 3.13(c) shows the generation of depth cues using a light source changing the illumination of the context layer, i.e. the skin. Instead of light being visible for the humans eye, infrared light can be used to illuminate the skin for **MATERIAL2**. Infrared light is reflected by the skin and is visible for cameras that are not equipped with infrared filters. In this case, the homogeneous illumination conditions in ORs are not disturbed, which would reduce the acceptance of the system.

As stated before the geometry based approach of *contextual in-situ visualization* works fine for anatomy that is hardly deformed during operation. The geometric properties being primarily used only correspond to a certain point in time, however do not adapt to changing topology. A future project could involve real-time deformation tracking using stereo color cameras that use light patterns projected onto the human body to adapt the geometric parameters. Alternatively an additional time-of-flight (TOF) camera could be used. TOF cameras will revolutionize the many fields of research in computer vision. It is imaginable that in a future version of the video see-through HMD the third infrared tracking camera will be replaced by a TOF camera as soon as these devices become smaller and lighter and provide a higher resolution image and depth accuracy.

Both approaches have been implemented in cooperation with the students Felix Wimmer and Maxim Kipot within the scope of their master theses [242] and [121] that have been supervised by the author of this thesis. Felix Wimmer has extended the initial idea of using ghosting techniques (see section 3.2) with a new approach that enables AR-driven cutaway visualization [243].

CHAPTER

FOUR

NOVEL INTERACTION TECHNIQUES FOR MEDICAL AUGMENTED REALITY

Human-Computer Interaction (HCI) is a multidisciplinary field of research incorporating knowledge from computer science, psychology, ergonomics, cognitive science and other scientific fields according to the target application of the particular system. Likewise other specialized working environments, the operating theater (OR) as the operating environment of the targeted AR system has specific requirements that have to be addressed. In particular the highly optimized space organization as well as the sterility regulation within the personal space of the surgeon pose an interesting challenge. According to Peters "It ultimately must be possible to operate computer systems in the OR without the use of a keyboard or complicated switching devices." [168]. With this principle in mind, a set of interactive user interfaces to control the AR environment has been developed that come up with the conditions of an intraoperative working space. Section 4.1 addresses the control of logic and continuous parameters of the AR scene from outside using pedal, hand gestures or voice control. Section 4.2 introduces a virtual mirror as a virtual media being part of the AR scene to interact with the scene. In the latter case, interaction rather affects the perception of the AR environment, for example enabling additional views on virtual objects using the mirror reflection.

4.1 Control of Intraoperative Devices

Current interaction devices within surgical environments consist of delegated control, keyboard and mouse, foot pedals, touch screens and more rarely speech recognition. Grange et. al. [87] report that delegated control, especially when affecting the change of continuous or multi-dimensional values, may lead to large temporal delays and deviations from the regular workflow of the procedure. Although mice and keyboards may be built to be sterilizable [103], they can be distracting and require delegated control for proper usage. Foot pedals, despite being commonly used in ORs, may lead to problems like losing the contact with the pedal, hitting the wrong switch or physical discomfort. In addition, the foot pedal obstructs the freedom of movement of the surgeon [227]. Touch screens, as in the BrainLAB VectorVision system, are used in Computed Assisted Surgery (CAS) systems, because they provide an intuitive user interface and don't occupy much space in the OR.

Speech recognition is an inherently sterile user interaction scheme and does not require the surgeon to look away from the current operating site. On the other hand, speech recognition is not appropriate when the input data is geometric or has some continuous value. Alapetite et al. [5] showed fail-proof speech recognition when suitable microphones are used.

The *M/ORIS project*[87] proposes gesture interpretation for touchless human computer interaction. With fixed stereo cameras, M/ORIS creates a depth map and rough 3D reconstruction of the scene for separation of the surgeon from the background [63]. Hand and head detection from the depth information, the surgeon's silhouette and color information are used to control a non-contact mouse by moving the hand inside a certain area. Wachs et. al.[232] propose gesture recognition system to build a sterile medical MRI image browser. The system itself uses color segmentation, background subtraction and the CAMShift algorithm to detect and track the hand.

The *FAce MOUSe* project [158] uses a surgeon's head to control the position of a laparoscope. Samset et al.[186] describe a surgical user interface based on the concept of occlusion.

Mine et al. have explored various methods for moving objects in virtual environments using the virtual proprioception, or a person's sense of the position and orientation of his body and limbs. One of their approaches is based on defining a *pick ray* [152], which is fixed to the head and propagates along the direction of view. The menu item intersecting with the pick ray is then selected. This gives a quite intuitive way of selecting items by simply looking at them.

For the assessment of the quality of user interfaces, various standardized questionnaires have been developed. For evaluating the usability of three interaction modes that have been integrated in our HMD based AR system, we use the standard questionnaire USE (Usefulness, Satisfaction, and Ease of Use) proposed by Lund [137]. USE consists of standardized questions (0 = strongly disagree to 6 = strongly agree) assigned to the four categories Usefulness, Ease of Use, Ease of Learning and Satisfaction. The USE questionnaire has been chosen because it has the intention to assess the quality for a wide range of user interfaces. Other common standard questionnaires do not cover the interlaced structure of AR systems consisting of hardware and software modules. For example SUMI rather covers the analysis of desktop software or software applications and MUMMS as an advancement of SUMI addresses also multi-media software.

The complete USE questionnaire is attached to the appendix of this thesis B.1.

4.1.1 Three AR UIs for Video See-Through Head Mounted Displays

Together with the trauma surgeons of our team, we defined a set of conditions that have to be fulfilled by the interactive interface to become useful and acceptable for its intraoperative application.

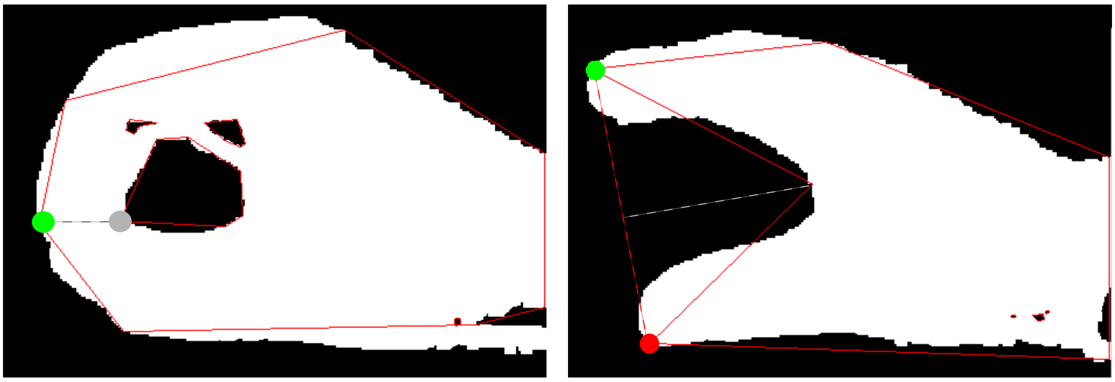
- The interface has to offer control of binary as well as continuous parameters. Interactions comparable to pressing buttons or shifting a slide bar should be controllable by the interface.
- The operator himself should be able to control the parameters without getting help from an assistant.
- The interaction procedure should be easy, intuitive and learned quickly to avoid loosing control in stressful situations.
- The interface has to be integrated into the working environment without wasting space and jeopardizing the sterile area around the operation site.
- Control of parameters has to be robust and unambiguous.

Derived from these conditions and inspired by previous work proposed in the literature, we developed a set of three promising interaction methods (AR UIs) to be used with the HMD based AR system. These AR UIs have been implemented in cooperation with the student Samuel Kerschbaumer within the scope of his master thesis [115] that has been supervised by the author of this thesis.

1. **PEDAL: Combination of head motion and foot pedal.** The user controls 3D graphical UI objects positioned in 3D space close to the region of interest within the AR scene with his ray of sight serving as a courser. In addition, a foot pedal is pressed for activation and control of virtual buttons and sliders.
2. **VOICE: Combination of head motion and voice recognition.** The user again controls 3D graphical UI objects positioned in 3D space with his ray of sight (courser), similar to **PEDAL**. Activation of objects is performed with univocal commands.
3. **HAND: Hand and gesture recognition.** One hand is moved in front of the HMD cameras in order to select objects from a 2D graphical UI drawn into the projection plane of the display in the foreground of the scene. Activation is performed with the gesture of a closed hand. The courser corresponds here to the tips of the touching thumb and index finger.

For the PEDAL based interactive interface, a *SpeedLink USB Racing Wheel* with foot pedals is used. The interaction with the pedal works through the Human Interaction Device interface of the *Microsoft Driver Development Kit*. To simplify the interaction, only one button of the pedal is used. To start the interaction and avoid unwanted activations, the user has to press the pedal for one second and then release it. Similar to a normal mouse button, the surgeon can use the pedal to click, drag and drop, after the user interface is activated.

For the VOICE based interactive interface, The *Microsoft Speech API* is used for speech recognition. The API allows the programmer to create its own grammar file, which specifies the sentences the system can understand. A limited set of phrases is used, which allows robust recognition rates.



(a) Closed hand with inner contour. (b) Open hand with convexity defect.

Figure 4.1: Using HAND to interact with HMD-based in-situ visualization.

When interacting with the interaction mode HAND operators wear distinctively *colored gloves* to guarantee robust and easy segmentation of their hands. This approach is also used by Kutter et al. [128] to increase the quality of depth perception by masking the virtual part of the scene after the segmentation of the hand area with a one-dimensional histogram analyzing the hue color information. Noise in the binary segmentation is reduced using morphological operations, namely erosion followed by dilation. The contour of the hand shape is determined with the *Douglas-Peucker algorithm* [174, 54]. *Convexity defects* [100] in the contour are calculated by computing the convex hull for an object. The contour is then compactly represented as a *Freeman Chain Code* [76]. In order to make the detection more robust, we track the contour over time using the *continuously adaptive mean shift* (CAMSHIFT) algorithm [36]. For detecting the *open hand* (Fig. 4.1(b)), several heuristics are applied to the contour of the segmented hand. The contour area is checked for a reasonable minimum size and for convexity defects of the shape. The open hand can only have one big convexity defect, which has to be on the left side of the contour. The start and end points of the convexity defect need to have approximately the same x coordinate and the point with the biggest distance to the convex hull has to lie vertically in the middle of the start and end point. To detect the fingertip, the whole contour is rotated counterclockwise by 20 degree and the leftmost point is taken. For detecting the *closed hand* (Fig. 4.1(a)), it is assumed that there have to be two contours with reasonable proportions, an outer contour of the hand and an inner contour for the hole. The inner hole has to lie approximately in the middle of the outer contour. To detect the fingertip of the closed hand, the leftmost point of the inner hole contour is taken. The intersection with a horizontal line going through this point and the outer contour is then the guessed position of the finger tip. To decrease the probability of unwanted hand detection, the user has to hold the hand in the open position in the upper left corner for a few seconds, as shown in image. Only then does the system get activated and the hand tracked. During the interaction, some filtering is applied to the results of the hand detection. If the tracking of the hand is lost, the last tracking result is used, if it is not older than a few seconds. Additionally, a closed hand can only be detected if an open hand was detected right before.

4.1.2 Evaluation

To assess the quality of the proposed AR UIs, a user study has been conducted. With respect to new types of image guided surgery systems Peters mentions that the "IGS system must not add time" or complication to the surgical procedure [168]. For this reason, time has been a major criteria for the comparison of the AR UIs. Preliminary tests by the developer team assume that HAND will consume more time than VOICE and PEDAL in terms of learning how to use the AR UI and performance in general. We, however, do not expect significant differences regarding the accuracy of performance.

In contrast to users without medical background, we expect physicians and medical students to pay more attention to criteria such as high robustness, sterility, experience with intraoperative UI solutions (e.g. pedals for x-ray devices) when judging the AR UIs. Physicians may be better aware of problems that appear when new equipment has to be integrated within their scarce and sterile workspace at the operating table. Pedals are already highly used for other devices such as mobile C-arms or bedding control. Voice control may be negatively affected by noise and communication unrelated to the system control. For this reason we expect HAND to be the surgeons' first choice. We expect user's without medical background to favor PEDAL and VOICE since PEDAL is a standard user interface e.g. for driving a car and VOICE is easy to learn. Such user's may, however, not be aware of todays infrastructure in the OR. We also expect an effect of previous experience with HMDs on the performance. Novel users may get distracted by the HMD device and the new experience of AR vision, which may affect their performance.

According to these assumptions, we state the following hypotheses:

1. Hypothesis 1: A new user will quicker understand and feel prepared to use the UIs PEDAL and VOICE than HAND.

2. Hypothesis 2: Users perform faster with PEDAL and VOICE than HAND.

3. Hypothesis 3: Users perform uniformly well with VOICE, PEDAL and HAND.

4. Hypothesis 4: Users with medical background favor HAND over VOICE and PEDAL for intraoperative usage.

5. Hypothesis 5: Previous experience with HMDs positively correlates with performance (time and accuracy).

6. Hypothesis 6: Users without medical background favor VOICE and PEDAL over HAND.

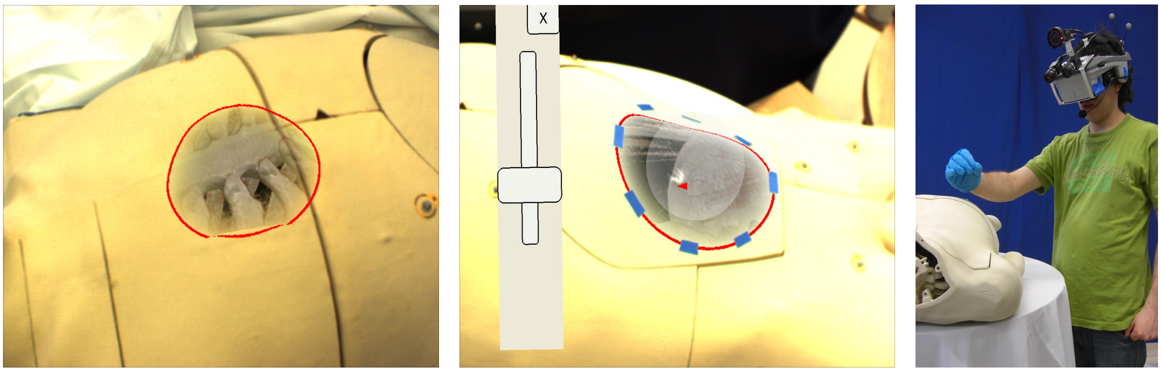
4.1.2.1 Method

In order to assess and compare the user's performance and acceptance with respect to the proposed user interfaces, we designed an experiment that requires interactive control of binary and continuous parameters. The experiment uses the AR system described in 2.1 and the Visible Korean Human Phantom (VKHP) introduced in section 2.2.2. Figure

4.2(c) shows an external view of a proband interacting with the AR scene. A pedal is positioned below the table carrying the VKHP and a microphone is attached to the HMD.

When probands have arrived, we first give information about the AR system, the VKHP, the three AR UIs and the purpose of the study. We then explain the task that has to be fulfilled using all three AR UIs. The order of AR UIs is randomized for every new proband.

The task requires the navigation (drag and drop) of a virtual vision channel as proposed in section 3.1 and [27, 128] from an initial, predefined position (see figure 4.2(a)) to a designated area on the skin surface of the VKHP (see figure 4.2(b)) and the adjustment of its diameter. The navigation of the vision channel has been chosen to allow probands with and without medical background to participate in the study. However, medical probands shall become aware of a future intraoperative application scenario of the AR system.



(a) Initial position (b) Final position (c) External View

Figure 4.2: Initial and final position of the virtual transparent window.

The AR scene probands are exposed with, consists of graphical UI elements such as a slider and a courser and contextual in-situ visualization of the VKHP's CT data.

In a first **warm-up stage**, we instructed subjects how to use the three AR UIs and give further information if requested. In this stage, we measure how long the instruction of each AR UI takes. The workflow of HAND has been explained as follows:

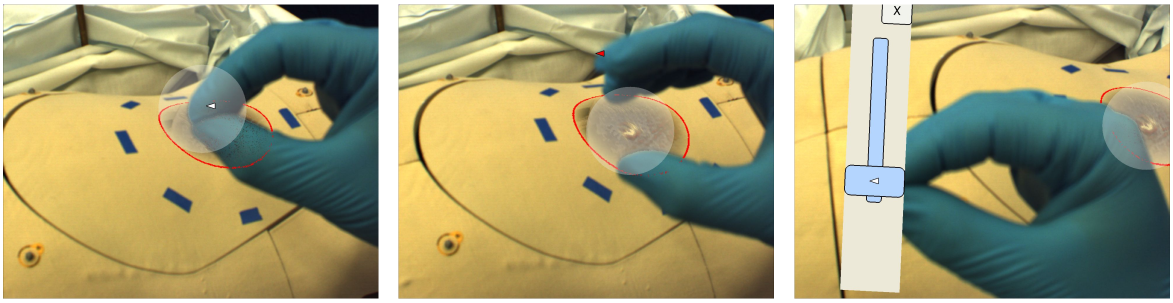
1. Activate the AR UI by positioning the hand in front of the camera as shown in figure 4.1(b). A slider and a semi-transparent 2D disk that is projected into the image plane will appear.

2. Activate the courser of the AR UI by touching the index finger with the thumb resulting in a shape as shown in figures 4.1(a) and 4.3(a).

3. Grab and drag the semi-transparent 2D disk with the courser as shown in figure 4.3(a). The center of the disk projected onto the skin surface of the VKHP corresponds to the center of the vision channel.

4. Drop the disk and correspondingly the vision channel above the target area as shown in figure 4.3(b).

5. Activate again the AR UI and grab the slider with the courser to adjust the radius of the vision channel as shown in figure 4.3(c).



(a) HAND drag (b) HAND drop (c) HAND slider

Figure 4.3: Interaction with HAND.

Probands have been instructed how to use PEDAL as follows:

1. Activate the AR UI by pressing and releasing the pedal. A virtual slider registered with the 3D AR scene, a semi-transparent 2D disk that is projected into the image plane and a courser in the center of the image plane will appear.

2. Navigate the courser by head motion to the 2D disk and press the pedal to drag the disk and correspondingly the vision channel to the target position as shown in figure 4.4(a). The center of the disk projected onto the skin surface of the VKHP corresponds to the center of the vision channel.

3. Drop the disk and correspondingly the vision channel above the target area by releasing the pedal as shown in figure 4.4(b).

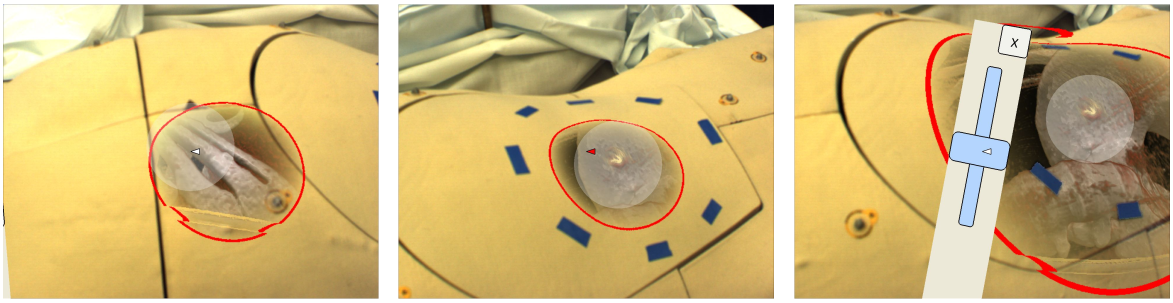
4. Activate again the AR UI by pressing and releasing the pedal and navigate the courser to the slider. Grab the slider by pressing the pedal and adjust the radius of the vision channel by head motion as shown in figure 4.4(c). Stop radius adjustment by releasing the pedal.

The workflow of VOICE has been instructed as follows:

1. Activate the AR UI by saying "NARVIS show". A virtual slider registered with the 3D AR scene, a semi-transparent 2D disk that is projected into the image plane and a courser in the center of the image plane will appear.

2. Navigate the courser by head motion to the 2D disk and say "NARVIS drag" to drag the disk and correspondingly the vision channel to the target position as shown in figure 4.4(a). The center of the disk projected onto the skin surface of the VKHP corresponds to the center of the vision channel.

3. Drop the disk and correspondingly the vision channel above the target area by saying "NARVIS drop" as shown in figure 4.4(b).



(a) PEDAL and VOICE drag (b) PEDAL and VOICE drop (c) PEDAL and VOICE slider

Figure 4.4: Interaction with PEDAL and VOICE.

4. Activate again the AR UI by saying ”NARVIS show” and navigate the courser by head motion to the slider. Grab the slider by saying ”NARVIS drag” and adjust the radius of the vision channel as shown in figure 4.4(c). Say ”NARVIS drop” to stop radius adjustment.

When feeling prepared, in a second **performance stage** probands start with the actual task.

For highlighting the destination area of the vision channel, we take the advantage of the occlusion technique described by Kutter et al. [128] that was originally used to mask the AR scene with the user's hand. For this reason the destination area is marked with blueish stripes in order to guarantee its permanent visibility (see Fig. 4.2(b)). In the performance stage probands have to perform the task with each of the interaction methods twice. In addition to the duration of the performance, we measure the deviation of the radius and pose to a predefined reference vision channel.

Right after the performance assessment, we ask probands to fill out an online survey (see appendix B.1) consisting of the USE questionnaire, demographic questions (gender, age, medical backgound) and questions addressing the experience with and usability of the present AR system. The order of items belonging to the USE questionnaire has been randomized. The questionnaire is accessible on one of our computers in the same room.

We invited 21 subjects ($N = 21$) having an average age of M = 28.1 (SD = 5.48). Seven probands have medical background and 14 probands were male. Voluntary probands were surgeons and medical students of the trauma surgery department, Klinikum Innenstadt, LMU, Munich and students of the computer science faculty. They were not compensated in any form for their participation in the study.

4.1.2.2 Results

Due to obvious erroneous measurement in the warm-up stage, we removed three performance cases for HAND, two performance cases for both PEDAL and VOICE. One performance case in the performance stage has been removed for VOICE. In the warm-up stage some of the probands did not complete the task, however, claimed to have understood how to use the AR UIs and to be prepared for the performance stage.

The Kolmogorov-Smirnov-test indicates a normal distribution for all variables except the accuracy of pose for PEDAL and HAND, both in the warm-up stage. Probands show differences in medical knowledge, experience with AR systems (yes/no: 13/8) and interactive computer games of the newer generation (yes/no: 10/11). For this reason, for normally distributed variables the one-tailed t-test for paired samples, for non-normally distributed variables the Wilcoxon-test has been applied.

Table 4.1 shows the statistics of performance time and accuracy of the size (radius) and pose of the vision channel for all three AR UIs.

Table 4.1: Results of the Experiment given as M (SD) (N=21)

	HAND M(SD)	**PEDAL** M(SD)	**VOICE** M(SD)
warm-up stage			
time (sec)	225.33 (131.01)	90.81 (84.18)	68.10 (51.96)
performance stage			
time (sec)	116.29 (56.34)	64.93 (25.34)	56.28 (24.24)
radius (mm)	5.57 (7.20)	3.59 (2.65)	5.12 (4.02)
pose (mm)	13.86 (8.61)	12.48 (8.77)	11.27 (9.88)

The tests of the defined hypotheses show the following results.

As expected, in the warm-up stage, a significant shorter preparation time for VOICE can be shown when being compared with HAND, which entails the acceptance of $H_1(t_{(.05,15)} = 3.970$, $p \leq .05)$. Also the comparison of performance time of HAND and PEDAL in the warm-up stage shows a significant shorter preparation time for PEDAL($t_{(.05,16)} = 4.398$, $p \leq .05$).

We further expected faster performance with PEDAL and VOICE than HAND in the performance stage. Comparison of HAND and VOICE shows a significant faster performance with VOICE and entails the acceptance of H_1 ($t_{(.05,19)} = 4.156$, $p \leq .05$). Also PEDAL shows significantly shorter performance time than HAND ($t_{(.05,20)} = 4.098$, $p \leq .05$).

Regarding the homogeneity of accuracy of performance (radius and pose) when comparing all three AR UIs, the F-test (MANOVA) confirms that the effect of the AR UI on the radius ($F(2, 18) = 0.177$, $p \geq .05$) and the pose ($F(2, 18) = 0.360$, $p \geq .05$) of the vision channel was at best small, however, does not lead to significant differences.

Unexpectedly, the analysis of bivariate correlation (PEARSON) shows no significant relation between medical expertise and performance ($p \geq .05$)(see table 4.2).

With respect to a favor for the intraoperative application of HAND over the other two UIs when considering only probands with medical background, comparison of HAND and VOICE even entails the acceptance of H_1 for the opposite effect ($t_{(.05,6)} = -2.449$, $p \leq .05$) telling that probands with medical background have a significant preference for PEDAL instead of HAND. No significant preference for HAND can be shown when being compared with PEDAL ($t_{(.05,6)} = -1.808$, $p \geq .05$).

Table 4.2: Bivariate correlation (PEARSON) of experience with HMD and performance in the performance stage. For all tested combinations no significant correlation can be shown ($p \geq .05$).

	HAND	**PEDAL**	**VOICE**
	r	r	r
time	-.174	-.202	0.026
radius	.025	-.013	.047
pose	-.159	-.007	.231

The items of USE categories are combined by the arithmetic mean. Figure 4.5 and table 4.3 show the results for each category and UI. We expected non medical probands to prefer VOICE and PEDAL over HAND. Comparison of HAND and VOICE shows a significant favor for VOICE with respect to the category *Ease of Use* ($t_{(.05,13)} = -2.64$, $p \leq .05$) and the category *Satisfaction* ($t_{(.05,13)} = -3.027$, $p \leq .05$) of the USE test. However, probands without medical background do not significantly favor VOICE over HAND in the categories *Usefulness* and *Ease of Learning*. In addition, H_0 is accepted for comparison of HAND and PEDAL in all USE categories ($p \geq .05$).

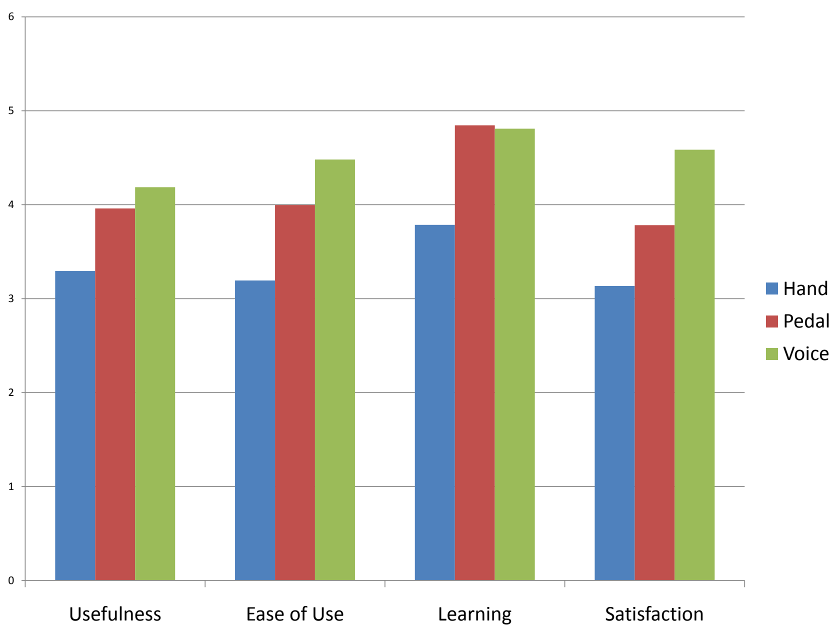


Figure 4.5: Results of the USE questionnaire

Complete answers of open questions are provided in the appendix B.1. The feedback is summarized and discussed in the following section.

Table 4.3: Results of the USE questionnaire.

	HAND M(SD)	**PEDAL** M(SD)	**VOICE** M(SD)
Usefulness	3.29 (1.62)	3.96 (0.90)	4.19 (0.89)
Ease of Use	3.19 (1.39)	3.00 (1.01)	4.48 (0.84)
Ease of Learning	3.79 (1.60)	4.85 (1.10)	4.80 (1.05)
Satisfaction	3.14 (1.75)	3.78 (1.18)	4.59 (1.37)

4.1.2.3 Discussion and Interpretation

Feedback gathered from open questions and discussions after the experiment disclose that probands with medical background favor VOICE since hands and feet are busy, hinder the field of view or require interaction outside the visible working space. However, one surgeon proposes a combination of HAND and PEDAL. Future work will include the evaluation of further combinations of the proposed AR UIs. Also a combination of HAND and VOICE has been proposed by a non medical proband.

According to the feedback of probands it is difficult to move ones view through the scene while keeping the hand within the narrow field of view of the cameras. As soon as the hand is out of sight, the HAND based UI gets deactivated, which resulted in some recognizable frustration of the probands. However, we also noticed that probands adapted to this restriction of the HAND based UI and learned how to deal with it. Adjustment of camera optics will achieve a bigger field of view and a reduction of the current optical magnification of the scene to better imitate natural optics of humans.

Although HAND performed worse in the present evaluation, some probands gave arguments for going further with this AR UI. It has been claimed that interaction with hands is more intuitive than pointing with the eyes and ”it is not easy to control the curser by moving the head”. In addition, head motion would remove the field of view from the operation site and ”gestures are easier to remember than voice commands and pedal control”. Future work will include the integration and evaluation of advanced methods for hand tracking such as a condensation algorithm [218], particle filters or machine learning based algorithms as described in the work of Kölsch [122] to improve HAND.

We observed that subjects sometimes had problems to keep theis balance when controlling the pedal while wearing an HMD and several probands claimed that PEDAL offered a poor degree of freedom of interaction and interaction has to be performed out of sight. In addition, balance problems and interferrence with other existiing pedal based devices can have negative effects. However, PEDAL can be used with hands free and robustly works even under bad conditions, such as noisy or unfavorably illuminated environments.

VOICE shows the best quantitative results and was favored by the subjects since it allows for hands free interaction and showed in most cases good robustness. Two female participants having a hoarse voice had problems with the voice recognition. Since the recognition engine was used in the untrained mode, we assume that this problem could be solved by tuning the recognition engine (by setting the age and gender) or by voice training.

We observed that most of the participants could easily remember the four commands. However, its potential sensitivity in noisy environments has been mentioned several times and has to be evaluated more extensively.

4.2 Virtual Mirror

With the combined view on the real operation site and the medical imaging data, AR visualization can provide perceptive advantages. However, one of the reasons for the absence of AR technology in todays operating rooms (OR) might be the deficient interaction of surgeons with visualized 3D data to get all desired views on the region of interest. "Magic lenses" [147] have been proposed to discover hidden structures, which are presented distinctively compared to the objects laying outside the field of the lens. Alternatively, the visualization of 3D information can be locally changed. Lenses provide additional functional views, however, they do not allow for additional geometric perspectives to entirely explore 3D data from all sides.

When observing the AR scene with head worn AR systems such as HMDs, simply changing the view position by walking around the patient becomes impractical. Space around the operating table in the OR is usually extremely scarce due to further surgical staff and all kind of equipment. Views from beneath the operating table onto the registered medical data is not possible at all and repositioning the patient is no option. Classical interaction methods for transformation parameters known from 2D interfaces can not be applied anymore. In this case, the advantage of presenting imaging data registered with the patient would lapse and mental mapping of image information with the real anatomy is again necessary. To allow intuitive interaction with and observation of virtual objects with having only one or a few points of view on the AR scene, we take advantage of an interaction paradigm, which is strongly used in our everyday life, however novel for AR applications. We introduce an interactively guided Virtual Mirror for generating additional views from any desired perspectives on the virtual part of an AR scene [154].

Overall, we determined four important benefits for Virtual Mirrors in AR scenarios.

1. **Visually accessing visually restricted areas:** Whenever a virtual object can not be viewed completely because the observer's viewpoint is visually restricted from a region of interest, the Virtual Mirror can provide the desired perspective.

2. **Supporting navigational tasks:** The Virtual Mirror is not restricted by physical barriers. For this reason, it can enter for instance the patient's body to provide helpful additional views for navigational tasks in several intra-operative procedures.

3. **Understanding complex structures:** Whenever a complex 3D structure can be observed from only a few points of stereo or monocular views for analytical or navigational tasks, a Virtual Mirror is capable of integrating additional perspectives on that structure into the original view. This can help to better understand hidden or self occluding structures such as complex blood vessel trees.

4. **Improving depth perception:** Simple superimposition of virtual anatomy on the patient or on organs inside the patient results in misleading depth perception of

involved objects. Virtual anatomy seems to be located outside the patient's body or the organ surface. A tangible Virtual Mirror with the mirror image reacting on user interaction is capable of communicating information about relative and absolute positions of affected objects.

Mirrors are well known and widely used in computer graphics to enhance the realism of virtual scenes. Within ”virtual worlds in the context of desktop workstation environments” the perceptive advantage of a ”magic mirror” is discussed in [90]. In this paragraph, we provide an overview of work employing the mirror as an interaction interface. However, none of these will present an interaction paradigm compatible to the present Virtual Mirror.

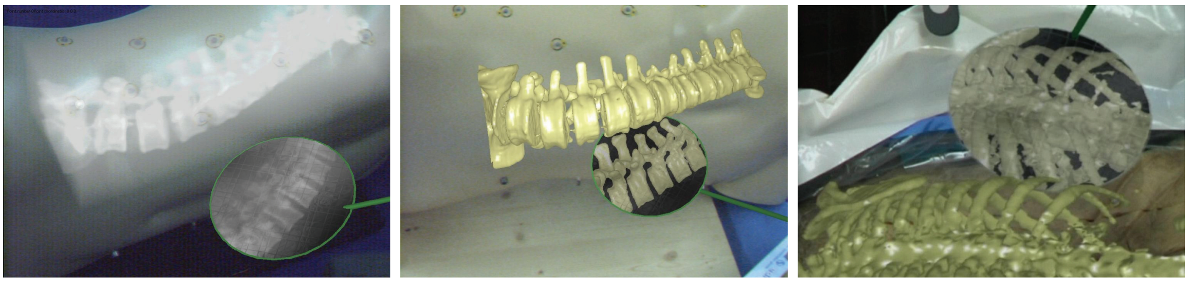
The mirror metaphor combining reality with an augmentation of virtual objects has been presented for different commercial issues. The customer is able to view oneself in a stationary mirror-like display that augments for instance clothes. Eisert *et al.* [166] and Vigano *et al.* [228] present stationary, large mirror-like displays to augment new customized sneakers on the user's feet. François *et al.* [75] introduce a ”handheld virtual mirror” consisting of a camera mounted on a display monitor and a magnetic tracking system. The recorded camera images are presented on the display and provide the mirror-like reflection, however, they do not report on registration of virtual objects to create an AR scene with the mirror. Darrell *et al.* [50] describe a system that integrates a face tracking algorithm and different image processing techniques to realize a stationary mirror display that is capable of distorting the mirror reflection.

The mirror concept was already determined to provide the driver of a car beside the information from the traditional rear view mirror with additional information in a mirror-like display that is not directly visible. Pardhy *et al.* [162] report on the idea of extending the rear view mirror in a car with additional information from a DGPS (Differential Global Positioning System), ”an onboard geo-spatial database” and ”radar or inter-vehicle communication” to provide the driver with an entire knowledge about position, distance and speed of other vehicles and objects close to his or her own car. Donath from the same group filed a patent on this topic and called it ”virtual mirror”. Kojima *et al.* [125] work on a similar concept ”to provide views of hazardous areas that drivers can not see directly”.

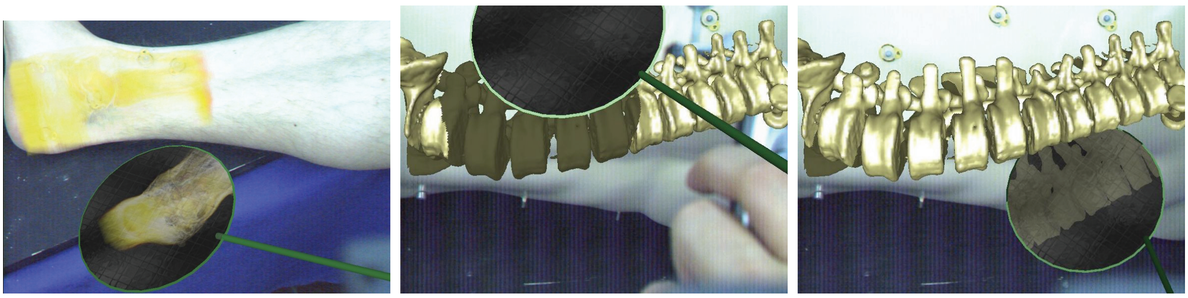
All of the mentioned mirrors referred to mixed or augmented reality applications used as a real display reflecting the world like a mirror. Some of them are capable of augmenting the reflected world with additional information registered with real objects. Our requirements on a mirror for medical AR scenarios differ from the concepts mentioned above. For our applications the mirror is one registered object among others within an AR scenario presented on a display device such as a monitor or the display of an HMD. The mirror can be guided like a dentist mirror through the AR scene having the advantage of accessing also areas inside real objects like the patient. Our mirror is completely virtual and is capable of reflecting virtual objects only. It can be attached to and guided with any tracked object, for instance a mouse pointer, a surgical drilling device, or a laparoscope. For this reason, the user can guide the mirror to a suitable position according to the particular application. Even though the mirror can only reflect virtual objects of the AR scene, virtual counterparts of objects relevant for a particular navigational tasks can be registered and visualized within the AR scene to be reflected in the mirror image like the surgical drill, the endoscope, or the endoscopic instrument.

Both, reflections on the mirror plane and shadows, are projections of the scene onto another surface, except that the mirror image contains much more information about the scene, such as color, shading, shape, and structure of complex objects, as well as information about the order of mirrored objects gained from the depth cue *occlusion*. Figures 4.6(e) and 4.6(f) show the combination of both projective light effects with one of the first versions of the Virtual Mirror.

In 2006, the concept and prototype of the virtual mirror has been presented for the first time. It can be guided through an AR scene by a hand-held, tracked, remote mouse pointer [26]. The mirror was successfully tested with a phantom (Figures 4.6(a)-(b)), a cadaver (Figure 4.6(c)), and a patient (Figure 4.6(d)). At that time, however, we have not evaluated its usefulness for any kind of application. The first generation of the virtual mirror as shown in figures 4.6 has been presented to 20 surgeons within the scope of an earlier study [200]. Physicians proposed several applications to support surgical procedures supposed that the



(a) Digital Reconstructed Radiographs (DRR) (b) Polygonal surface models (c) Cadaver study CT data



(d) Direct volume rendered MRI data set (e) Direct volume rendered MRI data set (f) Direct volume rendered MRI data set

Figure 4.6: The hand-held Virtual Mirror reflects virtual objects rendered with different visualization techniques. Both light effects shadow and reflection cause visual cues to enhance information about relative position of objects in an AR scene as shown in Fig.4.6(e), 4.6(f). Here, a spinal column is registered and superimposed on a plastic thorax phantom. The Virtual Mirror guided with a remote mouse pointer casts virtual shadow on the vertebrae.

AR system will become part of their working environment. They suggested to used the additional perspective to better understand 3D structures in diagnostics such as fracture pattern and border structures, entire vertebral segments, The mirror might also help in navigational tasks for instance the position control of implants, all kind of osteotomy close to nerves and vessels, control of the resection plane and resection border for tumor and metastases surgery, placement of osteosynthesis materials and prostheses, control of screw length, pose of drill canals and positioning screws, and get side views in general when an instrument approaches orthogradly.

In addition it could also server as an aid for the assistant, for instance when drilling has to be performed from the opposites site.

4.2.1 Real Time Mirror Techniques

In computer graphics mirror reflection is a standard effect that can be performed in real time with basically two different techniques. Both require the virtual scene to be rendered twice, once for the reflected scene and another for the original scene. For this reason the frame rate for both techniques is similar and strongly depends on the rendering of the original scene to be reflected. Both techniques have been implemented using the OpenGL and tested with our Medical Augmented Reality software framework CAMPAR [203].

The first approach referred to as render-to-texture stores a snapshot of the reflected scene to a designated texture and maps it view-aligned onto the virtual object representing the mirror. This approach allows for manipulation of the mirror image for instance distortion to magnify reflected objects. Another approach to create mirror effects employs the stencil buffer that can be used to mask the frame buffer. The stencil buffer is an additional buffer besides the color buffer and depth buffer found on modern computer graphics hardware. It can be used for instance to limit the area of rendering. Within this context the mirror is rendered first to define the mask. The mask defines the area that is allowed to be modified by the first rendering pass drawing the scene reflected on the mirror plane within the mirror mask. In the second pass the stencil function is disabled and the original scene is added to the mirror.

4.2.2 Integration of the Virtual Mirror

Different ideas how to interact with the virtual mirror have been developed strongly fitting to the needs of the particular application.

For navigated drilling in spine surgery (see section 5.2) the virtual mirror has been attached to the tracked drill. The mirror can be rotated on a circular path around the drill axes (radius=10cm) by rotating the drilling device around its axes. For ease of use the rotation angle of the drilling device is multiplied by an adjustable factor to change the position of the mirror. This enables the surgeon to move the mirror around the target while only rotating the drill by small angle. Thus only slight motion of the drilling device provides all desired side views (see figures 4.7,5.3). The mirror can be further fixed to a useful location in space.

The mirror has also been used with an augmented laparoscope to explore complex blood vessel structures for liver resection [25, 155]. Here the mirror is attached either to the laparoscope camera or to an additional tracked device (see Fig. 4.8). In both cases the

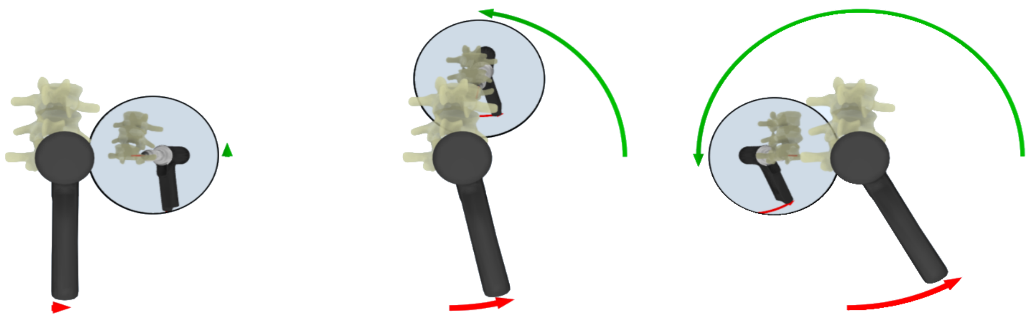


Figure 4.7: Descriptive Illustration: Angle of mirror rotation around the drill tip(green/bright arrow) is the multiplied rotation angle of the drilling device around the drill axes (red/dark arrow).

mirror can be placed again at a useful location in space.

Currently a new method is developed and evaluated to automatically position the mirror using pose information of the operation site, the tip of the surgical drill and the HMD. The new interactive positioning of the mirror is applied to a system for navigated shoulder surgery.

4.2.3 Evaluation

The stated benefits **visually accessing physically restricted areas** and **understanding complex structures** have been addressed in [25, 155]. In that case a laparoscopic camera captures the video images, which are augmented with 3D angiographic images and then presented on a monitor.

The present evaluation shall further investigate the advantage of the Virtual Mirror for **improving depth perception** and **supporting navigational tasks** in medical AR scenarios. For this reason, we designed an abstract task that is performed at a simulated minimally invasive interface to the operation site. In this case, the AR scene is presented with the HMD based system described in section 2.1. Figures 2.15 let assume that a Virtual Mirror can improve depth perception of objects within the field of view. This assumption is supposed to be evaluated with the present study. Regarding the benefit of a mirror for navigation tasks, we believe that errors can be avoided or at least they can be early detected and quickly corrected. Although the mirror reflection may be advantageous for perception and navigation issues, the study of its function and determination of its benefits for a task will consume time.

With respect to these assumptions we have the following hypotheses:

1. Hypothesis 1: A Virtual Mirror enables higher accuracy in navigational tasks.

2. Hypothesis 2: A Virtual Mirror helps the early detection and correction of erroneous navigation.

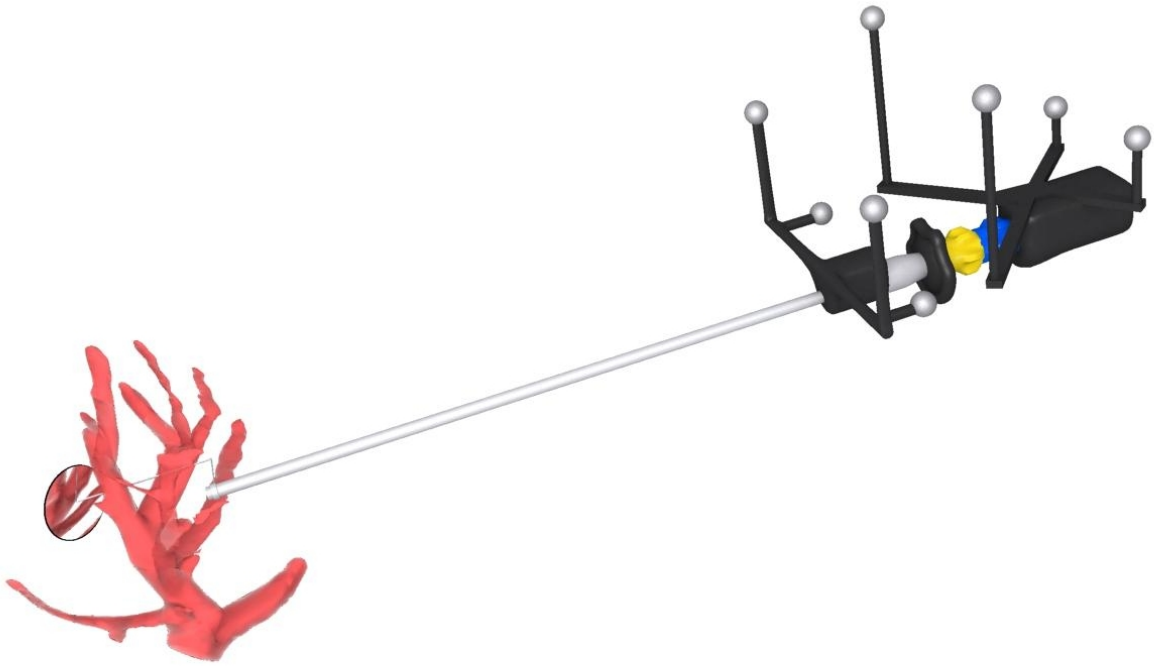


Figure 4.8: Integration of the virtual mirror for laparoscope augmentation being applied to liver resection.

3. Hypothesis 3: A Virtual Mirror can improve depth perception and enables therefore a more determined navigation.

4. Hypothesis 4: A Virtual Mirror will increase time of navigation.

4.2.3.1 Method

The experimental setup simulates an operation area, which is typical for minimally invasive spine or liver procedures (see Figure 4.9). During such interventions, surgeons usually view only a small area of the patient's skin, which is bordered by covers to secure the sterility of the operation site. Within the sterile area, different trocars are penetrated through the skin and offer ports to the interior anatomy of the patient.

The test bed consists of an aluminum frame that is covered with elastic synthetic leather to simulate the skin surface (see Figure 4.9). The leather is masked with surgical covers usually used for real operations. In the center of the masked skin area, we installed one trocar serving as the port to the backside of the skin, i.e. to the inside of the patient. The frame is positioned on a table and tracked by the outside-in tracking system, which is described in section 2.1. Probands are seated in front of the aluminum frame. This reduces their freedom of movement [163] and simulates the situation in the OR when the surgeon's space is restricted by further clinical staff and equipment. However, subjects can still move the upper part of their body, head and thorax, to get different views.

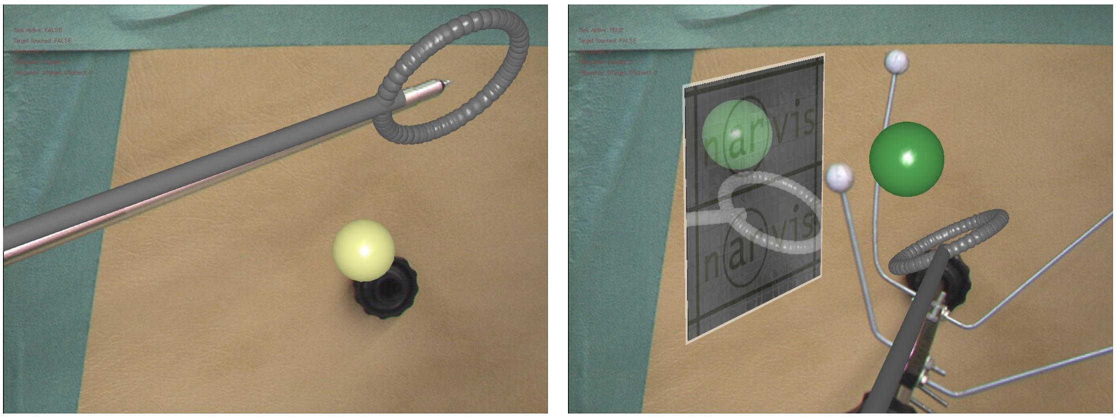
We instructed probands to penetrate a tracked endoscopic instrument through the mentioned trocar (see Figure 4.9) and showed figures 4.10 to explain the AR system and

Figure 4.9: Subjects have to insert an endoscopic instrument through the trocar within the simulated operating area.

the task. Probands received the following information:

The instrument is augmented with a virtual cylinder aligned with the instrument body and a virtual ring at its tip. In addition, the virtual part of the AR scene consists of twenty virtual green balls, only one displayed at the same time. Each ball has an individual position behind the frame and represents one of twenty tasks that have to be fulfilled. This task requires the user to move the virtual ring over the augmented green ball without touching it. The green ball changes its color to blue when it collides with the ring. To release the collision state in order to restart and finish the task, one has to move the ring away from the ball until it is green again (distance $\geq$ diameter of the ball). Once the ring has been successfully guided over the green ball without any collisions, the green ball disappears and a start button (see yellow ball in Figure 4.10(a)) is drawn at a fixed location close to the trocar. One can individually decide when they want to continue with the next ball by touching the start button with the ring. For each subject ten of the twenty balls are visualized together with a Virtual Mirror that is positioned at a suitable location that has been individually defined for each green ball. Figures 4.10(a) show the AR scene and are presented to the subjects before the experiment started. We uniformly distribute four experiment modes among the subjects that differ in the presence of the Virtual Mirror and the index order of the target balls.

Every time, when the start button is touched, we start measuring the number of collisions (NC), number of tries (NT), duration of collisions for each task (DC), duration of each task (DT), path length of the tip of the tracked instrument for each task (PL) and motion of the instrument along its axes for each task, which we call depth motion (DL). The latter two items are measured from tracking data of the instrument that is recored while the subject is performing.



(a) Start button (b) Target ball and the Virtual Mirror

Figure 4.10: In the left figure the virtual instrument and the green ball is reflected to provide additional perspectives during the performance. The right image shows the yellow start button that has to be touched to continue the task.

During first tests with the experimental setup, we noticed that moving the instrument rapidly through the scene in order to estimate the relative positions from the intersection of the virtual ring and the balls is much easier than using the mirror reflection as the primary depth cue. However, if the ball was a tumor inside an organ, picking around the target region could seriously damage the surrounding tissue. A deep seated tumor inside an organ can sometimes not be physically accessed at all with the instrument and no virtual intersection giving visual feedback can be achieved. For this reason, we instructed probands to avoid collisions: *Faultless performance is preferred over trial and error strategies.*

DL and PL are supposed to investigate, how straight forward the task can be accomplished. We assume that if subjects follow the trial and error strategy and move the ring close to the ball, they notice very late that they will fail. In this situation they will collide or almost collide with the sphere. In any case they have to bring the ring in a save position next to the ball and start a new approach, which is measured with NT.

The present test bed takes the advantage of the AR specific problem of misleading depth perception (see section 2.3). Using the naive approach of in-situ visualization, the virtual anatomy occludes the real skin. For this reason, the virtual anatomy seems to be outside the body. We try to disable as many natural depth cues as possible to negotiate the perceptive effect of a Virtual Mirror in an AR scene. The depth cue *occlusion* among real and virtual objects is ineffective, since virtual balls behind the leather cover occlude the simulated skin. Since position and size of the balls change, also the depth cue *relative* *size* does not provide perceptive hints. *Familiar size* has no effect because the spheres are not textured or do not correspond to known objects. *Motion parallax* is restricted by placing the subject on the chair. The depth cue *stereopsis*, which is ranked as one of the most powerful information sources for depth perception is available in every situation during the task due to the stereo HMD. However, the fixed orientation of the cameras attached to the HMD limits the contribution of *convergence*. The resulting AR scene has very few depth information to estimate the position and depth of the green balls, which allows us to reduce the falsification of data analysis caused by other visual cues than the Virtual Mirror. We consider DL as a strong indicator for depth perception. DL measures the ability of the proband to guide the tip of the instrument to the same depth level as the target object, a green ball. Back and forward motion of the instrument along its axes through the port would increase DL.

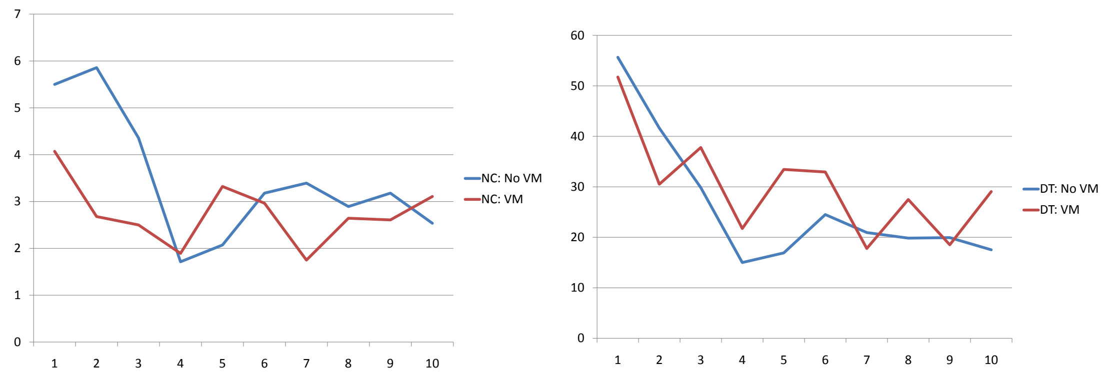
With respect to the previously stated hypotheses, we assume that the presence of the Virtual Mirror will reduce NC and NT because the additional perspective allows for side views to control the relative position of the ring and the balls and for a more determined and direct navigation of the instrument. We further assume that probands will need more time to complete one task when using the Virtual Mirror (DT). Subjects may feel motivated to use this additional information source for their task to avoid errors. However, understanding the reflection, learning to interact with the Virtual Mirror and double checking of the ring position in the reflection image will consume some time. With the additional perspective due to the Virtual Mirror DC will be reduced. Deep or unnoticed collisions costing time like complete penetration of the instrument into the ball can be rather avoided and reduce duration of collision. DC is computed for each navigation mode and each proband as follows:

$$\frac{\sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} NC_i} \tag{4.1}$$

t_i measures how long the ring has collided with the ball in one of n = 10 tasks. NC_i is the number of collisions in one of n = 10 tasks. For this reason, DC measures the average collision time of one proband.

After the practical task of the experiment, we ask the probands to fill out an online questionnaire consisting of standardized questions (scale 0-4: *I strongly agree* | *I agree* | *I am undecided* | *I disagree* | *I strongly disagree*) and open questions (see appendix B.2). The questions were posed in German, which is the native language of most of the subjects. Probands can complete the survey either on a computer in the same room as the experiment takes place or they are provided with a link and a password to complete the survey when they have time.

We invited 31 probands to participate our study. The result and feedback of the first three candidates were used to optimize the experimental setting. We then evaluated the performance of 17 female and 11 male probands (N=28). The average age is M = 29.26 years (SD = 4.53). Although medical background or anatomic knowledge is not required to participate in the study, 10 subjects (5 female, 5 male) are surgeons or medical students at the trauma surgery department, Klinikum Innenstadt, LMU München, Germany. Voluntary probands were not compensated in any form for their participation



(a) NC (b) DT

Figure 4.11: Line diagrams show mean performance with increasing experience when fulfilling the tasks with present Virtual Mirror (VM) and absent Virtual Mirror (No VM).

in the study.

4.2.3.2 Results

On the significance level of $\alpha = 0.05$ the Kolmogorov-Smirnov-test indicates a normal distribution of all samples except PL with present Virtual Mirror. Due to differing experience with the instrumentation and vision abilities of probands, the one-tailed t-test for paired samples has been chosen to test the hypotheses stated above. When analyzing the correlation of at least ordinally scaled data the one-tailed, bivariate correlation analysis (SPEARMAN) has been used. For cardinally scaled data, the one-tailed, bivariate correlation analysis (PEARSON) has been chosen.

Table 4.4 shows the descriptive statistics of samples.

The measuring data shows that the subjects need some time to get used to the system and the task. The learning curves (see figures 4.11) when analyzing DT and NC show that subjects quickly attune to the tasks.

Table 4.4: Statistics of all measured data (N=28): Number of Collisions (NC), Number of Trys (NT), Duration of Task (DT), Duration of Collision (DC), Path length (PL), Path length in depth direction (DL).

	With Virtual Mirror M(SD)	**Without Virtual Mirror** M(SD)
NC (count)	2.75(2.18)	3.47(3.11)
NT (count)	3.41(1.59)	3.8(2.27)
DT (sec)	30.101(21.47)	26.178(18.43)
DC (sec)	0.45(0.11)	0.47(0.31)
PL (mm)	1583.28(1460.82)	1432.34(924.28)
DL (mm)	737.81 (647.65)	692.61(446.61)

The test indicates significant less collisions with the balls (NC) when performing with a present Virtual Mirror ($t_{(.05,27)} = 2.062$, $p \leq .05$). This confirms our first hypothesis that the additional perspective due to the mirror image allows for more accurate navigation in AR environments.

Unexpectedly, data does not shown that a present mirror helps to early detect and correct of erroneous navigation ($t_{(.05,26)} = 0.254$, $p \geq .05$). For this reason, the second hypothesis can not be confirmed.

The analysis of DL, which is our main criteria to investigate the potentials of Virtual Mirrors for improving depth perception, cannot confirm our third hypothesis and leads to the acceptance of H_0 ($t_{(.05,27)} = -0.401$, $p \geq .05$). Also PL is not influenced by a present Virtual Mirror ($U = -0.660$, $p \geq .05$) and leads to the conclusion that the mirror does not support the navigator to work more straight forward.

Unexpectedly, no effect due the Virtual Mirror can be shown when analyzing the time of performance ($t_{(.05,27)} = -1.156$, $p \geq .05$). This let us conclude, that the understanding and analysis of the mirror reflection to be used for the navigational tasks does not negatively affect the time of performance.

While one proband is performing, investigators can observe the AR views of subjects on monitors. We noticed a difference in the performance comparing subjects with and without medical background regarding the smoothness and cautiousness of motion (see also Figure 4.12).

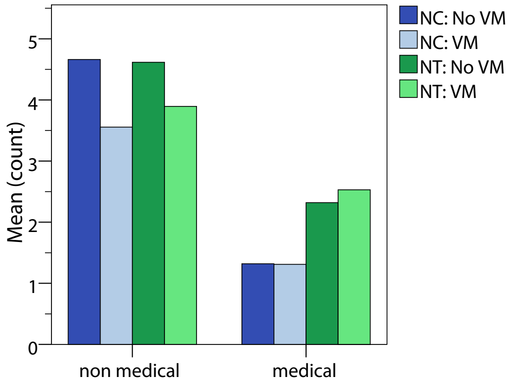
Further exploratory data analysis shows that medical expertise significantly correlates with NC for both navigation modes and with DT without mirror. The correlation of all measured samples can be reviewed in table 4.5.

Table 4.5: Correlation of medical expertise (medical background coded as 1, no medical background coded as 0) and samples: Number of Collisions (NC), Number of Trys (NT), Duration of Task (DT), Duration of Collision (DC), Path length (PL), Path length in depth direction (DL).

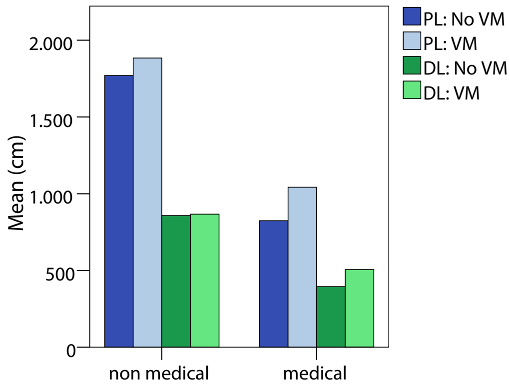
	With Virtual Mirror	**Without Virtual Mirror**
	r, p	r, p
NC	$-.503, \leq .05$	$-.525, \leq .05$
NT	$-.420, \leq .05$	$-.494, \leq .05$
DT	$-.251, \geq .05$	$-.440, \leq .05$
DC	$.317, \leq .05$	$-.078, \geq .05$
LP	$-.281, \geq .05$	$-.499, \leq .05$
DL	$-.272, \geq .05$	$-.506, \leq .05$

The results of the questionnaire are given as means and standard deviation. The original questions are provided in the appendix of this thesis B.2.

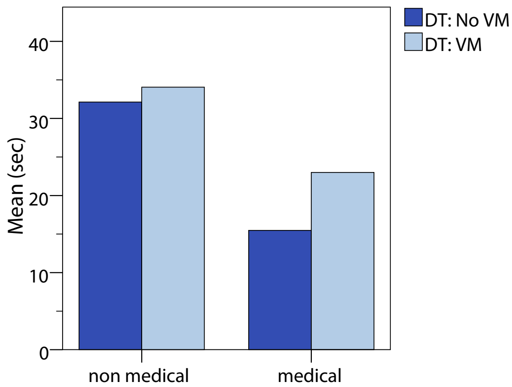
54.8% *strongly agreed* or *agreed* that they had problems to perceive the position of the green spheres. 25.8% were undecided on this question (M = 1.61, SD = 0.84). 63.3% *stronly disagreed* or *disagreed* that the Virtual Mirror is not beneficial for the performance of the task (M = 2.7, SD = 1.05). 80.6% *stronly disagreed* or *disagreed* that the mirror was not hindering the performance of the task (M = 3.26, SD = 0.93). 51.7% *strongly*



(a) PL and DL



(b) NC and NT



(c) DT

Figure 4.12: Diagrams show mean performance grouped in present Virtual Mirror (VM) and absent Virtual Mirror (No VM) as well as probands having medical background or not.

agreed or *agreed* that they used the Virtual Mirror if avaible for the task. 19.4 % were *undecided* if they used the Mirror(M = 1.68, SD 1.28). 64.4% *strongly agreed* or *agreed* that the Virtual Mirror would be more beneficial, if they can position it manually. 25.8% were *undecided* (M= 2.32, SD = 0.832).

Regarding different subtasks of one task 54.8% *strongly agreed* or *agreed* that the Virtual Mirror was helpful for the navigation of the instrument tip **to** the sphere (M = 1.58, SD = 1.12), 63.5% *strongly agreed* or *agreed* that the Virtual Mirror was helpful for the navigation of the ring **over** the sphere (M = 1.32, SD = 1.05), 51.6% *strongly agreed* or *agreed* that the Virtual Mirror was helpful to avoid collisions (M = 1.52 , SD = 1.12) and 77.4% *strongly agreed* or *agreed* that the Virtual Mirror was helpful to control the position of objects due to the additional perspective (M = 1.23, SD = 0.85). Finally 61.3% *strongly agreed* or *agreed* that they felt safer when navigating with the Virtual Mirror. 22.6 % were *undecided* (M = 1.25, SD = 0.99)

We asked the subjects to propose ideas that can improve the navigational task. They suggested to add shadow effects, digits telling the distance to the target object, acoustic feedback and the use of several mirrors.

4.2.3.3 Discussion and Interpretation

NC is considered as the most important variable that indicates high accuracy of performance. Statistical analysis has shown significant less collisions of the ring with the balls when having the Virtual Mirror as a navigational aid at hand.

Figures 2.15 assume a positive effect of mirror reflection on the correct estimation of depth order and even relative distances of objects. Unexpectedly, DL being considered as the most important indication for depth perception has not shown an improvement when adding the Virtual Mirror to the scene. Longer average pathways of the tip in three dimensions and along the instrument axes when the Virtual Mirror is present (not significant) even let us assume that its presence rather negatively affects depth perception. During the experiments, we observed that some of the probands experimented with the mirror reflection in order to negotiate the most promising position and trajectory of the ring before they started with the actual tasks. In fact, after those probands had finished all tasks, some of them claimed that they needed some time to get used to the Virtual Mirror and explore its advantage. Once they have learned how to use the mirror reflection, they found it beneficial. Interestingly, the Virtual Mirror had no effect on the time of performance (DT). Others mentioned that they had difficulties at the beginning to interact with the Virtual Mirror, which decreased their motivation for further exploration of its potential navigational advantage. In those cases, subjects reported that they decided not to use the mirror for accomplishing the remaining tasks.

Inspired by the present AR scene, one subject proposed an interesting idea on how to design a similar system for tumor resection. He suggested to attach a virtual semi-transparent sphere to the tip of the instrument. The color and transparency of that sphere can be coded with the distance and relative orientation to the target region. A virtual sphere won't injure real tissue surrounding the tumor, however, it can be used as a visual park distance control system to safely guide the instrument to the pathological tissue.

The experiment modes are not uniformly distributed within the two groups having different medical expertise. For this reason, we have to postpone the evaluation of performance regarding the expertise to future studies.

Some of the subjects and in particular the ones with medical background guided the instrument slowly and carefully through the scene with a strong focus on accurate work, i.e. avoiding collisions and finishing with the first try. For this reason, in a follow up study the questionnaire has to be extended to determine the subject's preference of working either fast or accurately. The motivation to work accurately can also increase by reactive feedback of the program due to the performance of the subject. For instance, acoustic feedback like beep sounds with increasing frequency similar to those used for announcing vital signs in the OR can produce stress when the subject fails. In addition, we will test two groups separately according to their medical expertise and compare their data. In future experiments, we plan to qualify collisions. It is interesting whether the ball was only touched slightly or the ring completely penetrated the ball. In order to evaluate the frequency of using the Virtual Mirror if present, we plan to measure the view direction of subjects. This tells us where the subjects looked at and how often the mirror reflection is used.

The reason for the predefined position of the mirror is that we want to provide similar test conditions for every new subject without evaluating their capabilities of handling two instruments at the same time. However, some of the subjects mentioned that it would be preferable to reposition or reorient the mirror to a more helpful location. We believe that in a real scenario the mirror can be a virtual add-on feature for an instrument as shown with figures 4.7 and 4.8.

CHAPTER

FIVE

MEDICAL APPLICATIONS

The previous chapters have technically described methods how to improve the composition of real and virtual entities and the interactive component of a medical AR system. The following chapter addresses potential medical applications of the AR system. Navigated spine surgery has been addressed in section 5.1.2 and training of optimally placing ports for minimally invasive surgery is described in section 5.3. This chapter begins with an overview of medical applications of head worn AR systems that have been proposed by our clinical partners and the literature.

5.1 Proposed Medical Applications: A review

The following section summarizes proposed medical applications for head worn AR systems. The summary does not include AR windows or endoscope augmentation. The applications proposed by our clinical partners strongly reflect their daily tasks in trauma surgery.

5.1.1 Applications Proposed by Our Clinical Partners

In 2006 Sielhorst et al. [200] presented a user study that has been arranged to evaluate the perceptive benefit of various in-situ visualization methods using the previously described stereo video see-through HMD. The author of this thesis is one of the coauthors of [200]. The practical part of the experiment consisted of a guidance task. Twenty trauma surgeons moved an endoscopic instrument to a virtual spinal column that has been installed inside a thorax phantom. After the experiment, subjects were asked to fill out a questionnaire. Regarding the targeted clinical application of the experiment, which is minimally invasive spine surgery, the subjects responded as shown in table 5.1 Subjects are keen to use the system for preoperative data analysis and preparation of the operation site as well as for intraoperative navigation tasks. In particular, they want to use the system for instrument positioning in open surgery and minimally invasive procedures. Consequentially, the AR system presented in section 2.1 has been evaluated with respect to *navigated drilling* (see section 5.2) and *port placement* (section 5.3).

Table 5.1: Different clinical applications of a video-see through HMD based AR system, N = 20 (scale: 0 = I completely agree to 4 = I do not agree at all).

I can imagine that such a system would support my work ...	M
... in general.	0.8095
... for preoperative analysis outside the OR.	0.8095
... for preoperative analysis inside the OR.	0.6190
... at certain moments during the intervention, e.g. to control the position of instruments respective the operation site.	0.2381
... for navigation in drill applications, e.g. implantation of pedicle screws.	0.2381
... for port placement at the beginning of an intervention.	0.6190
... for keyhole surgery to visualize instruments inside the body.	0.3810

We further asked the subjects to propose medical procedures that might benefit from the AR system. Some of the ideas have already been proposed by the literature, which are presented in section 5.1.2. They proposed to use the system in general for the visualization of pathways to the operation site. The resection of tumor affected soft tissue has been suggested for parenchymatous tissue, liver, lung and adrenal gland. Surgeons add nephrectomy, interventions concerning retroperitoneal tissue (kidney, nerves, lymph nodes) and all laparoscopic and arthroscopic operations when relevant tissue could be injured. In particular, operations may benefit that require complex navigation within and mental understanding of the interior 3D space, for instance arthrodesis and operations close to critical anatomy (aorta, spinal cord). Further applications that aim at instrument positioning are the implantation of portacaths in vascular surgery, the placement of drill canals, for example for IM nail locking, needle puncture & biopsy and drilling tasks for interventions concerning osteochondritis dissecans. They further mention the AR supported control, reposition and axis correction of multiple fractured bones with osteosynthesis, endoprosthesis in general, and treatment of fractures of the spinal column in particular for dens and pedicles. In combination with arthroscopy, surgeons also suggest knee and ligament surgery, shoulder surgery (glenoid), foot surgery (ankle joints, talus, calcaneus, cartilage therapy) and hand surgery (scaphoid and carpus fractures). They also suggest interventional procedures in neurosurgery and eye surgery.

Physicians distinguish the instructional impact of the system when the workflow of specific and unusual interventions has to be communicated. They can imagine to arrange workshops and courses employing the AR system to instruct young but also professional medicals in such interventions.

5.1.2 Applications Proposed by the Literature

This section summarizes reported clinical studies using head worn AR systems that have been reviewed within the scope of this thesis. The review might not be complete, however, is capable of reflecting the major application trends. The given ID can be used to refer the proposed target applications to the body shown in Fig. 5.1.

Authors have proposed to apply head worn AR systems intra operatively for US guided needle biopsy [229] in adrenals surgery(ID: 2) [79], breast surgery(ID: 4) [15, 79, 190], kidney surgery(ID: 11) [79], liver surgery(ID: 12) [79] and pancreas surgery(ID: 15) [79]. In addition, guided needle biopsy has been proposed in conjunction with CT data for abdominal surgery(ID: 1) [193] and with MRI data in general [112, 246]. AR guided biopsy might be beneficial in particular when dealing with "complex anatomy, where vital structures like nerves or blood vessels have to be avoided while the needle is advanced towards a target like a tumor" [192]. Neurosurgery (ID: 13) has been addressed by several authors to be a promising application domain for AR. They propose to use AR view for US based impending surgical incision [15], for intracranial procedures for tumor, vascular disease or functional disorder [95], for visualizing hidden critical structures [119], for preoperative planning regarding the limits of tumor resection, avoidance of more important structures, accurate execution of the procedure planned and guidance to small lesions such as solitary metastases for excisional biopsy [180], for craniotomy, tumor resection and needle placement and minimally invasive procedures in general [191] and image-guided stereotactic navigation in tumor surgery [234]. For oral and maxillofacial surgery AR visualization has been suggested for the insertion of endsteal implants [31, 239], for superimposing osteotomy lines to guide the saw to cut bones [235] and for osteotomy of the facial skeleton and correction of posttraumatic enophthalmos [236]. For foot and leg surgery Heining et al. [97] proposed AR views for interlocking of intramedullary nails (im-nail locking) to fixate bone fractures. Regarding spine surgery(ID: 20), authors have determined AR to be beneficial for minimally invasive treatment of herniated vertebral disc, pedicle screw implantation [219, 199] and to assist surgery to correct scoliosis [169] and [169]. Further proposed intraoperative applications are cyst aspiration and needle placement in liver surgery(ID: 12) [184], catheterization(ID: 6) [182] and foreign body removal(ID: 10) [79].

With respect to non-invasive clinical tasks, researchers have proposed cardiac diagnoses(ID: 5) [15], the examination of pregnant women (ID: 17) [15, 207] and breast diagnoses (ID: 4) [190].

Using AR as an instructional media, authors propose teaching anatomy, physiology and pathology (ID: 3) [30, 181] and radiographic positioning (ID: 18) in particular knee joints [30, 181, 9, 182], educating patients before an intervention (ID: 16), explaining medical diagnoses, procedures, treatments, prognoses, and consequences of lifestyle choices such as obesity and high-impact exercise [181], training complicated situations in delivery(ID: 7) [201], and teaching to use and interpret US images [33]. In Fig. 5.1 green disks show proposed AR aided surgical procedures, orange disks diagnostic procedures and blue disks show suggested AR aided training and educational procedures. Disks with a red border indicate already existing in-vivo studies with patients.

Table 5.2 summarizes and categorizes the clinical studies reported by the literature. Primarily clinical studies concerning the head have been performed, which can be explained by the ability of accurate tracking due to rigid skull anatomy and a wide range of state of the art systems for navigated procedures concerning the head. The non-invasive AR supported diagnoses of pregnant woman is easily to achieve and does not affect the patient except the expenditure of time. The other mentioned applications have been proposed to

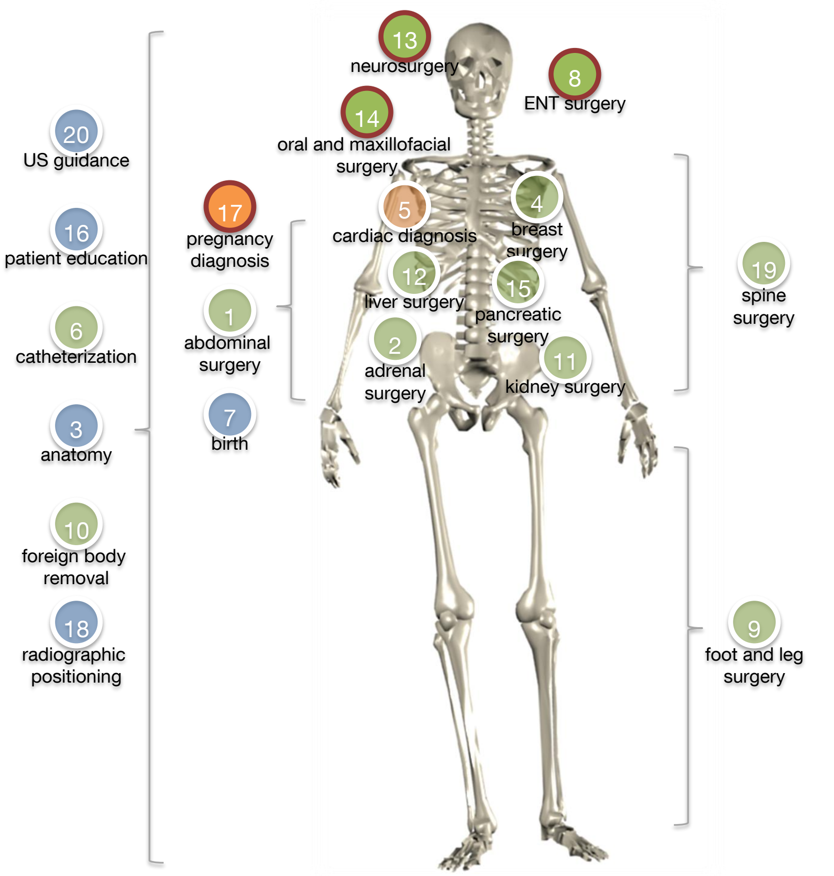


Figure 5.1: Medical applications proposed by the literature that can benefit from in-situ visualization. Green disks show proposed AR aided surgical procedures, orange disks diagnostic procedures and blue disks show suggested AR aided training and educational procedures. Disks with a red border indicate already existing in-vivo studies with patients.

theoretically benefit from AR views or have been evaluated with few animal and cadaver studies and many phantom studies.

Table 5.3 summarizes head worn AR systems being proposed by the literature for instructional applications. Together with surgeons of the Trauma Surgery Department, Klinikum Innenstadt, LMU, Munich, Germany an experiment has been arranger to evaluate the benefit of the virtual mirror described in section 4.2 when being used as a navigational aid.

5.2 Pedicle Screw Implantation

Implantation of pedicle screws is a frequently performed procedure in spine surgery. The pedicle approach is not only used to perform minimally invasive spinal interventions like vertebroplasty and kyphoplasty in osteoporotic fracture conditions but also in most dorsal stabilization procedures. The region around an injured vertebra is stabilized with an

Table 5.2: AR aided training and educational procedures proposed by the literature and the medical collaborators of the author (see also Fig. 5.1).

(ID) Domain	**Procedures**
(13) neurosurgery	3 cases reported in [119]: removal of one of two bilateral petrous apex cysts, parietal cerebral arteriovenous malformation (AVM), undiagnosed lesion near the teniclualte gnagnlion causing facial palsy 6 cases reported in [165]: brain cavernomas and low-grade brain tumors 1 case reported in [180]: left frontoparietal glioblastoma multiforme
(8) ENT surgery, (13) neuosurgery	13 cases reported in [56]: Carotid lesion, Petrous apex cyst, Superficial AVM, Facial nerve angioma, Clivus meningioma, Vestibular Schwannoma, Parietal meningioma, Vestibular Schwannoma (NF2), Subfrontal meningioma, Olfactory neuroblastoma, Subfrontal meningioma, Recurrent ethmoid meningioma, Adenocarcinoma
(8) ENT surgery	4 cases reported in [56]: Epidermoid cyst, Glomus jugulare, Ethmoid carcinoma (2x)
(14) Oral and maxillofacial surgery	5 cases reported in [235]: Removal of a foreign body in the left zygoma originating from a gunshot, suspect recurrence of osteosarcoma in the deeper infraorbital region, intraoperative virtual visualization of osteotomy lines, virtual visualization of temporomandibular joint (TMJ) movements, CT-guided placement of dental implants (Simultaneous visualization of the inferior alveolar nerve results in avoidance of atrogenic nerve trauma)
(14) Oral and maxillofacial surgery	1 case reported in [236]: Le Fort I osteotomy
(14) Oral and maxillofacial surgery	1 case reported in [234]: bioptic excision of a tumor located in the upper jaw
(17) Pregnancy diagnosis	3 cases reported in [15, 207]: US based examination of pregnant volunteer

internal fixateur attached to intact vertebrae with screws drilled into their pedicles. Drilling canals into intact pedicles around injured vertebrae is an essential, preparative procedure for implantation of pedicle screws.

The surgical task of pedicle screw placement in the lumbar and thoracic spine remains interesting even after a decade of image-guided surgery in the spine, which has led to a variety of computer-aided techniques using different imaging modalities [93]. Basic techniques use anatomic descriptions of the entry points in different spinal levels and the typical directions for pedicle screws (”droit devant” [185]) or/and static x-ray control after instrumentation as well as intra-operative 2D-flouroscopic control. Advanced techniques employ CT-based surgical navigation or/and intra-operative Iso-C-3D-control. Regarding this procedure, state-of-the-art navigation systems consist of an optical tracking system that locates surgical instruments and the patient at the operation site. Imaging data is presented as three orthogonal slice views of the operation site on an external monitor.

Table 5.3: AR aided training and educational procedures proposed by the literature and the medical collaborators of the author (see also Fig. 5.1).

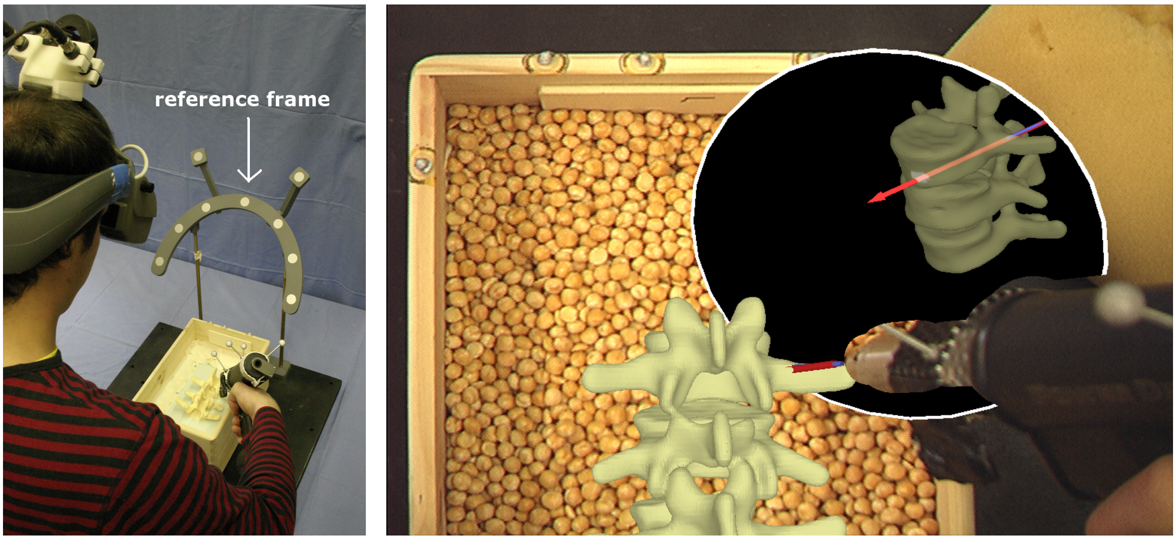
(ID) Domain	**Visualization & Learning Unit**
(18) radiographic positioning	superimposed bone structure and realistic simulation of joint motion to learn how to position joints for obtaining x-ray data beneficial for diagnostic tasks [30, 181, 9, 182]
(7) delivery simulation	superimposed virtual baby's & mother's anatomy, expert/trainees forceps paths (synchronized) and the feature After Action Review (AAR) to learn how to use forceps for delivery in critical cases of a birth stop [201]
(20) US diagnostic	US simulation from CT in combination with *contextual in-situ visualization* of CT, expert/trainees US probe paths (synchronized) and After Action Review (AAR) to learn the protocol Focused Abdominal Sonography for Trauma (FAST) [33]
patient education (16)	explain patients medical diagnoses, procedures, treatments, prognoses, and consequences of lifestyle choices such as obesity and high-impact exercise
(3) teaching anatomy	instruct the anatomy and physiology of dynamic bone structures in general, study joint changes related to lifestyle choices and pathologies [181]

Position of slices follows the drill, this means the intersection line is aligned with the drill axes. In addition, some systems provide a 3D visualization of the region around the operation site using polygonal surfaces or direct volume rendering.

5.2.1 AR-Driven Navigation

The preparation of drill canals for pedicle screw implantation has been selected as one promising application in trauma surgery that can benefit from in-situ visualization of available imaging data. In addition, the virtual mirror described in section 4.2 has been determined as a useful device to control drill depth and orientation. For this reason, we developed an experimental setup described in section 5.2.2.1 to perform drilling of synthetic vertebrae first with a classical monitor based navigation system and second with a new AR driven navigation system.

In contrast to the monitor based procedure the AR-driven mode incorporates an additional planning step. Virtual components of the AR scene include a polygonal surface model of the vertebrae, a red arrow supporting the interactive planning of the drill canal, the virtual mirror and a blue cylinder representing the tracked drill. Depending on the position of objects, a virtual model of the drilling device is visible in the mirror image. A simple method to handle occlusions among real and virtual objects in AR scenes can be achieved when there is a virtual counterpart of the real object. For example Fischer et al. [74] extract the virtual skull of a CT data set. The skull model stays invisible, however, it is employed to discard fragments of a virtual instrument, e.g. a needle for



(a) View from outside. (b) Surgeon's view through HMD.

Figure 5.2: Virtual and real objects of the AR scene. Tracked targets are surgical drilling device, phantom and the reference frame with 9 fiducial markers.

biopsy, approaching and entering the skull structure. In a similar way, we use a virtual model of the drilling device to create a depth mask with the stencil buffer[1]. This enables the direct view onto the drilling device at any time and allows for occlusion of virtual objects such as the segmented vertebrae or the virtual mirror.

5.2.1.1 Planning

The red arrow is initially orientated to the drill direction and positioned at its tip. The surgeon moves the drill to the visible entry point on the vertebrae and orientates the red arrow to the optimal drill canal. To ensure the correct position of the drill canal, the mirror is used to provide side views of the semi-transparent vertebrae. It can be rotated around the vertebrae using the tracked drill as described in section 4.2.2. When the canal is positioned correctly, it can be locked. It will then remain at a fixed position inside the vertebrae during the following steps of the procedure.

5.2.1.2 Drilling

Once the canal is defined and locked, the mirror automatically moves to a position in front of the drill, orthogonally to the drill direction. Now the surgeon moves the mirror to a desirable position that enables a view on the exit point on the bottom of the vertebrae but also allows for control of depth insertion. The mirror can be positioned automatically. However, we suggest to let the surgeon define its position according to the particular scenario, e.g. pose of the patient, position of equipment, surgeon and surgical staff during

[1]The stencil buffer is an additional buffer besides the color buffer and depth buffer found on modern computer graphics hardware. It can be used to limit the area of rendering.

the operation and position of the operation site. When the ideal position for the mirror is found, it can be locked and remains respective the vertebrae. A virtual spot light is attached to the drill tip and orientated to the drill direction. The non-realistic behavior of spot lights in OpenGL [2] here turns into an advantage. Spot lights are not blocked by surfaces. Even surfaces not visible for the light source are illuminated, if they are located inside the cone of the spot light. Therefore the spot light illuminates the entry point at the pedicle as well as the exit point on the opposite side of the vertebrae, which is visible only through the mirror (see figure 5.2(b)). Regarding the surgical task, the drill has to be aligned with the defined drill canal at the entry point using visual cues due to the spot light and intersection of drill, vertebrae and drill canal. Thereupon the drill has to be reoriented until the visible spot light is aligned with the exit point on the back of the vertebrae seen through the mirror.

Figure 5.3: Mirror is rotated around the operation site to check the position of the drill canal. Box was filled with peas to simulate the restricted view on the operation site.

5.2.2 Evaluation

The proposed AR-driven navigation system incorporating a Virtual Mirror has been quantitatively compared with a standard monitor-based navigation system as it is used in today's ORs. We expect, that the AR-driven approach consisting of a planning stage and a performance stage will consume more time than the workflow of the monitor-based system, which comprises only a performance stage. However, we believe, that the combined in-situ presentation of navigational information and 3D visualization of imaging data will allow for more accurate performance.

Following these assumptions, the hypotheses are defined as follows:

- Hypothesis 1: The AR-driven navigation can be performed with higher accuracy.
- Hypothesis 2: The AR-driven navigation will consume more time.

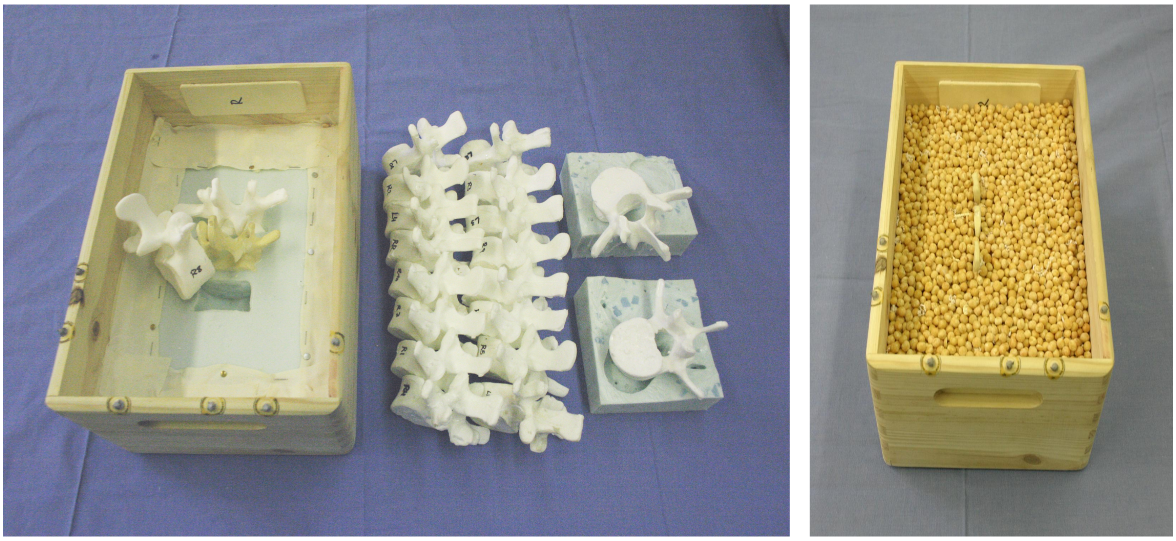
5.2.2.1 Method

For primary tests and prospective evaluation, we built a phantom consisting of three vertebrae embedded in a silicon mold. The outer two vertebrae are replaceable. The

[2]Open Graphics Library: www.opengl.org

use of the silicon mold avoids multiple scans for every new experiment. Therefore, we created further silicone molds (see figure 5.4(a)) to reproduce the outer two vertebrae using synthetic two-component resin, which has similar physical properties as real vertebrae. According to the surgeons, the material is slightly harder than real bone, which, however, does not affect the drilling procedure. The reproduced vertebrae fit precisely into the original silicon mold.

In a real scenario the surgeon only views a small area of vertebrae and therefore is able to roughly estimate the pose of the spinal column. Real pose of the spine can only be estimated. The silicon mold holding the three vertebrae is installed into a wooden box, which is filled with peas to simulate such restricted direct view in spine surgery (see Fig. 5.4(b)).



(a) Phantom, silicone mold and reproduced vertebrae. (b) Restricted view on vertebrae.

Figure 5.4: Phantom for drilling experiments.

Probands are asked to drill preparative canals ($\oslash_{drill} = 4mm$) for pedicle screw implantation into the two replaceable lumbar vertebrae. In real surgery, surgeons prepare the pedicles to avoid gliding off and injuring anatomy before they start drilling. In this experiment, however the surgeons start drilling directly into bones. Each subject has to consecutively drill four canals using each method. Both, the starting mode and the order of pedicles to be drilled, are defined randomly. The quality of the drill canals is judged (double blind) by an expert surgeon of the trauma department of Klinikum Innenstadt. For this reason, a CT scan of all drilled vertebrae is achieved. Drilling quality is scaled as *perfect* = *0*, *acceptable* = *1* or *perforation* = *2* following the method of Arand et al. [8].

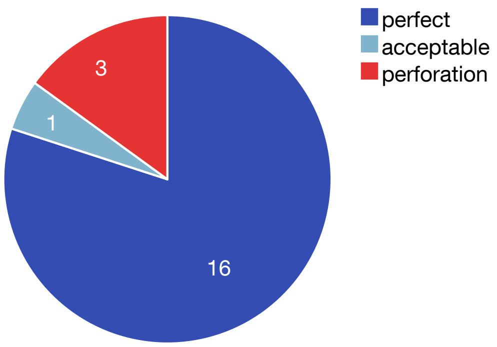
The experiment was conducted at our lab space [3]. Probands were five (N=5) trauma surgeons of the trauma surgery department at Klinikum Innenstadt. Regarding navigated drilling for pedicle screw implantation, three of the surgeons are highly experienced, one

[3]NARVIS lab: http://campar.in.tum.de/Chair/NarvisLab

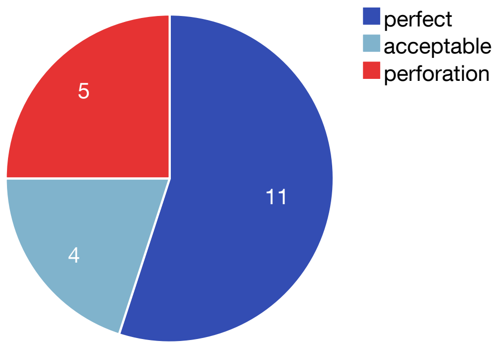
has low experience and one has no experience. All of them have been exposed to our AR system within the scope of a previous evaluation study.

5.2.2.2 Results

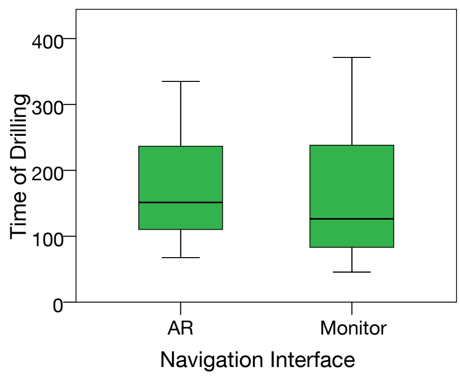
Overall, we measured the quality of 20 canals for each method and time of the procedure without missing values. Figures 5.5(a) and 5.5(b) show the frequencies of drilled canals according to their qualities.



(a) AR-based



(b) Monitor-based



(c) Boxplots

Figure 5.5: Frequencies of drilling qualities and distribution of drilling performance regarding time duration (c).

The boxplots show the overall distribution of measured time of drilling respective the two navigation interfaces (see Fig. 5.5(c)). They indicate a slightly more consistent performance in terms of duration of drilling when using the HMD as the navigation interface. With respect to the HMD-based mode, the graduated median shows that the additional planning stage costs more time, however, the span length as well as the interquartile range (IQR) of measured values shows a more consistent performance. Also an asymmetric distribution of time duration is shown by both navigation modes. Time duration increases with variance.

Statistics presented in table 5.4 let assume more time efficient performance when using the monitor based method. When comparing the quality of drill canals, the AR-driven navigation method seams to lead to more accurate results (see table 5.4).

Table 5.4: Descriptive statistics (N=5 surgeons). Drilling quality has been scaled as *perfect = 0*, *acceptable = 1* or *perforation = 2.*

	HMD-based	**Monitor-based**
	M (SD)	M (SD)
time	173.75 (84.13)	168.95 (103.59)
	Median	Median
quality	acceptable (1)	acceptable (1)

Since the five probands show different levels regarding surgery experience, application of surgical navigation systems and experience with the AR system, we compare their individual time of performance pairwise using the one-tailed t-test for paired samples. The quality of drilling is analyzed with a one-tailed Wilcoxon-test for paired, ordinally scaled data. A Kolmogorov-Smirnov-test indicates a normal distribution for quality and time data for each of the navigation user interfaces.

Regarding the number of perforations, no significant difference can be shown when comparing both navigation methods ($t_{(.05,4)} = -1.633$, $p \geq .05$). However, the number of perfect drill canals is significantly higher for the AR-driven mode ($t_{(.05,4)} = 3.162$, $p \leq .05$).

Regarding hypothesis 2, time measurement does not show a significant difference in time of performance when comparing both navigation modes ($t_{(.05,4)} = 0.136$, $p \geq .05$).

Further exploratory data analysis shows that the one-tailed, bivariate correlation analysis (SPEARMAN) of time and quality for the HMD-based ($r_s = .224$, $p \geq .05$) method as well as for the monitor-based ($r_s = .030$, $p \geq .05$) navigation system shows no significant correlation.

5.2.2.3 Interpretation and Discussion

Although the t-test has not shown significant differences, descriptive statistics show at least a tendency that the additional planning stage of the AR-driven navigation consumes more time. After the experiment, surgeons told us that they favor the additional planning step since it enables a final check of canal position and allows for collaborative decision making before the vertebra is actually drilled. In addition, they claimed that the time lag of a few seconds in favor of an intraoperative planning stage would not negatively affect the outcome of the surgery.

Probands wear the HMD in average 11 min and 34 sec to drill the four canals into the vertebrae. Unexpectedly they did not state that their performance is negatively affected by the weight of system on their head. However, they reported that looking down to the operation site causes some discomfort and shifts the displays away from the user's eyes. This happens in particular when the HMD can not be strapped correctly to the everybody's head since an appliance for universal head shapes and sizes is missing in this prototype.

The commonly tight workflow of surgeons makes it difficult to access potential end users of the proposed system. In order to draw stronger statistical conclusions, the experiment has to be extended to a higher number of probands and completed in the future.

5.3 Port Placement

Together with our clinical partners, we identified the definition of entry ports and the consequential preparation of skin incisions as one of the early crucial tasks in minimally invasive surgery that can potentially benefit from the surgeon's augmented vision. Positioning ports has to cause as few lesions as possible and provide maximal access for the instruments to reach the site of operation. In many cases, palpation of superficial bone structures provide important cues to decide on pose and shape of the incision. However, when the operation site is deep-seated and multiple tissue layers and risk structures have to be passed, information from superficial palpation is not sufficient. In these cases, the surgeon's ability of generating a mental model of the patient's individual anatomy becomes important. When available and applicable intraoperative imaging modalities such as US or fluoroscopy are used to support the creation of this model.

The literature proposes several methods using 3D visualization for the manual planning of port positions [46, 220] as well as automatic computation [2, 45, 140] of optimal access points for surgical instruments. AR supported port placement has been discussed by Feuerstein et al. [66] and Traub et al. [220].

5.3.1 Evaluation

We arranged a user study with surgeons and medical students of our partner hospital, in order to evaluate the effect of AR vision on the quality of ports to be placed. As a second objective of the study, we intend to investigate the potentials of the AR system to be used as an educational training tool. Here, we aim at assessing the ability of the combination of AR vision and palpation to create a sustained mental model of the interior anatomy to be applied to future port placement.

We selected three minimally invasive orthopedic operations for which a series of probands has to define suitable ports. In addition, we defined a set of criteria for each of the operations to judge the quality of port placement. The selected operations, usually performed under fluoroscopic control, progress in difficulty from selecting one entry point over setting a cloud of ports for the second task, to defining a skin incision.

In case of anterior pelvic ring fracture with transpubic instability, a screw osteosynthesis can be carried out percutaneously. This procedure is delicate as the available space is narrow and on both sides structures are at risk.

Thoracoscopically assisted anterior spine surgery is a minimally invasive procedure used in severe vertebral fractures in thoroco-lumbar levels. Deflating one of the patients lungs opens a workspace and three to four intracostal ports are placed directly above the injured vertebrae. The ruptured disc material has to be removed and using cortical bone graft from the iliac crest a spondylodesis is then performed. In our test scenario, we asked for a Th10 to Th11 procedure defining one port for a camera and two for instruments entering from the left hemithorax.

In the case of degenerative disc disease, lumbar disc replacement using an anteriorly placed disc prosthesis is a minimally invasive procedure requiring a small paramedian incision. According to the patient's weight and size, the incision site is chosen under fluoroscopic control and marked on the skin. We asked the participants to define left paramedian incision for a L5 to S1 intervention.

When designing a study incorporating humans as the unit of observation, one of the major premises is to ensure that probands do not feel frustrated or worse due to the study. To avoid this, probands have to be fully informed about the intents of the study and all remaining questions have to be discussed and solved before they are released from the study procedure. In this case, probands have to fulfill tasks with a prototypical AR device that still suffers from ergonomic constraints. For instance, the HMD can not be individually adjusted to all probands head shapes. Also the proband's ability of compensating the weight of the HMD to achieve comfortable or at least acceptable working conditions may be strongly affected by many unknown, individual parameters. The task probands are exposed with in our study has characteristics of an exam situation, that evaluates their professional competence. Although, this aspect is not intended to be assessed, the experimental scenario might cause additional stress to the probands. The combination of the proband's potential experience to be situated in an exam situation and the suboptimally designed HMD prototype may affect the proband's mental state. In order to assess whether probands had the same mood when they started with and were released from the experiment, we use the Befindlichkeitsskala (BFS) proposed by Abele-Brehm and Brehm [1](scale 0-4: I do not agree – I agree). The BFS has to be filled out twice and builds the starting and ending event of the experiment schedule.

We incorporated a second test into the workflow of the study that assesses proband's abilities of stereo vision. For this reason, we developed a set of Random-Dot-Stereograms with Photoshop [4] that are similar to the standard Lang-Test [5] used by ophthalmologists. To display the stereo images along separate optical paths for the right and left eye, we use the advantage that the stereo HMD has separated displays. The objects hidden in the stereograms only become visible when candidates use both eyes while wearing the HMD. The used stereograms are provided in the appendix B.3. Two variations of the test are used within the schedule of the experiment. Before probands are exposed with the AR scene, we want to make sure that subjects have mounted the HMD correctly to actually take the advantage of the stereo see-through system. For this reason, we asked subjects whether they can see an elephant hidden in the Random-Dot-Stereogram shown in figure 8.1 attached to the appendix B.3. In a second test later in the experiment schedule, we want to gradually evaluate their ability to see in stereo. Following the Stereo-Test Düsseldorf [57] each stereogram consists of four different shapes in this case a square, circle and two differently orientated triangles. The levels of difficulty in our test are defined by

[4]graphics editing software of Adobe Systems, USA: adobe.com

[5]Test to observe the abilities of stereo vision by Joseph Lang

the number of pixel shifts: 1x, 3x, 4x, 6x, 9x, 12x. Shifts range from $3.53\mu m$ (difficult) to $42.33\mu m$ (easy).

We hypothesize that

1. Hypothesis 1: Available AR vision will improve planning of ports.
    - Ports leading to impossible instrument trajectories through blocking and risky anatomic structures can be avoided.
    - A set of ports or an incision is better arranged with AR vision to achieve the best balance of maximum freedom of movement for the instruments and accessibility of the operation site.
2. Hypothesis 2: The experience of AR-vision showing a port set defined by an expert as well as a previously set defined by the learner will show a positive learning effect in subsequent classically performed procedures.
    - Ports leading to impossible instrument trajectories through blocking and risky anatomic structures can be avoided avoided.
    - A set of ports or an incision is better arranged with AR vision to achieve the best balance of maximum freedom of movement for the instruments and accessibility of the operation site.

5.3.2 Method

The AR system described in section 2.1 is used to project CT data into the Visible Korean Human Phantom (VKHP) (see section 2.2.2) with an advanced version [128] of the contextual in-situ visualization introduced in section 3.2).

The experiment is scheduled as follows:

1. **Questionnaire Phase 1:** Initial BFS-test.
2. **Instruction Phase 1(IP1):** Candidates are provided with information about three operations from a surgeon who assists us in the experiment. This includes a text with illustrations, a life size skeleton model and other plastic bone models relevant to the operations. The assisting surgeon points out the optimal position and arrangement of the ports and gives further explanations if requested. However, the VKHP is not used during this briefing.
3. **Port Placement Phase 1 (PPP1):** Applying the information given to them in IP1, subjects point with a tracked hand held target to a port position on the skin surface of the VKHP.
4. **Stereo Vision Test 1& HMD Setup SV1:** Ensure correct mounting of the HMD to see in stereo.
5. **Port Placement Phase 2 (PPP2):** Subjects wear the HMD to adapt themselves to the augmented view inside the patient. When the subjects feel comfortable, they start to point again with a tracked hand held target to define the same set of port positions but now with an augmented CT dataset.

6. **Instruction Phase 2 (IP2):** Subjects view their defined port positions from PPP1, their port positions from PPP2 and also the port positions from an expert surgeon, which has been previously saved with enabled AR vision.
7. **Stereo Vision Test 2 (SV2):** Gradual testing of subjects' abilities to see in stereo.
8. **Port Placement Phase 3 (PPP3):** Subjects place the ports again without the use of the HMD.
9. **Questionnaire Phase 2:** Along with the second instance of the BFS-test, we pose further questions to assess the quality of the AR system and the experiment.

SV2 is scheduled between IP2 and PPP3 in order to establish a small break and allow the candidates to focus on another task other than port placement. This may allow us to evaluate the learning effect rather than simply the abilities of a candidate's short term memory for remembering landmarks on the VKHP [24] for accomplishing PPP3.

The probands define their preferred port positions with a tracked pointer by tapping on the skin surface of the VKHP (see Fig. 5.6(a), 5.6(b)). For each participant, we measure one port for the pelvis surgery, three ports for the thoracic spine surgery and two points defining an incision for the lumbar spine surgery (see Fig. 5.7(a), 5.7(b), 5.7(c)).

(a) Defining the ports for thoracic spine surgery (b) With AR vision (c) Reviewing the ports for lumbar spine surgery

Figure 5.6: One of 20 subjects performs port placement.

The following data is measured and calculated to judge the quality of posed ports.

According to our expert surgeon, in many cases there is not only one optimal set of ports. Slight variations would still allow for a reasonable instrument setting. An expert surgeon assessed qualitatively (double blind) the participants' ports positioning (**QualPos**) and arrangement where applicable (**QualArr**) on an ordinal scale from 1-3 (0 is optimal, 1 possible, 2 is impossible).

In addition to the qualitative judgment, the similarity of the expert's ports to the subjects' ports is measured metrically for each clinical operation with regards to the number and type of employed surgical instruments as well as the accessibility of the operation site. Here, we analyze geometric functions of the procedure specific port set, i.e. a line for the incision in lumbar spine surgery and a triangle in thoracic spine surgery. We use the probands' samples and the expert's port set as the input of these functions and compute the geometric deviation of the results. This deviation is then used as a metric criterion to judge how similar the proband's performance is compared to the one of the expert.

For the pelvis surgery, the geometric distance of the expert's port to the proband's port is measured (**Centroid**) in addition to QualPos.

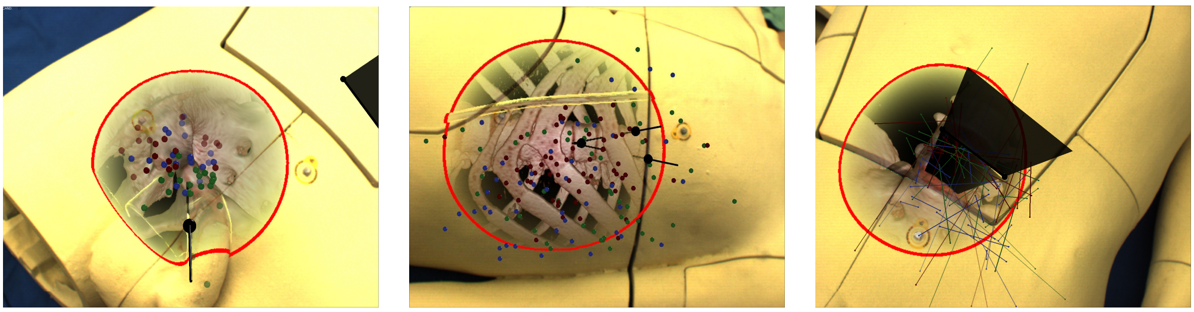
With respect to the thoracic approach, along with QualPos and QualArr, we analyze two geometric functions of the triangle that is defined by the three ports. The geometric functions are the centroid (**Centroid**) of the triangle and the distance of centroid to the operation site (**PathLen**). PathLen servers here as an indicator for the quality of the port setup so that instruments can be inserted close to orthogonally the body to create best balance of freedom of movement and accessibility of the operation site.

For the lumbar spine surgery, one geometric function analyzes the parallelism (**CutAng**) of the expert's and subjects' cuts. Further geometric functions are length of the cut (**CutLen**), the centroid (**CutCent**) and the distance of the centroids to the operation site (**PathLen**). QualPos and QualArr again play a major role when judging the port placements. Precise orientation of the skin incision of this mini-open procedure is crucial in order to fully reach the intervertbral space L5-S1 for removal of the degenerated disc and insertion of the implant. The expert's incision is shown by a colored plane between the disc and the skin incision (see Fig. 5.7(c)).

The questionnaires before and after the port placements are accessible on one of our computers in the same room as the experiment takes place. The items of the BFS questionnaire are attached to the appendix B.3.

We invited 20 surgeons and medical trainees, i.e. students in the last year of their medical studies. The average age is M = 26.01 (SD = 3.23), seven of them are female, six wore glasses, two had experience with 11 to 20 endoscopic interventions and 18 had no previous experience in the surgical operations used in the study. Subjects have the option to stop the experiment if requested, however, everyone completed all stages of the study.

5.3.3 Results



(a) Pelvis surgery (b) Thoracic spine surgery (c) Lumbar spine surgery

Figure 5.7: Subjects' ports of PPP1 in blue, PPP3 in red and PPP2 in green. Black spheres represent the expert's reference ports.

Since the twenty probands show different levels regarding surgery experience and experience with the AR system, we compare their individual performance (cardinally scaled data) pairwise using the one-tailed t-test for paired samples. The quality of port placement as well as the non normally distributed samples are analyzed with a one-tailed Wilcoxon-test for paired, ordinally scaled data. A Kolmogorov-Smirnov-test is used to control normal distribution of assessed data. When analyzing the correlation of at least ordinally scaled data the one-tailed, bivariate correlation analysis (SPEARMAN) has been used. For cardinally scaled data, the one-tailed, bivariate correlation analysis (PEARSON) has been chosen.

The results from the questionnaire (scale 1-10: I don't agree – I strongly agree) show that all participants found the AR vision into the body helpful for their task (42.86% strongly agreed)). 71.43% of proabands claimed that the experiment was mentally demanding and 23.8% said that it was physically demanding.

In SV2 test 13 candidates saw all objects in the stereograms, one person only saw objects on the left area of the screen, two saw no objects and four probands missed a minor number of objects. The results of the survey shows that 85.72% used both eyes (stereo vision) during the performance.

5.3.3.1 Pelvis Surgery

Table 5.5 shows the statistics of the measured data during PPP1, PPP2 and PPP3 for pelvis surgery.

Table 5.5: Pelvis surgery (N = 20)

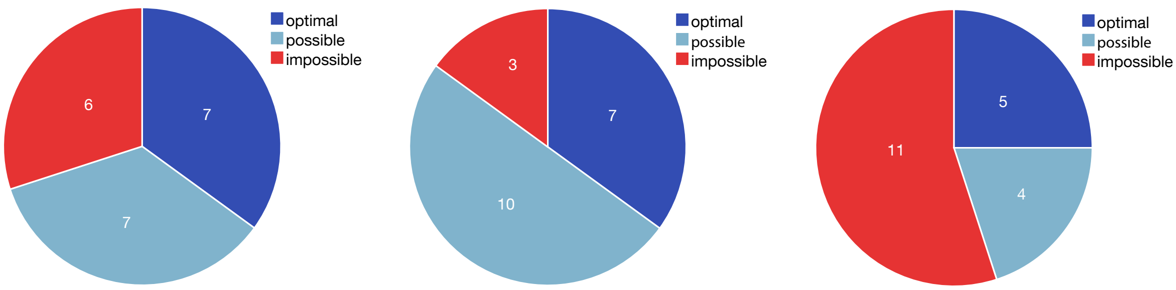
Data	**PPP1**	**PPP2**	**PPP3**
	M (SD)	M (SD)	M (SD)
Centroid [mm]	52.18 (14.56)	33.49 (7.96)	64.87 (20.04)
	Median	Median	Median
QualPos	possible (1)	possible (1)	impossible (2)

All metrically scaled variables indicate a normal distribution. Figures 5.8 shows results of the expert's judgment of the quality of port position.

With respect to the first hypothesis figure 5.8(b) indicates less impossible port placements with available AR vision. Further analysis shows that the quality of ports in PPP2 has at least the same quality than in PPP1, however, does not significantly improve. Comparison of QualPos in PPP1 and PPP2 shows no significant difference in port quality ($U = -0.660$, $p \geq .05$). However, the port is geometrically positioned (Centroid) significantly closer to the one of the expert in PPP2 ($t_{(.05,19)} = 7.478$, $p \leq .05$).

With respect to the second hypothesis, no learning effect can be shown when comparing QualPos as well as Centroid of PPP1 with PPP3 ($U = -1.425$, $p \geq .05$). In fact statistics and figure 5.8(c) let assume that probands performed even worse in PPP3 than in the earlier phases of the experiment. We want to state that also no significant opposite effect can be shown, which would tell that AR experience would result in worse port qualities of subsequent, classically performed procedures.

In addition to previous findings, exploratory data analysis shows a strong positive correlation of Centroid and QualPos in PPP3 ($r_s = .718$, $p \leq .05$). 47.62% claimed that



(a) QualPos in PPP1 (b) QualPos in PPP2 (c) QualPos in PPP3

Figure 5.8: When AR vision is available the impossible port placements in pelvis surgery can be reduced. A learning effect can not be shown.

the weight of our prototype HMD was rather cumbersome.

5.3.3.2 Thoracic Spine Surgery

Table 5.6 shows the statistics of the measured data during PPP1, PPP2 and PPP3 for thoracic spine surgery.

Table 5.6: Thoracic spine surgery (N = 20)

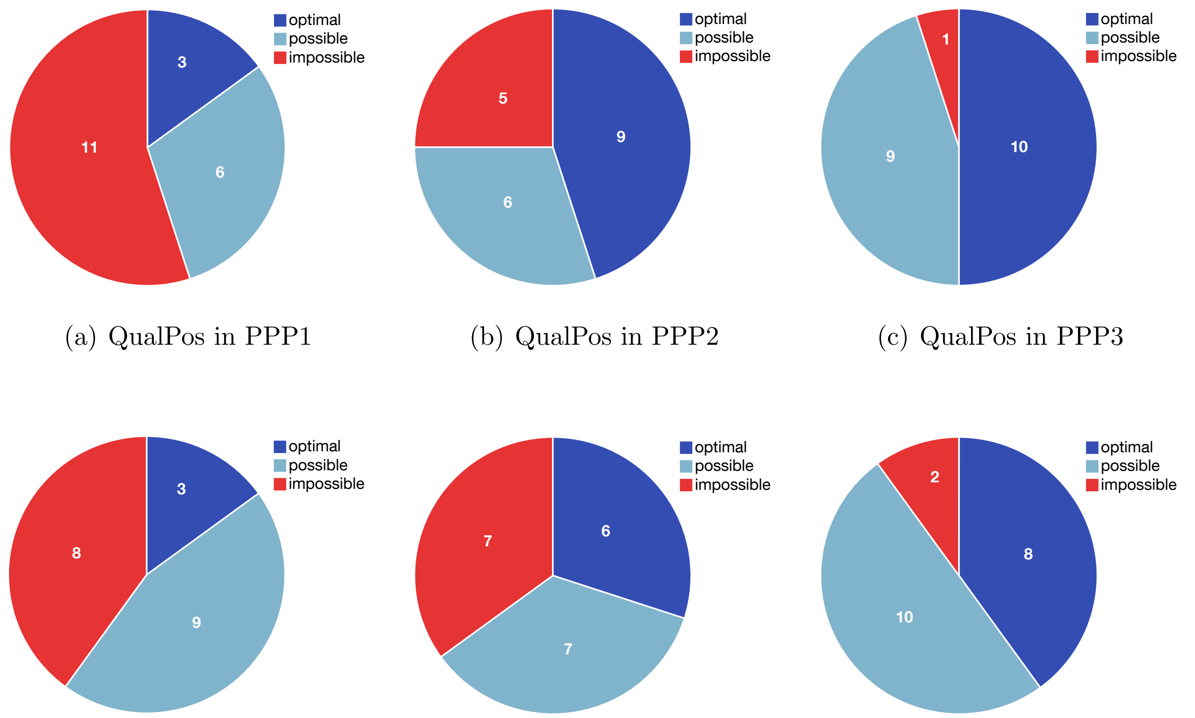
	PPP1	**PPP2**	**PPP3**
Data	M (SD)	M (SD)	M (SD)
Centroid $[mm]$	91.97 (43.41)	68.45 (34.97)	50.25 (22.60)
PathLen $[mm]$	6.09 (5.09)	7.74 (4.60)	3.49 (2.30)
	Median	Median	Median
QualPos	impossible (2)	possible (1)	"almost" optimal (0.5)
QualArr	possible (1)	possible (1)	possible (1)

All cardinally scaled samples variables indicate a normal distribution. Figures 5.9 shows results of the expert's judgment of the quality of port position and arrangement.

First, we analyze the effect of AR vision on setting suitable ports and compare performance of PPP1 and PPP2. Statistics and figures 5.9(b) and figure 5.9(e) indicate improved performance when AR vision is available. Advanced analysis discloses that QualArr is close to significant improvement of performance with AR views($U = -1.633$, $p = .051$). QualPos shows that the arrangement of ports significantly improves with augmented views of the patient ($U = -2.377$, $p \leq .05$).

Analysis of metrically scaled data shows no significant shorter trajectories with AR vision ($t_{(.05,19)} = -1.327$, $p \geq .05$). Centroid shows significantly better arrangement of ports ($t_{(.05,19)} = 2.851$, $p \leq .05$).

When investigating the data to measure a potential learning progress between PPP1 and PPP3, we identify a significant better positioning and arrangement of port sets in the



(a) QualPos in PPP1 (b) QualPos in PPP2 (c) QualPos in PPP3

(d) QualArr in PPP1 (e) QualArr PPP2 (f) QualArr PPP3

Figure 5.9: When AR vision is available the impossible port placements in thoracic spine surgery can be reduced. PPP3 shows further reduction of impossible ports.

last phase for QualPos ($U = -3.017$, $p \leq .05$) as well as QualArr ($U = -2.653$, $p \leq .05$).

Analysis of PathLen path length shows a significant shorter trajectory ($t_{(.05,19)} = 2.088$, $p \leq .05$). Also the centroid of proband's triangular port set becomes significantly more similar to the one the expert ($t_{(.05,19)} = 4.475$, $p \leq .05$).

Further exploratory data analysis shows significant positive correlation between QualPos and Centroid in PPP1 ($r_s = .517$, $p \leq .05$), PPP2 ($r_s = .517$, $p \leq .05$) and PPP1 ($r_s = .772$, $p \leq .05$).

5.3.3.3 Lumbar Spine Surgery

The statistics for lumbar spine surgery are shown in table 5.7.

Except CutAng for PPP1 and PPP2 all cardinal variables indicate a normal distribution. Figures 5.10 shows results of the expert's judgment of the quality of port position.

First, we analyze the effect of AR vision on setting suitable ports and compare performance of PPP1 and PPP2. Figures 5.10(b) and 5.10(e) indicate slightly higher quality of the incisions performed with augmented views. Further statistical analysis, however, shows that this tendency does not result in a significant difference for both QualPos ($U = -1.155$, $p \geq .05$) and QualArr ($U = -0.302$, $p \geq .05$). Also the analysis of metric criteria to judge the similarity of probands' performance to the one of the experts

Table 5.7: Lumbar spine surgery (N = 20)

Data	**PPP1** M (SD)	**PPP2** M (SD)	**PPP3** M (SD)
CutAng [°]	19.38 (17.61)	17.87 (16.57)	13.87 (30.36)
CutLen [mm]	12.49 (10.32)	14.58 (9.46)	8.93(7.96)
CutCent [mm]	40.73 (20.82)	34.48 (25.54)	28.29 (26.48)
PathLen [mm]	24.44 (10.82)	21.11 (14.91)	19.15 (13.94)
	Median	Median	Median
QualPos	possible (1)	optimal (0)	optimal (0
QualArr	possible (1)	optimal (0	optimal (0

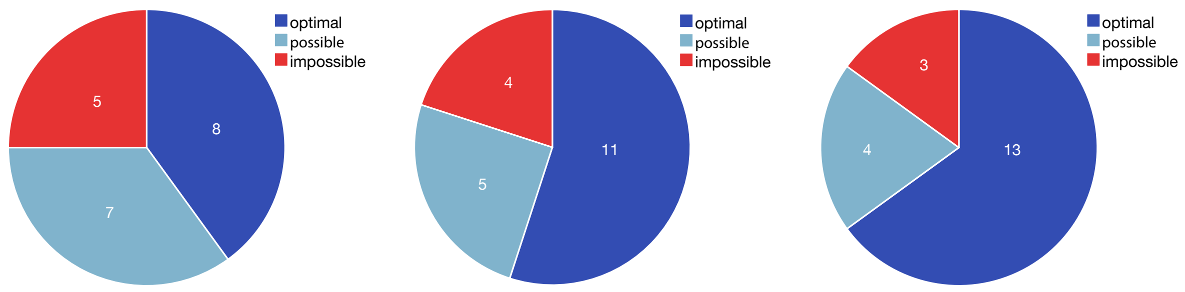
does not show significant improvement with available AR vision. This is the case for the length of the incision (CutLen) ($t_{(.05,19)} = -0.910$, $p \geq .05$), the orientation of the incision CutAng ($t_{(.05,19)} = 0.598$, $p \geq .05$), the centroid of the cut ($t_{(.05,19)} = 1.477$, $p \geq .05$) and the distance between the centroid of the incision and the target vertebrae PathLen ($t_{(.05,19)} = 1.614$, $p \geq .05$).

With respect to the second hypothesis, we analyze a potential learning effect due to the AR experience and compare the measured outcome of PPP1 with PPP3. Analysis of QualPos($U = -1.807$, $p \leq .05$) as well as QualArr($U = -1.933$, $p \leq .05$) shows significantly improved performance. With respect to the analysis of metrically scaled criteria the probands' cut length CutLen does not become significantly closer to the one of the expert($t_{(.05,19)} = 1.549$, $p \geq .05$). This is also the case for the orientation of the incision CutAng($t_{(.05,19)} = 1.326$, $p \geq .05$). In contrast, CutCent is positioned significantly better with present AR vision ($t_{(.05,19)} = 2.938$, $p \leq .05$). Also, the distance of the centroid of the incision to the target vertebrae PathLen becomes significantly closer to the one of the expert ($t_{(.05,19)} = 2.553$, $p \leq .05$).

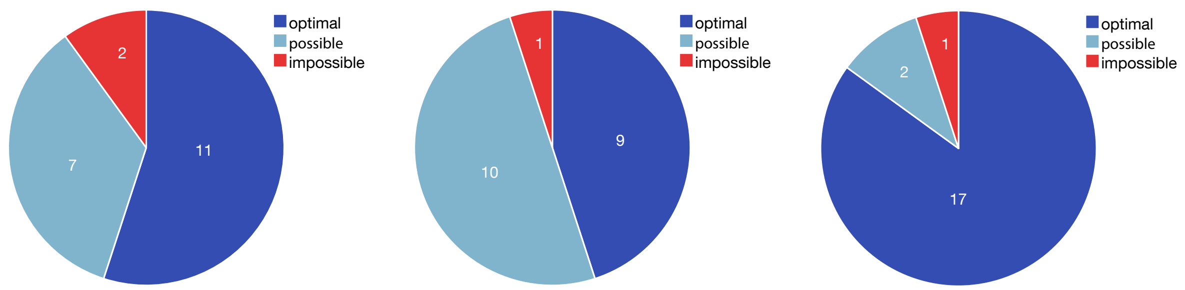
Further exploratory data analysis shows that CutCent and QualPos positively correlate in PPP1 ($r_s = .760$, $p \leq .05$), PPP2 ($r_s = .728$, $p \leq .05$) and PPP3 ($r_s = .779$, $p \leq .05$). QualArr and CutLen positively correlate in PPP1 ($r_s = .791$, $p \leq .05$), PPP2 ($r_s = .478$, $p \leq .05$) and PPP3 ($r_s = .575$, $p \leq .05$).

5.3.3.4 BFS-test

Statistics of the BFS-test are shown in table 5.8. The results also incorporate the answers of two expert surgeons who helped to optimize the experimental setup, however, they did not fully complete all answers. For this reason 20 to 22 responses were given for each questionnaire item. The analysis of the BFS-test shows neither a strong positive nor a strong negative change of the probands' mood. There is a slight tendency that they feel more *exited* (7.1%) and *contemplative* (5.8%), which are neutral evaluated categories of the BFS-test with a tendency to *arousal* rather than *resolution*. Figure 5.11 shows the results of the BFS categories.



(a) QualPos in PPP1 (b) QualPos in PPP2 (c) QualPos in PPP3



(d) QualArr in PPP1 (e) QualArr PPP2 (f) QualArr PPP3

Figure 5.10: QualPos and QualArr indicate for lumbar spine surgery slightly better performance with AR vision. Candidates seam also to learn from the AR experience to better position their ports.

5.3.3.5 Interpretation and Discussion

With respect to the results, AR vision being used for port placement procedure is mostly beneficial when the operation site is rather deep seated like vertebrae in thoracic spine surgery than superficial close to the skin layer like scew osteosynthesis. In contrast to superficial operation sites, the trajectory of instruments has to pass several layers and regions at risk to reach the target region. Target areas close to the skin allow the surgeon for a broader spectrum of possible incisions due to the soft skin. In particular for pelvis surgery, we noticed that there were strategic differences when defining ports. One part of probands rather decided to follow geometric arguments and positioned the port at the intersection point of the trajectory that passes the pelvis bone structure (see figure 5.7(a)). Others mentioned after the experiment that they would cut in an area being less painful for the patient and move soft skin to insert the instrument rather than define the port position according to geometric criteria. For lumbar spine surgery, having a target region being less deep seated than the one of thoracic spine surgery, we identify at least statistical tendencies that argue for AR vision to support the procedure of port placement.

In general, we did not notice any disadvantage due to AR vision. Comparing data of PPP1 and PPP2 shows at least similar performance qualities, however, in all cases a

Table 5.8: Statistics of the BFS test are given as M(SD) and the percentaged difference.

	Before experiment M (SD)	After experiment M (SD)	Difference %
anger	0.17 (0.06)	0.18 (0.11)	1.17
excitement	0.91 (0.26)	1.32 (0.55)	7.11
activity	2.44 (0.21)	2.61 (0.16)	-1.17
elation	2.67 (0.30)	2.35 (0.44)	3.42
calmness	2.23 (0.48)	1.72 (0.36)	-3.12
contemplation	1.00 (0.25)	0.76 (0.48)	5.77
fatigue	0.85 (0.21)	0.30 (0.06)	-3.62
depression	0.29 (0.09)	0.14 (0.07)	-0.49

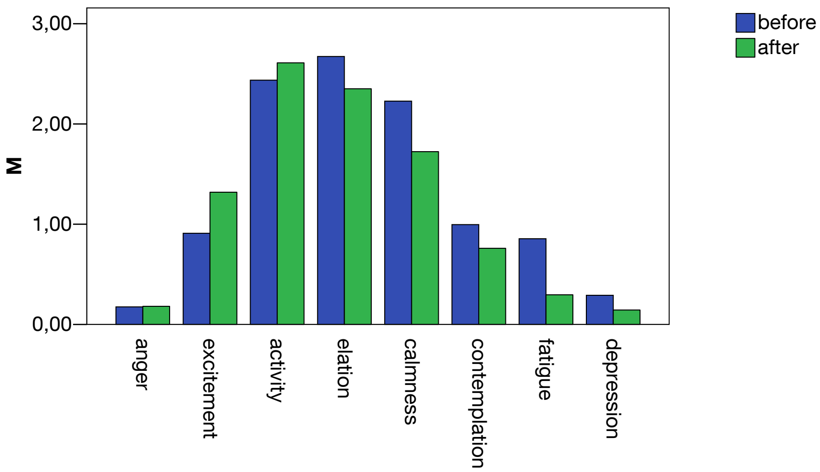


Figure 5.11: BFS-test

positive tendency with available AR vision. Data of all interventions shows that equal or less impossible port positions or more optimal port positions can be achieved (QualPos).

Except for the port placement of pelvis surgery, a significant positive effect of previous AR experience in combination with the presentation of expert ports and previous probands' port arrangement can be shown when analyzing data of PPP3 that is performed again without augmented views. For both lumbar and thoracic spine surgery less impossible and more optimal port positions QualPos and arrangements QualArr can be shown.

For both hypotheses, metric criteria tend to support the results of qualitative judgment of samples. However, we want to note that metric criteria rather shows the similarity to only one expert port placement while qualitative criteria (QualPos and QualArr) takes also slight variations of ports placed by the expert, that still allow for sophisticated instrument handling and access of the target region.

Since the physical patient model (VKHP) consists of rigid material the ability to touch and feel anatomic structures such as the ribs or the pelvis is reduced. However, the available tactile shape of the hip and the chest as well as a life size human skeleton model placed nearly could be used by the subjects to help placing the ports without AR vision.

The evaluation of the learning effect requires further studies involving a control group that learns port placement without AR vision. However, tendencies shown with the present experiment may indicate a benefit from AR views to create a mental model of the target anatomy that incorporates optimal reference ports of an expert in comparison with reviews of earlier own performance. We noticed that time is required to explain the complex technology on their head, for them to understand the nature of AR and to begin to immerse themselves in the AR scene and take the full advantage of the technology to allow to create a mental model of anatomy. For this reason, in a subsequent experiment, subjects are supposed to learn how to place ports with a longer training period in which they can carefully familiarize themselves with the AR vision.

In our survey that has been arranged after an experiment, subjects were asked whether they were able to view the AR scene in stereo. We noticed that some of the subjects either could not take the advantage of the stereo images due to individual vision problems. However, we also discovered that some subjects simply did not wear the HMD correctly so that one eye could not get exposed with the images of the corresponding display. For this reason, we want to suggest, that the test can become inherent part of every future experiment incorporating the HMD in order to ensure that the HMD is correctly mounted and subjects evaluate targeted items with the best configuration of the AR system.

A simple extension of the present study can be the evaluation of procedure specific trajectories for instance to drill the pelvis. This will incorporate an intraoperative planning procedure, that potentially benefits from AR vision.

CHAPTER

SIX

DISCUSSION AND OUTLOOK

This chapter starts with the introduction of concepts and strategies in section 6.1, how the AR system can be smoothly transferred into its designated working environment in order to start its clinical evaluation. Section 6.2 proposes further promising medical applications for head worn and alternative AR systems. For some of these applications, collaborations between computer scientists and physicians have already been established, while others are still in the conception stage.

6.1 Stepping into the OR

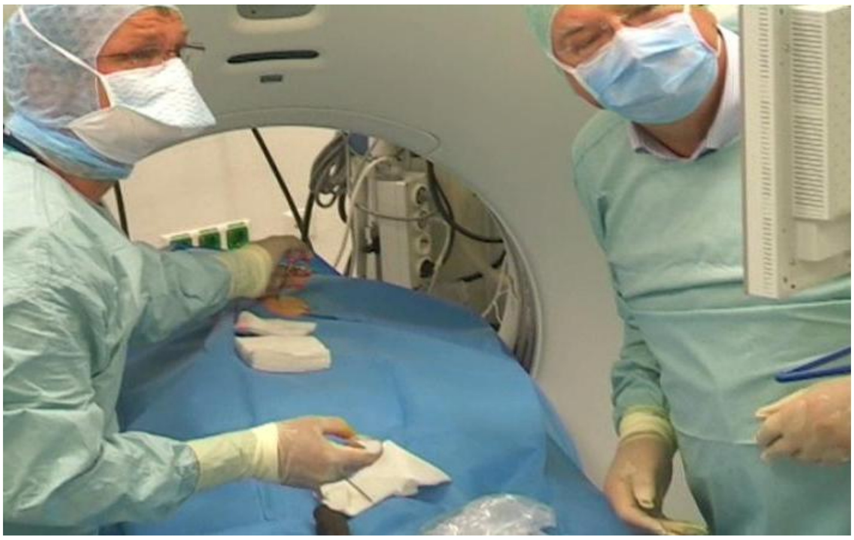
The final goal of the project is to evaluate the impact of the AR system in its designated working environment. Earlier works have reported the clinical evaluation of head worn AR systems. However, researchers have either avoided invasive procedures [15] or existing systems such as operating microscopes [29, 120] have been equipped by only minor modifications with AR capabilities. Such add-on solutions facilitate the integration of AR into the operative environment. Issues such as the acceptance of a new technology in a highly organized working environment, space restrictions and sterilization can be handled easier. However, innovation and creativity for the development of a new solution that takes full advantage of AR technology can be restricted due to the implicated limitations of an existing product that has been created on earlier conditions for solving a different problem. For instance, optical systems [72, 56] present the operation site in the highest perceivable resolution. However, the most promising video see-through approach for medical AR applications regarding synchronization, image composition, and registration of real and virtual objects [194, 182] can not be applied for such systems.

In a first feasibility study [23], the present AR system has been installed in an operating environment. In order to analyze potentially existing problems that do not become obvious in our lab space and to plan the smooth integration of the system into the workspace to evaluate a broad set of applications, we first chose a simple procedure in spine surgery. Vertebroplasty might not be the *killer application* for which AR technology could really revolutionize therapy, however, there are several good reasons why vertebroplasty was chosen.

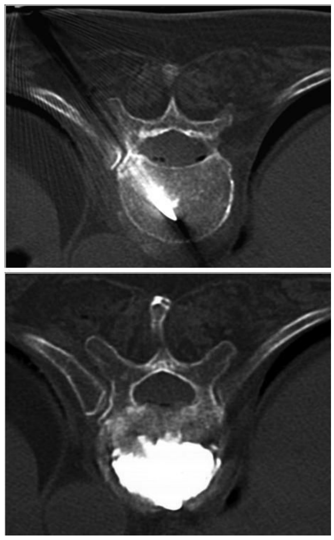
First, vertebroplasty is a minimally invasive approach without any direct view of the operation site. Second, frequently updated volumetric imaging data is available to be prepared for individual visualizations adapted to the current working step. Third, surgeons have to deal with two focuses of attention as shown in Fig. 6.1. One workspace is the monitor showing imaging data of the anatomy. The other workspace is the patient lying on the bed. In addition, vertebroplasty is a frequently performed common standard procedure of our clinical partner and has a clearly specifiable workflow. Finally, the operating environment is not highly staffed with standard equipment and surgical personnel, which leaves psychological and physical elbowroom to install the prototype AR system.

Vertebroplasty is a minimal invasive spine surgery with the objective to insert bone cement into weak and brittle vertebrae through a trocar for stabilization. First, the trocar has to be positioned dorsal above the target vertebra and then inserted by hammering until it reaches the target region. Second, bone cement is carefully injected through the trocar. Too little bone cement would leave the vertebra in an unstable condition that potentially requires an additional surgery for reoperation. Too much bone cement would affect, in the worst case strangulate, sensitive nerves and vessels around and within the spinal canal. Physicians control and discuss each step with frequently updated x-ray scans that are presented on a monitor (see Fig. 6.1). The most critical phases of the vertebroplasty are the penetration of vertebrae with the tocar and the injection of bone cement, which mark entry points for the AR system. During these phases, AR can be employed more intelligently by adapting the visualization to particular surgical work steps. For example, the opacity of augmented scene could be changed, depending on whether the surgeon is hammering (focus on real-world scene) or whether the surgeon is evaluating penetration depth (focus on augmented vision). Based upon the current phase, different objects should be furthermore segmented and highlighted automatically. During trocar insertion, the location of the tool tip could be pointed out while during bone cement filling, the spreading of bone cement in the vertebra could be segmented and given special focus in the augmentation. Apart from that, additional visual cues such as a virtual mirror [156] could be offered to the surgeon through a context-sensitive user interface at given times. We have identified more potential features for such a user interface, many of which being particularly helpful to inexperienced surgeons or surgical trainees. Tips for tool handling could be displayed in the augmentation during corresponding work steps, such as the direction of the trocar tip for proper steering of the tip to the desired location during *hammering*. Furthermore, an augmented stopwatch could indicate remaining stirring time during the temperature- and time-critical task of bone cement compound mixing. Monitoring the patient's O2-saturation during bone cement injection and augmenting it in the user interface is another medically relevant task, since a sudden rise can indicate an impending complication.

For a smooth adaption of the clinical staff to the AR technology, it is planned to stepwise install AR systems in the OR. Three stages with different invasive characters are considered as shown in figure 6.2. In a first stage a synchronized sensor data set consisting of a video, tracking and imaging data shall be captured before the operation starts. This procedure may be called AR Tomography (ART). As soon as the surgeon has to start the operation, we remove the equipment from his working space. The data can then be used



(a) Mental fusion of two workspaces



(b) Radio graphs

Figure 6.1: An anesthetist, the trauma surgeon and a radiologist are the protagonists in the operating theater. The latter two check (upper 6.1(b)) the position of the trocar and the amount of bone cement (lower 6.1(b)) with slices views presented on a monitor.

to develop and evaluate offline different visualization techniques to communicate imaging data and navigational information with the corresponding pathological case. These offline scenarios can then be iteratively discussed with the operating surgeon after surgery and to improve the AR UI according to his feedback.

In a second stage an AR system will be present during the whole operation. For this reason, a color camera and a tracking system will be installed and the AR scene consisting of augmented instruments and the augmented patient is presented on a monitor. With this system we avoid a complex setup. The surgical staff can slowly adapt to the augmented views and the surgeon can predefine potential applications that might be worse to present the AR scene with an HMD.

In the third stage the HMD system will be brought to the OR to be used following the novel concept of the *shadow surgeon*. The *shadow surgeon* is a member of the clinical staff, for instance an assisting surgeon or a medical student, watching the intervention through the HMD. Although the *shadow surgeon* is not actively contributing to the operation and influencing its outcome, we hope that a discussion after the surgical procedure among clinicians with different vision on the same workspace can result in precious feedback for further development of the system, better acceptance of medical AR, detection of benefits for vertebroplasty and determination of further promising clinical applications. This concept would allows us to perform clinical experimentation without the need for extensive ethical approval.

With robotics systems for image guided surgery in mind, Cleary notes that part of the reason why such systems are only sparsely used in real life ”is that the medical environment is a very complex one, and the introduction of new technology is difficult” [48] In contrast to robotics systems, which require the physician to sit at a remote controller desk to

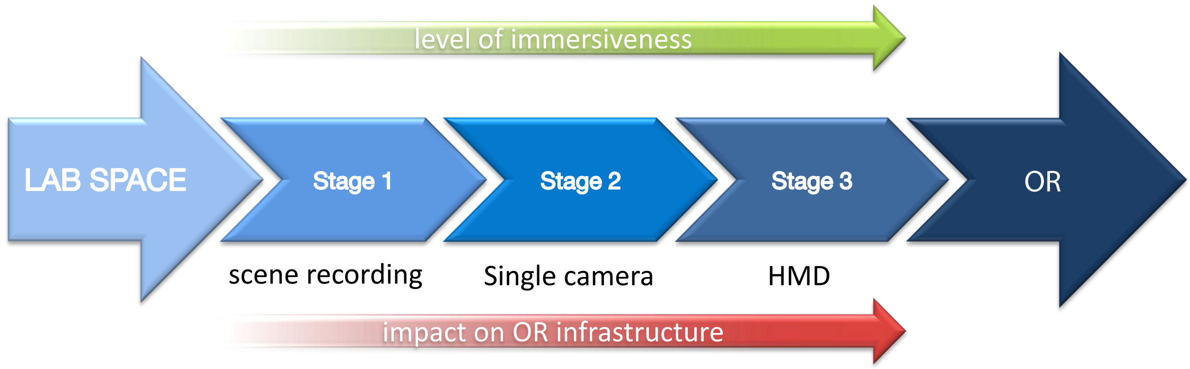


Figure 6.2: Three stages to transfer AR smoothly from lab space to OR space.

navigate the robot above the operating table, in our case the surgeon only puts a device on his head and the patient is still within reach of his hands. Anyhow the integration of the proposed AR system into the OR means a big impact to the existing infra structure. Only a simple to use, compact, mobile and robust system with a sterilization concept, if required, has the chance to survive the "baptism of fire" and to get accepted by the surgical staff. It is furthermore of high importance to provide an exit strategy, which means that the system can be quickly removed when not working or not needed without disturbing the surgical workflow. Part of the third integrative stage will be the development of a cart onto which the entire AR system can be installed. The cart will be equipped with a flexible arm construction to first compensate the weight of the prototype HMD and to be able to quickly remove it from ones head when being in the way. In addition, also the tracking system has to be installed on that cart to track the patient and used instruments over the shoulder of the *shadow surgeon*.

6.2 Future AR Supported Instructional Applications in the Medical Field

Contextual in-situ visualization may have an impact in different application areas in the medical field. The following section will introduce further ideas how to use AR as an instructional tool.

6.2.1 Teaching Anatomy, Physiology and Pathology in Formal and Informal Learning Environments

The education of anatomy, physiology and pathology is an inherent part of the medical curriculum. Some of the traditional media forms reviewed in section 1.1 embed a certain feature to be learned within the context or surrounding anatomic structures to better create a big, sustained picture of the body. AR is predestined to render 3D anatomy in context with the real body. Learning where to place beneficial ports for different procedures as described in section 5.3 is only one of many instructional units that may benefit from the context preserving view into the body. An exemplary learning unit of todays medical curricula of massager and physiotherapists that comes visually extremely close to AR views is known as "anatomy in vivo" [176, 177]. This instructional involves the superimposition of anatomy renderings onto the skin of a volunteer. However, the rendering is actually manually drawn with color pencils. The learners enhance sensory information from palpation with visual information from the drawing. If a future AR system is capable of registering deformable patient based, virtual anatomy with the body of a real volunteer, the application of AR to this learning unit becomes obvious. Besides the work referenced in section 5.1.1 that mainly employs head worn see-through displays, other groups of the AR community have proposed alternative AR systems to instruct human anatomy related subjects. However, only few user studies have been arranged to actually prove the instructional impact of AR.

Lenz et al. [129] present the interactive four page magic book "My Inner Body". The first application/lesson aims to explain nutrition on its way through the alimentary system. Visualization techniques such as exploded views of 3D organs and their interactive coloring, labeling and animation using real buttons installed in the book have been realized. An audio guide asking intermediate questions to reach the next level/page provide the user with acoustical instructions and feedback. Instead of text or illustrations the magic book contains a set of ARToolkit markers for augmenting the pages with animated 3D organs. A user study with 16 participants has been arranged to control and prove the positive learning effect of the proposed instructional tool. Thomas et al. [214] present an AR tool to teach anatomy using ARToolkit makers that are attached to the removable components of a plastic male torso model. Due to the markers, virtual descriptive labels can be registered with the organs. The pose information from tracking data is also used to control a volume rendering of CT data presented in an virtual environment. The user interfaces allows for cross sectioning of the volume rendering for a better exploration of the volume data. However, the volume rendering was not superimposed onto the anatomy of the plastic torso. Thomas et al., 2007 [215] and John et al., 2007 [106] from the same group created anatomic models from CT data, e.g. of the liver, using rapid prototyping. Their advanced approach allows now for registering a volume rendering of the CT data with the RP model and present the AR scene with a HMD. With a hand held tracked device the user can virtually resect portions of the anatomy rendering. So far, no user study has been presented. Hedegaard et al.[96] report on an AR system for learning of electrocardiography (ECG) analysis (EKGAR). Interviews with medical students reveal the limitations of traditional learning material. Hence, the authors propose a single camera based AR system to augment the heart onto an ARToolKit marker. The system allows for interactive transformation and clipping of the augmented heart. According to the authors, the collocation of EKG diagrams next to the AR scene, the interactive visualization of the heart and the spatial awareness of the anatomy due to the 3D presentation, using the proposed AR system "becomes a constructionistic experience". Except interviews and "a handful of user evaluations" the authors present no user study that might approve the benefit of the proposed system. Juan et al., 2008 [109] propose a see-through HMD based anatomy learning tool for children. Their systems uses 2D images of the interior body that are registered in AR space. A user study with children as subjects is presented, which compares the instructional impact of a monitor based system and the AR system. However, authors only report few information of the experimental setup and quantitative results.

6.2.2 Teaching Diagnostic and Surgical Procedures

The data set of the VKH [164] that has been chosen to create the VKHP described in section 2.2.2 comes with an MRI, CT and photographic data. When performing and visualizing US Simulation from CT data as part of the AR application, a wide spectrum of imaging modalities that is frequently used in clinical routine can be presented simultaneously. For this reason the AR system can become a powerful learning tool for the instruction of using and interpreting different imaging modalities. In particular, the interpretation of ultrasonic imagery can be difficult for untrained users. When providing a joint data presentation combining imagery from all available imaging modalities as slice views, the user can browse through the body of the Visible Korean Human and identify and compare structures having different appearance in different data sources. In contrast to the monitor based VKH viewer [1] the introduced AR visualization presents the anatomy right on a three dimensional patient model. The surgeon can palpate the body in conjunction with scanning it using various visualization methods to present the anatomy. 3D/3D registration methods allow for replacing certain parts of the data sets with pathological counterparts to adapt the learning experience to real tasks in the medical professional life.

The VKHP phantom is equipped with several removable windows to access the operation sites of various traumatological procedures (see section 2.2). In particular surgical drilling may benefit from AR views. Members of the group started to target the preparation of drill canals to implant fixateurs and stabilize fractured shoulder joints. Applications in hip, neuro and spine surgery will follow. The phantom can not only be used to evaluate the system in terms of finding the most suitable visualization of anatomy and navigation information. It can also be considered as a training environment to teach physicians being unfamiliar with a target procedure. Using the AR system and the VKHP, surgeons can repetitively train most parts of the procedure. They have the drill in their hands and can observe the drill progress inside the bone by using the AR views.

Many current medical simulators target only a small part of a procedure, sometimes only one working step. However, "most medical errors result from problems in the systems of care rather than from individual mistakes". In contrast to other simulation domains "medical education has neglected the importance of teamwork" [195]. AR might fill the gap to create a team-oriented training environment consisting of virtual and real items. Low et al combine a HMD and a projector-based displays to create a surgical training scenario that involves learning "to work in an environment that is realistically filled with both necessary and distracting objects and events" [135]. The wearer of the HMD can experience the perspectives of different virtual protagonists standing around an operation table, while getting cognitive input from additional sources of the 3D scene. Dieker et

[1] Visible Korean Human Project: http://vkh3.kisti.re.kr/

al. [52] proposed to control virtual avatars by capturing the motion of an actor to fill a class room with students. Here teachers are prepared for difficult situations in their future professional life. This concept can be transferred to an intraoperative training scenario. The learner, wearing an HMD can be for example the operating surgeon or the nurse. He has to interact with her colleagues that can be virtually placed into the scenario and played by an actor. This training concept may reduce costs since parts of the OR infrastructure can be replaced and modified using virtual objects while for example the operating table remains real. In addition, only one actor can play several roles by switching to another avatar.

6.2.3 Teaching the Patient

Frank Moss from MIT media lab identifies new media forms to enhance communication between patient and medical expert as a major issue to improve the current health care system. He mentions three major principles that motivate the establishment of various research projects in this direction. First "patients are the most underutilized resource in health care", second "the revolution must take place in our everyday lives, not in the doctor's office or the lab" and third "information transparency, not just information access, is the solution" [2].

AR technology can have a major impact in improving the knowledge transfer. Bluteau et al. [34] proposed to use a projector based AR system that superimposes virtual anatomy onto the patient's body. On one hand the system should support the physician to explain his diagnosis and consequential treatment "in a simple way, even if the patient has no anatomic knowledge". On the other hand the patient should be supported to explain "the feelings that are difficult to localize and specifiy" [34].

Together with our clinical partners at the *Department of Plastic and Reconstructive Surgery, Klinikum Rechts der Isar, TUM, Germany* a project has been launched that aims at developing an AR system for preoperative and postoperative patient education and consultation. The planned target system is a digital mirror having also been introduced as *magic mirrors* [67] [149] or virtual mirrors [3]. Such mirrors usually consist of a projection screen and a video camera attached to the screen. The video data showing the environment in front of the screen is then projected onto the screen to create a mirror effect. Virtual objects can be registered with a person standing in front of the screen.

Our currently targeted medical application is breast reconstruction. The system is supposed to support physicians to explain to patients the limitations, risks and potentials of an intervention, the surgical procedure itself, potential complications, and prescription for preventive and postoperative behavior, treatment and medicamentation. It is necessary for a patient to completely understand the treatment to consider and communicate the personal value they place on the benefits versus the harms [160].

Today's standard media providing the patient with such information is a printed brochure. Both, the physician and the patient review together the items of this brochure, which are mainly based on text and 2D sketches. An example of such brochures being used

[2] MIT Media Lab: http://www.media.mit.edu/research/groups/new-media-medicine

[3] Adidas Virtual Shoe-Fitting Mirror

in Germany are the information and consent sheets of the company proCompliance [4] or Perimed [5]. This form is then signed by the patient in order to confirm that the information was entirely transferred for legal protection of the surgeon and the hospital. However, in fact in many cases, time is scheduled too shortly to ensure a complete, informative conversation and information is often too complex for the standard patient.

The workflow of conventional therapy in breast reconstruction can be described with the flowchart shown in Fig. 6.3. In a first stage, the surgeon explains to the patient all applicable types of breast reconstruction, their limitations, potentials and risks. During this conversation the expectations, wishes and fears of the patient are discussed. Also the psychological and medical background of the patient is investigated to evaluate whether the patient is capable of withstanding the operation. In most cases the patient takes extra time to think well about all provided information and consequences. Usually there is a second stage during which the patient decides on the type of procedure fitting best to her situation.

When the decision is made, the patient has to sign a legal document declaring that she is fully aware of the risks and consequences of the operation. Then photos of the patient are taken for documentation. In the third stage the breast reconstruction is performed sometimes taking up to six hours. In many cases, a second surgery is performed to adapt the shape and size of the breast for a symmetric result and/or to restore the mamilla (fourth stage). After the operation(s), again photos of the patient are taken to compare the preoperative situation with the postoperative outcome. During the fifth stage the healing process of the patient is monitored by medical staff.

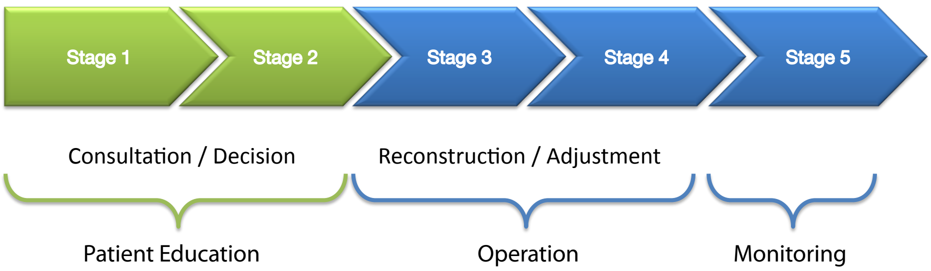


Figure 6.3: Therapy workflow for breast reconstruction

The system focuses on supporting the first and the second stage of the therapy workflow, when the patient has to understand complex information, decides on the procedure and gives the surgeon the approval of being aware of all legal information. We believe that the transfer of complex information and communication barriers caused by medical terminology being usually not familiar to the standard patient can be better managed when the briefing procedure currently performed with conventional spoken, written and 2D illustrated explication is augmented with information presented in-situ on the patient's body.

[4] Thieme Compliance GmbH
[5] perimed Fachbuch Verlag

First concepts of how to present different sources of information in terms of technical but also instructive methodology have been presented at [247, 248].

CHAPTER

SEVEN

CONCLUSION

The present doctoral thesis *Immersive, Interactive and Contextual In-Situ Visualization for Medical Applications* addresses the application oriented improvement and evaluation of an innovative Human Computer Interface Design (HCID). The targeted Medical Augmented Reality (AR) system is supposed to support intraoperative data visualization and instrument guidance as well as instructional units of the medical formation. The two major aspects that have been addressed by this thesis aim at the seamless composition of virtual and real objects and the interaction with these objects in the designated medical AR environment. Both aspects aim at the enhancement of the user's immersion into the AR scene.

With respect to the fusion of reality and virtuality, two approaches of contextual in-situ visualization have been introduced to overcome the problem of missleading depth perception in AR worlds (see section 2.3.1). Both approaches introduced in sections 3.1 and 3.2 advance techniques known from the research domain focus and context visualization in virtual worlds to make them applicable and useful for medical AR scenarios. Feedback from many users who have experienced the contextual in-situ visualization or have seen related video and images confirm the improvement of image composition compared to previous attempts proposed by the literature. However, no objective, quantitative evaluation of the effect of the proposed methods on depth perception has been shown so far. The complex setup of the prototype AR environment involves many hidden factors that might bias the targeted effect to be measured. Such factors can be for example caused by the heavy and bulky prototype HMD consisting of many cables and sensitive devices. Also time pressure, tiredness or fear of using complex technology that might break can bias measuring results. One major tasks of future research will be the interdisciplinary development of an evaluation environment that assesses the quality of different visualization methods. In particular the evaluation of perceptive effects needs intensive collaboration with psychologists having a strong background in perception and methodology to guarantee reliability, validity and objectivity of the upcoming user studies.

As a second contribution, a set of interaction techniques has been evaluated that is supposed to interact with objects of the AR scene. First, three physical interfaces presented in section 4.1 have been compared that allow for the control of continuous and binary parameters. The result of a user study entails to rather apply voice recognition and a foot pedal than hand gestures to control the parameters.

Second, a Virtual Mirror, which is a virtual but tangible interface has been introduced and evaluated in section 4.2. Results of a user study show that a virtual mirror providing a second view of the operation site allows for more accurate instrument guidance in minimally invasive instrument settings.

Finally, two user studies are presented that address surgical applications that have been proposed by partner surgeons. The first user study concerns navigated drilling in spine surgery. The sample shows a positive tendency for AR driven navigated drilling compared to a classical monitor based system. Port placement has been selected as the second target application. A user study with 20 surgeons and medical students shows that AR vision being used for port placement procedures becomes beneficial when the operation site is rather deep seated like vertebrae in thoracic spine surgery and the trajectory of instruments has to pass several layers and regions at risk to reach the target region. Besides the technical advances and user studies that have been presented with this thesis, a second objective of this work is to point out the potentials of see-through head mounted displays providing AR for medical purposes. The introduction of this thesis highlights different media forms that have been proposed in the past to present anatomy for the generation of a sustained mental model of the body. Consequentially, different criteria has been determined to better distinguish the capabilities of such media forms (see table 7.1). It was evaluated whether

- the observer can interact with or interactively manipulate the anatomy,
- single anatomic components are presented in topological context with the complete body,
- the anatomy is presented in its natural dimension (3D or even 4D)
- the media can be used for surgical training and the anatomy that is manipulated during the training procedure can be reproduced,
- the media is or has the future potential to be accessible for a broad audience,
- the approach is flexible in terms of being capable of showing in addition to the standard anatomy also anatomic anomalities or pathologies,
- the anatomy can be explored in the context of a living subject.

Once this relatively young technology has progressed and found a first market, not necessarily the medical market, that invests money and efforts to improve this innovative human-computer interface, AR has the potential of fulfilling all the identified criteria. Users of future HMD-based AR systems will observe anatomy in context with the living body, e.g. of the patient to be examined or their school mate during the anatomy lesson. They will be able to walk around this body to gain all perspectives of the 3D subject. It will be possible to interactively exchange virtual anatomy with pathological counterparts served from a large anatomy database. Medical students can gain and surgeons refresh

Table 7.1: Analysis of different media regarding their capability of communicating anatomy and pathology. Criteria/media combinations marked with ”O” exist, however, are rare.

criteria/media	Dissection	Drawing	Ceroplasty	Plastination	Rendering	AR
interactive	✓	⊖	⊖	⊖	✓	✓
contextual	✓	✓	✓	✓	✓	✓
3D	✓	⊖	✓	✓	⊖	✓
operable	✓	⊖	⊖	⊖	✓	✓
reoperable	⊖	⊖	⊖	⊖	✓	✓
accessible	⊖	✓	O	O	✓	✓
variable	O	✓	O	O	✓	✓
alive	⊖	⊖	⊖	⊖	⊖	✓

knowledge of operative procedures by operating with semi-virtual instruments semi-virtual anatomy. Full control of virtual settings of the training environment will allow for after action reviews and endless repetition of training sessions. Recent trends of new game controllers released by the game industry illustrate the potential key to also make AR technology accessible to a wide audience.

The research community has proposed several medical application domains that may take advantage of see-through AR systems. Beside diagnostics, pre- and intraoperative planning and intraoperative navigation, also training scenarios have been proposed as described in chapter 5. Future applications of this portfolio may become the communication between physician and patient, training of surgical teams, and teaching anatomy in formal and informal learning environments (see section 6.2).

Within discussion with medical partners and colleagues, different important aspects have emerged that may be important when developing future AR user interfaces to be used in the surgical workflow. Some of those aspects may also be valid for other application domains or other technology.

1. **Workflow-driven application of AR:** It is important to understand, that the AR system won't be used during the whole procedure. The AR system is considered as a surgical tool like a scalpel. It is provided when needed and removed when dispensable. Situations during the workflow that might benefit from AR views can be automatically determined by a major workflow analysis of a particular intervention. A corresponding analysis system uses several sensor modalities. This has been done by members of the group for instance for vertebroplasty [3, 4, 23]. Using the data from workflow analysis allows for predicting certain working steps to prepare the system, for instance to update and change the visualization of imaging data or a

new augmented surgical instrument. Predictable information can also serve non-AR related tasks within the clinical workflow. For instance, the time and room schedule can be better planned or the next patient outside the OR can be prepared and anesthetized.

2. **Interdisciplinarity:** It is also important that a common physical, interdisciplinary platform is provided such as a *real world lab* to get to know and understand each other. In the case of collaboration between surgeons and engineers or computer scientists, it is essential that the latter group physically moves their workspace close to the surgeon's workspace. The time schedule of physicians is usually tight. As a computer scientist one should not expect that physicians, will have the time for frequent meetings or to spend much time on traveling to ones own remote research and developer lab. It has to be clear that it is the job of the engineer to physically bring his gadgets to the designated medical end-user to let him decide whether he wants to use it or not. Engineers and computer scientists have to understand the problems to be solved of the application domain. For this reason, they have to experience the medical working environment and the workflow of a target procedure. In the same token physicians have to get to know what is technically possible. For example they might find additional applications for an existing software or hardware solution that was realized for another hurt.

3. **Importance of the User Interface:** It is important to understand that only a sophisticated user interface will create an impact in the medical workflow. A proposed technical method that might improve a certain working step has to be provided in a way so that it can be easily controlled and used by the end-user. In addition, it has to fulfill the infrastructural requirements such as sterility rules. Otherwise the probability that the technology or features of it are really used would extremely decrease. Head worn AR devices might be the "ultimate display" [210] for many steps in the surgical workflow. However, see-through HMDs are still in a research stage. Mostly prototypes are available. Recent trends may result very soon in increasing markets in particular in the military and entertainment field. Once there is a market that demands lightweight, ergonomic HMDs, such devices will initiate and maybe revolutionize various application fields in the medical domain. As a preliminary solution one may think of a hanging HMD that is mounted like an OR lamp to a flexible arm construction attached to the ceiling. The user can pull the HMD onto his head when AR views are helpful and push it away when being obstructive.

4. **Stepping into the OR:** Besides the operating surgeon, there are many more protagonists in the OR and each one plays an important role within the surgical team. At first the operating surgeon has an interest to get support for his tasks, however, the operation is rather a complex network of interconnected procedures performed by different persons that have to be synchronized to finish the job. The indication of an additional device may have an impact on many of those sub procedures. It may affect not only the work of the surgeon, but also the job of the anesthetist or the assisting nurse. For this reason it is important that the whole surgical team gets to know

Workflow Orientated Application of AR

The AR system is considered as a surgical tool like a scalpel. It is provided when needed and removed when dispensable.

the objectives of a project that influences their jobs in order to achieve tolerance and acceptance. The transfer of the AR system into its designated intraoperative workspace has to be a smooth, iterative procedure as described in section 6.1. In particular in the initial phase of the integration, the AR system has to carefully augment currently available capabilities of the OR infrastructure without replacing them or restricting the staff's room for maneuver. Only if the surgical team identifies benefits due to the AR system that would implicate, however, justify a change of the workflow or the OR infrastructure, the integration of the AR system can be brought to the next invasive level.

One of the most important topics that have to be addressed by the research community in the future to further push AR to the identified application fields is the development of advanced see-through HMDs that have no eye to display parallax and provide large field of views, high resolution, light weight constructions and wireless data transfer. As Henry Fuchs pointed out at the AMI/ARCS workshop on the 24th of September, 2009 ”very slow improvements” in this area due a ”tiny research community”. Another field will be the ”real time model creation / scene capture” [1] to take the advantage of a completely geometrically known AR environment. This will facilitate many previous problems. With respect to surgical application environments, this is of interest for handling deformations of the patient due to breathing or operative treatment to ensure accurate tracking of the patient and registration of imaging data, composition of real and virtual entities and gesture based interaction. A third domain requiring progress will be the individualization of visualization and interactive features with respect to the user's preferences and the particular task that has to be supported.

I would like to conclude this thesis with motivating new talents having fresh ideas to further progress Medical Augmented Reality. In particular the user interface needs further attention to make this technology acceptable and useful to the final end users. On the telecommunication and personal PC sector touch-sensitive displays revolutionized man-machine interaction and show a strong request for a paradigm shift in designing human-computer interfaces. For other application domains, an ”ultimate display” [210] may become the user's choice.

[1] Henry Fuchs at the AMI/ARCS workshop on the 24th of September, 2009

CHAPTER

EIGHT

APPENDIX

A Authored and Co-Authored Publications

1. C. Bichlmeier, B. Ockert, S. M. Heining, A. Ahmadi, and N. Navab. Stepping into the Operating Theater: ARAV - Augmented Reality Aided Vertebroplasty. In *Proceedings of the 7th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 165–166, 2008.

2. C. Bichlmeier, S. M. Heining, M. Feuerstein, and N. Navab. The virtual mirror: A new interaction paradigm for augmented reality environments. *IEEE Trans. Med. Imag.*, 28(9):1498 – 1510, Sept 2009.

3. C. Bichlmeier, S. M. Heining, M. Rustaee, and N. Navab. Virtually extended surgical drilling device: Virtual mirror for navigated spine surgery. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2007, 10th International Conference*, Lecture Notes in Computer Science, Brisbane, Australia, October/November 2007. Springer Verlag.

4. C. Bichlmeier and N. Navab. Virtual window for improved depth perception in medical ar. In *AMIARCS - Virtual Window for Improved Depth Perception in Medical AR*, Copenhagen, Denmark, Oct. 2006. MICCAI Society.

5. C. Bichlmeier, B. Ockert, O. Kutter, M. Rustaee, S. M. Heining, and Nassir Navab. The visible korean human phantom: Realistic test & development environments for medical augmented reality. In *International Workshop on Augmented environments for Medical Imaging including Augmented Reality in Computer-aided Surgery (AMI-ARCS 2008)*, New York, USA, Sept. 2008. MICCAI Society.

6. C. Bichlmeier, S. M. Heining, R. Mohammad, and N. Nassir. Laparoscopic Virtual Mirror for Understanding Vessel Structure: Evaluation Study by Twelve Surgeons. In *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 125–128, Nara, Japan, November 2007.

7. C. Bichlmeier, T. Sielhorst, and N. Nassir. The tangible virtual mirror: New visualization paradigm for navigated surgery. In *AMIARCS - The Tangible Virtual Mirror: New Visualization Paradigm for Navigated Surgery*, Copenhagen, Denmark, Oct 2006. MICCAI Society.

8. C. Bichlmeier, F. Wimmer, S. M. Heining, and N. Nassir. Contextual Anatomic Mimesis: Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality. In *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 129–138, November 2007.

9. O. Kutter, A. Aichert, C. Bichlmeier, S. M. Heining, B. Ockert, E. Euler, and N. Navab. Real-time Volume Rendering for High Quality Visualization in Augmented Reality. In *International Workshop on Augmented environments for Medical Imaging including Augmented Reality in Computer-aided Surgery (AMI-ARCS 2008)*, New York, USA, Sept. 2008. MICCAI Society.

10. N. Navab, J. Traub, T. Sielhorst, M. Feuerstein, and C. Bichlmeier. Action- and workflow-driven augmented reality for computer-aided medical procedures. *IEEE Computer Graphics and Applications*, 27(5):10–14, September/October 2007.

11. N. Navab, C. Bichlmeier, and T. Sielhorst. Application no. pct/ep2007/003206: Virtual penetrating mirror device for visualizing of virtual objects within an augmented reality environment, April 2006.

12. N. Navab, M. Feuerstein, and C. Bichlmeier. Laparoscopic virtual mirror - new interaction paradigm for monitor based augmented reality. In *Virtual Reality*, pages 43–50, Charlotte, North Carolina, USA, March 2007.

13. T. Sielhorst, C. Bichlmeier, S. M. Heining, and N. Navab. Depth perception a major issue in medical ar: Evaluation study by twenty surgeons. In R. Larsen, M. Nielsen, and J. Sporring, editors, *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Lecture Notes in Computer Science, 2006.

14. F. , C. Bichlmeier, S. M. Heining, and N. Navab. Creating a vision channel for observing deep-seated anatomy in medical augmented reality. In *Proceedings of Bildverarbeitung fuer die Medizin (BVM 2008)*, Munich, Germany, April 2008. BVM.

15. P. Wucherer, C. Bichlmeier, M. Eder, L. Kovacs, and N. Navab. Multimodal medical consultation for improved patient education. In *Proceedings of BVM 2010*. Springer, Mar. 2010.

B Questionnaires and Material for User Studies

B.1 Evaluation: AR UIs

This section is related to the user study presented in section 4.1. The standard questionnaire USE (Usefulness, Satisfaction, and Ease of Use) proposed by Lund [137] (0=strongly disagree to 6 = strongly age) has the following items:

- Usefulness

  1. It helps me be more effective.
  2. It helps me be more productive.
  3. It is useful.
  4. It gives me more control over the activities in my life.
  5. It makes the tings I want to accomplish easier to get done.
  6. It saves me time when I use it.
  7. It meets my needs.
  8. It does everything I would expect it to do.

- Ease of Use

  1. It is easy to use.
  2. It is simple to use.
  3. It is user friendly.
  4. It requires the fewest steps possible to accomplish what I want to do with it.
  5. It is flexible.
  6. Using it is effortless.
  7. I can use it without written instructions.
  8. I don't notice any inconsistencies as I use it.
  9. Both occasional and regular users would like it.
  10. I can recover from mistakes quickly and easily.
  11. I can use it successfully every time.

- Ease of Learning

  1. I learned to use it quickly.
  2. I easily remember how to use it.
  3. It is easy to learn to use it.
  4. I quickly became skillful with it.

- Satisfaction

  1. I am satisfied with it.
  2. I would recommend it to a friend.
  3. It is fun to use.
  4. It works the way I want it to work.
  5. It is wonderful.

6. I feel I need to have it.
7. It is pleasant to use.

Additional questions that have been posed in the evaluation reported in section 4.1.2 are:

- Did you wear and test this or another HMD before? (0 = yes/1= no)(*Haben Sie zuvor schon einmal dieses oder ein anderes Head-Mounted-Display aufgesetzt und ausprobiert?*)

- How would you judge the usability of the proposed UIs for the productive application in the OR? (0=not useful at all to 6 = very useful)(*Wie schätzen Sie die Eignung der Systeme für den produktiven Einsatz im OP ein?*)

- Please explain your previous answer! (open question) (*Womit begründen Sie ihre letzte Entscheidung?*)

- Beside the reduction of the weight of the HMD, which problems have to be solved before the System becomes useful for intraoperative tasks? (open question)(*Welche Probleme neben der Gewichtsreduktion müssen für den Einsatz eines solchen Systems im OP Ihrer Meinung nach noch gelöst werden?*)

Complete answers of the extension of questions "How would you judge the usability of the proposed UIs for the productive application in the OR?" that asks for further feedback. In some cases the original answers are given in German. In these cases a translation into English is provided. Responses of probands with medical background are marked with **Med**.

Probands gave the following reasons for their UI preferences (QU3):

- **Med**: **PEDAL** is rather uncomfortable to use since the control of the position of the tool (virtual slide bar) becomes complicated and time- consuming. (*Das Pedal ist eher umstaendlich, da es kompliziert ist, wenn man zuerst das Tool schliessen muss und an anderer Stelle wieder oeffnen muss. Das braucht meines Erachtens ein bisschen zu viel Zeit.*)

- Interaction using the hands is more intuitive and precise than pointing with the eyes (moving the head). A combination of **HAND** and Sprache (**VOICE**) would be most beneficial in my opinion.

- **Med**: The HMD is heavy and gets warm. A pedal is practical, however, it is always difficult to find it in the OR. (*Head Mounted System ist schwer und warm. Fusspedal ist praktisch, muss man aber immer suchen im OP.*)

- It is not easy to control the cursor by moving the head.

- **PEDAL** is not within the field of vision. **HAND** restricts the vision and needs a free hand. (*Fusspedal ist nicht im Blickbereich. Handgesten beeintraechtigen Sichtfeld, bzw. benoetigen freie Hand.*)

- For me, **PEDAL** was the easiest UI to use. One has hands free and unnecessary noise, vocal sounds can be avoided. (*Das Pedal war fuer mich am einfachsten zu bedienen< man hat die haende frei, unnoetige geraeusche...gespraeche...werden vermieden.*)

- I think, there are already a lot of pedals in the OR and an additional pedal would be bothersome. **HAND** had problems to detect my hand and needed several tries. (*Ich denke, dass in dem OP raum schon sehr viele Pedale rumstehen und noch ein zusaetzliches Pedal stoerend waere. Die Hand wurde nicht immer gleich erkannt. Ich brauchte mehrere Anlaeufe, um die Hand zu erkennen.*)

- **Med**: Hands and feet are busy in the OR. (*Haende werden im Op gebraucht, Fuesse meistens auch.*)

- Restriction of vision with **VOICE**. Restriction of movement with **HAND** and **PEDAL**. (*Sichteinschraenkung durch Kopfbewegungen bei Sprachsteuerung, sonst Bewegungseinschraenkung bei Hand- und Pedalsteuerung*)

- **Med**: **HAND** had problems to effectively detect my gestures. Potential danger due to nonsterility when accidentially touching the HMD while moving the hand in front of the camera. Usually it is calm during the operation, for this reason **VOICE** is very comfortable, secure and fast. (*Die Handgesten werden nicht gut erkannt, man muss die Hand weit weghalten. Gefahr der Unsterilitaet bei Beruehrung mit Headset durch Heben der Hand vor die Kamera. OP meist ruhig, deswegen Sprache sehr komfortabel, sicher und schnell.*)

- **VOICE** needs no body motion. (*Die Sprachsteuerung erfordert keine Koerperbewegung*)

- Noise inside the OR. Aren't hands needed for other tasks? (*Laerm im OP. Braucht man seine Hand nicht fuer andere Dinge?*)

- The pedal approach is quite easy to use and learn, and it works. But in general I'm not sure how friendly it would be for the surgeon. The hand approach - in the end it does work, but it takes quite some efforts, and I think the visualization of the gray circle (cursor) should be more expressive.

- **HAND**: One has to remove the Hand from the situs, in addition it has not worked in a stable way. However, it is easier to remember gestures than voice commands or the pedal control. I think, it is more intuitive to touch objects I can see than talking to them. **VOICE**: Worked best and I could quickly finish the task. It is prefect to be used in the OR, because neither hands nor feet are needed and sterility can be guaranteed. However, I know that system being already used in the OR have problems with voice control due to environmental noise, which was no problem within lab conditions. However, it is difficult to remember many vocal commands. **PEDAL**: It is useful, however its too remote from the space I see and interact with. It has only few options for interaction. (*Hand: man muss die Hand vom*

Situs wegnehmen, um eine Geste durchzufuehren. Ausserdem hat die Handerkennung nicht gut funktioniert. Andererseits erinnert man sich an die Gesten besser als an voice commands oder an die Pedal-Befehle, es kam mir auch intuitiver vor, Dinge zu beruehren, die ich sehe, als sie anzusprechen. Sprache: hat sehr gut funktioniert und ich konnte das gewuenschte Ziel am schnellsten erreichen. Perfekt fuer den OP, da man weder Haende noch Fuesse braucht und Sterilitaet ist auch gegeben. Aber von echten Systemen im OP weiss ich, dass es mit den Umgebungsgeraeuschen nicht gut klappt. Hier im Labor war das ideal. Allerdings muss man sich viele Kommandos merken. Pedal: gut geeignet, aber Fuss ist weit weg von dem was ich sehe und bediene. Wenig Interaktionsmoeglichkeiten.)

- **HAND**: Does not work in most cases at the first go. There was a delay when moving the vision channel. **VOICE**: Worked good and is maybe easier to integrate into the intraoperative workflow than **PEDAL**.(*Hand: Funktioniert meistens nicht auf Anhieb. Lag beim verschieben. Sprache: Funktioniert gut und evtl. leichter in OP-Ablauf/Layout zu integrieren als Pedal.*)

With respect to problems of the AR system that have to be solved before it can be used intraoperatively the following answers are given. In some cases the original answers are given in German. In these cases a translation into English is provided. Responses of probands with medical background are marked with **Med**.

- **Med**: It has to be ensured that the HMD can be rigidly fixed to the user's individual head. (*Sie (HMDs) muessten so extrem fest auf dem Kopf sein, dass es auf keinen Fall verrutschen kann. Moeglichst genaue Anpassung!*)

- **Med**: The hand is difficult to be seen, because of the present optical zoom. For this reason it is difficult to coordinate the hand. (*Die Hand (alles andere auch) ist schwierig zu sehen, da sie naeher erscheint, als sie eigentlich ist. Deshalb ist es schwierig, sie zu koordinieren.*)

- Lag reduction, precision & jitter of tracking.

- Time lag reduction. (*Verzoegerung*)

- Frame rate of the video is low, which makes my eye uncomfortable soon (5 min later).

- Field of view should be increased. (*Blickfeld sollte groesser sein.*)

- Better fixation of the HMD to the user's head. Better cameras, since I saw the AR scene blurred. (*Noch bessere Anpassung an den Kopf, vielleicht noch eine bessere Kammera...z.t. etwas unscharf fuer mich (was machen Brillentraeger).*)

- Robustness (*Robustheit*)

- **Med**: Camera should have a bigger field of view. The offset of the cameras to the eyes is noticeable. I can imagine a combination of **HAND** and **PEDAL**, e.g.

PEDAL activates the gesture recognition, which would increase the sensitiveness of gesture recognition and reduce frustration of the operator. (*Die Kamera sollte mehr Weitwinkel haben, das Gesichtsfeld fuehlt sich sehr weit von den eigenen Augen entfernt an. Bei Steuerung durch Handgesten ggf. Aktivierung der Gestenerkennung durch Pedal (dadurch hoehere Sensitivitaet der Erkennung moeglich und weniger Frustration auf Seiten des Operateurs).*)

- I did not feel uncomfortable during the short period of time of the test due to the weight of the HMD. However, it would be beneficial if the displays can be swiveled out of the field of view when not needed. This would allow for a more comfortable posture. A bigger field of view would be beneficial, e.g. by using an optical see-through device. In addition, the position of the pedal is important to guarantee a good balance during its control. **HAND** restricted my field of view and I found it difficult to quickly drag objects to a more beneficial position in space. (*Fuer die kurze Dauer des Tests empfand ich das Gewicht des Displays nicht als stoerend. Allerdings sollte es eine Moeglichkeit geben, den Helm bzw. die Kameras auf dem Helm nach Bedarf nach oben oder unten zu schwenken, da es fuer diese Anwendung meiner Ansicht nach keine grosse Rolle spielt, ob die virtuelle und die tatsaechliche Sichtachse zusammenfallen. Man haette dann die Moeglichkeit, insgesamt in einer bequemeren Koerperhaltung zu agieren. Ein groesseres Sichtfeld waere sehr hilfreich, evtl. einfach dadurch, dass die Umgebung direkt sichtbar ist, z.B. mit einem durchsichtigen Helm. Es waere auch sehr guenstig, wenn man sich den Fussschalter so plazieren koennte, dass man bequem auf einem Bein stehen und den Schalter mit dem anderen Bein betaetigen kann, ohne die Balance halten zu muessen. Stoerend empfand ich, dass die Handsteuerung das Sichtfeld verdeckt, und dass es mir nicht leicht fiel, den virtuellen Regler schnell im Raum an eine andere, guenstigere Stelle zu verschieben.*)
- **HAND** would be very beneficial when becoming more robust. Then it would be comparable with **VOICE**. (*Die Verwendung der Handdektion waere sehr gut, wenn das System robuster waere. Dann waere es mit Sprache vergleichbar.*)
- Recution of image jittering, bigger field of view, improvment of **HAND**, sliderbar should also be drag- and dropable. (*Weniger Bildruckeln, groesserer Bildausschnitt, viell wuerde 16:9 natuerlicher aussehen; Handbedienung muss verbessert werden, UI muss man auch drag & droppen koennen*)
- Improvement of image quality, stereo effect, tracking, real time performance. (*Bildqualitaet (anstrengend fuer die Augen), Stereowahrnehmung nicht perfekt, zu vielen Fehlerquellen (z.B. Tracking funktioniert nicht, Lags).*)

B.2 Evaluation: Virtual Mirror

This section is related to the user study presented in section 4.2. The questionnaire of the evaluation of the virtual mirror consists of the following items (Given in German)

- I do not have problems to perceive depth or distances within an area of one meter. (*Ich habe keine Probleme Tiefe bzw. Distanzen im Bereich 1 Meter abzuschätzen.*)

- I have problems to perveice depth or distances within an area of three meters, e.g. parking a car (*Ich habe Probleme Tiefe bzw. Distanzen im Bereich 3 Meter abzuschätzen (z.B. einparken)*)
- I have problems to perceive 3D shapes within an area of one meter. (*Ich habe Probleme 3D Formen zu erkennen im Bereich 1 Meter.*)
- I do not have problems to perceive 3D shaes within an area of three meters. (*Ich habe keine Probleme 3D Formen zu erkennen im Bereich 3 Meter.*)
- I need glasses for my daily tasks (*Ich nutze eine Sehhilfe für meine Arbeit.*)
- The mirror does not provide advantages for accomplishing the task. (*Der Spiegel bringt keinen Vorteil für die Durchführung der Aufgabe*)
- I had problems to estimate the position of the spheres. (*Ich hatte Probleme die Position der Kugeln abzuschätzen.*)
- I used the mirror, if provided, to accomplish the task. (*Ich habe den Spiegel, wenn er vorhanden war, genutzt, um die Aufgabe durchzuführen.*)
- The task was easy. (*Die Aufgabe war einfach.*)
- The mirror constrained my performance. (*Der Spiegel war hinderlich für die Durchführung der Aufgabe.*)
- In my opinion, it makes no difference for the performance whether the mirror is present or not. (*Ich hatte nicht den Eindruck, dass es einen Unterschied macht ob der Spiegel vorhanden ist oder nicht.*)
- If the mirror had been at another position, it would have been more helpful. (*Wenn der Spiegel an einer anderen Position gewesen wäre, dann wäre er hilfreicher gewesen.*)
- If I had the option to control the position of the mirror, it would have been more helpful. (*Wenn ich den Spiegel selbst hätte positionieren können, dann wäre er hilfreicher gewesen.*)
- The mirror was helpful for navigating the ring to the sphere. (*Der Spiegel war hilfreich für die Navigation des Rings zur Kugel.*)
- The mirror was helpful for navigating the ring over the sphere. (*Der Spiegel war hilfreich für die Navigation des Rings über die Kugel.*)
- The mirror was helpful to avoid collisions between the ring and the sphere. (*Der Spiegel hat geholfen die Kugel nicht mit dem Ring zu berühren.*)
- I felt saver when navigating with the mirror. (*Ich hatte mit dem Spiegel ein sichereres Gefühl bei der Navigation.*)

- The mirror allowed me to control the position of objects using the additional perspective. (*Der Spiegel verhalf mir durch die zusätzliche Perspektive zu der Möglichkeit die Position der Objekte zu überprüfen.*)
- The mirror could not prevent collisions. (*Der Spiegel konnte mir nicht helfen eine Kollision mit der Kugel zu vermeiden.*)

B.3 Evaluation: AR for Port Placement

This section is related to the user study presented in section 5.3. Figure 8.1 shows a random-dot-stereogram. The image has to be divided at its vertical center into two pieces to see the hidden elephant. The left/right half of the image has to be exposed to the left/right eye of the observer, e.g. it has to be shown on the left/right display of the stereo HMD.

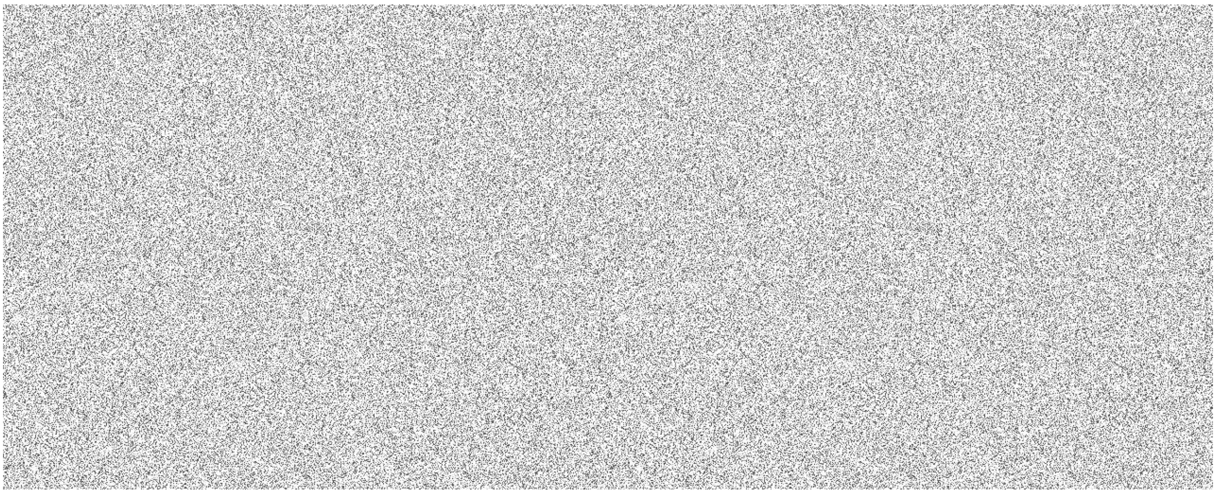
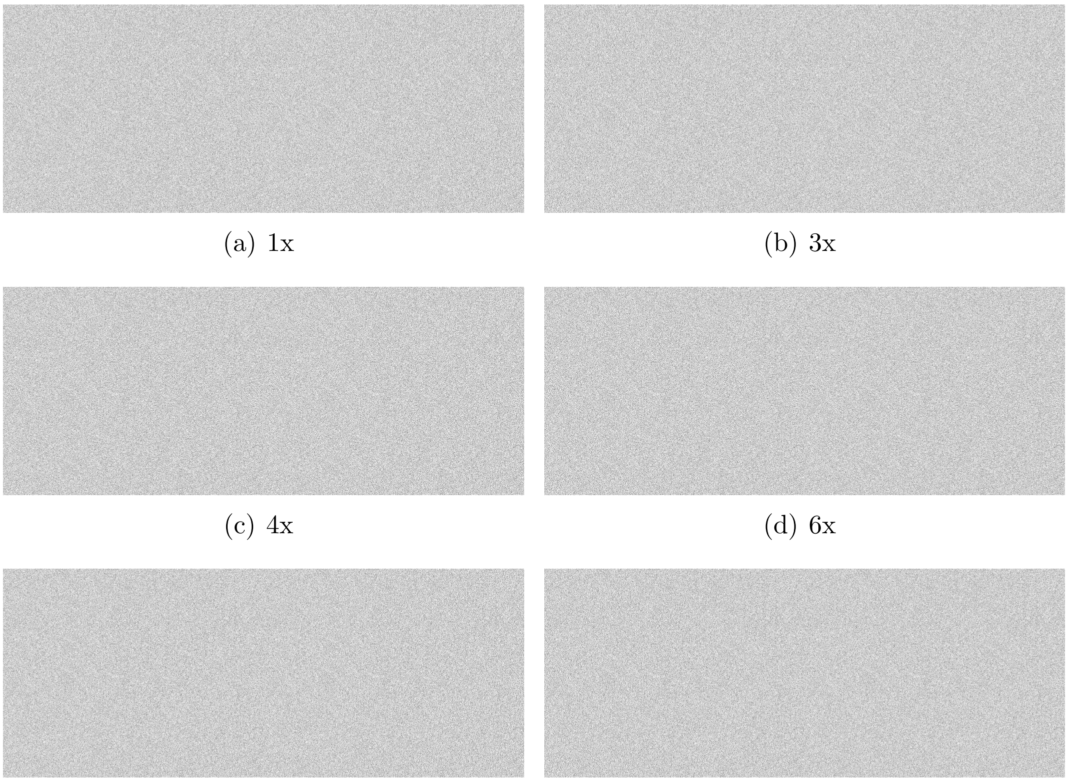


Figure 8.1: Elephant

Figures 8.2 show an advanced version of the Lang stereo test. Each of the random-dot-stereograms hides four objects while the horizontal shift of these objects increases.

The BFS test asks probands to evaluate their current mood (scale 1-5: I agree – I do not agree). The following list already categorizes criteria to describe the emotional state. The ability of verbalizing the emotional state is strongly affected by cultural and lingual background of the probands. Some of the items describe only slightly different emotions. Translation of the single items would bias understanding and interpretation. For this reason, we chose a German test since mostly German probands took part in our study. However, the main emotional categories have been translated: depression, fatigue, contemplation, calmness, elation, activity, excitement, anger. During the study, the items were presented randomized.

- Aktiviertheit (activity)
  - frisch



(a) 1x (b) 3x

(c) 4x (d) 6x

(e) 9x (f) 12x

Figure 8.2: Random-dot-stereograms with increasing pixel shifts (a-f) shows sets of four objects that can be a sphere, a square or two differently orientated triangles. The objects are randomly ordered.

  - angeregt
  - voller Energie
  - tatkraeftig
  - aktiv

- Gehobene Stimmung (elation)
  - unbeschwert
  - angenehm
  - ausgezeichnet
  - gut gelaunt
  - freudig
- Besinnlichkeit (contemplation)
  - nachdenklich

  - beschaulich
  - nach innen gekehrt
  - traeumerisch
  - besinnlich
- Ruhe (calmness)
  - locker
  - geloest
  - entspannt
  - ruhig
  - gelassen
- Ärger (anger)
  - missmutig
  - ärgerlich
  - sauer
  - gereizt
  - mürrisch
- Erregtheit (excitement)
  - ruhelos
  - nervös
  - verkrampft
  - angespannt
  - kribbelig
- Deprimiertheit (depression)
  - gedrückt
  - betrübt
  - traurig
  - niedergeschlagen
  - unglücklich
- Energielosigkeit (fatigue)
  - passiv
  - energielos

- lasch
- träge
- lahm

Further questions being part of the questionnaire after the port placements were posed as follows:

- Did you use in-situ visualization to position the ports or was in-situ visualization helpful for accomplishing the task? (*Wurde die in-situ Visualisierung genutzt, um die Ports zu setzen bzw. War die in-situ Visualisierung hilfreich fuer die Aufgabenbewaeltigung?*)
- Did you use both eyes for accomplishing the task? (*Wurde das Stereosehen genutzt bzw. wurde bei der Port Platzierung mit dem HMD die Szene immer mit beiden Augen betrachtet?*)

REFERENCES

[1] A. Abele-Brehm and W. Brehm. Zur Konzeptualisierung und Messung von Befindlichkeit. Die Entwicklung der Befindlichkeitsskalen (BFS). In *Diagnostica*, volume 32, pages 209–228, 1986.

[2] L. Adhami and È. Coste-Manière. Optimal planning for minimally invasive surgical robots. *IEEE Transactions on Robotics and Automation*, 19(5):854–863, 2003.

[3] S.-A. Ahmadi. Discovery and detection of surgical activity in percutaneous vertebroplasty. Master's thesis, Technische Universität München (TUM), 2008.

[4] S.-A. Ahmadi, N. Padoy, S. M. Heining, H. Feussner, M. Daumer, and N. Navab. Introducing wearable accelerometers in the surgery room for activity detection. In *7. Jahrestagung der Deutschen Gesellschaft für Computer-und Roboter-Assistierte Chirurgie (CURAC 2008)*, Leipzig, Germany, September 2008.

[5] A. Alapetite, H. Boje Andersen, and M. Hertzum. Acceptance of speech recognition by physicians: A survey of expectations, experiences, and social influence. *Int. J. Hum.-Comput. Stud.*, 67(1):36–49, 2009.

[6] R. Allen and M. Rubin. Chromostereopsis. *Surv Ophthalmol*, 26(1):22–7, 1981.

[7] J. Aloimonos. Shape from texture. *Biol. Cybern.*, 58(5):345–360, 1988.

[8] M. Arand, M. Schempf, D. Hebold, S. Teller, L. Kinzl, and F. Gebhard. Precision of navigation-assisted surgery of the thoracic and lumbar spine. *Unfallchirurg*, 106(11):899–906, 2003.

[9] Y. Argotti, L. Davis, V. Outters, and J. Rolland. Dynamic superimposition of synthetic objects on rigid and simple-deformable real objects. volume 26, pages 919–930. Elsevier, 2002.

[10] C. Arth, D. Wagner, M. Klopschitz, A. Irschara, and D. Schmalstieg. Wide area localization on mobile phones. pages 73–82, 2009.

[11] B. Avery, C. Sandor, and B. Thomas. Improving spatial perception for augmented reality x-ray vision. In *VR '09: Proceedings of the IEEE Conference on Virtual Reality*, pages 79–82, Louisiana, USA, 2009. IEEE Computer Society.

[12] R. Azuma. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.

[13] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001.

[14] Y. Baillot, J. Rolland, K. Lin, and D. Wright. Automatic modeling of knee-joint motion for the virtual reality dynamic anatomy (vrda) tool. *Presence: Teleoperators and Virtual Environments*, 9(3):223–235, 2000.

[15] M. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 203–210. ACM Press, 1992.

[16] A. W. Bates. Anatomical Venuses: the aesthetics of anatomical modelling in 18th- and 19th-century Europe. In *DAGM*, volume 1, 2006.

[17] B. Bauer, P. Jolicoeur, and W. Cowan. Distractor heterogeneity versus linear separability in visual search. *Perception*, 25:1281–1294, 1996.

[18] R. Behrens. The man who made distorted rooms: A chronology of the life of adelbert ames, jr. 1993.

[19] C. Bichlmeier. Diploma thesis: Advanced 3d visualization for intra operative augmented reality. Master's thesis, Technische Universitaet Muenchen, Department of Computer Science, Chair for Computer Aided Medical Procedures and Augmented Reality, 2006. 1.

[20] C. Bichlmeier, S. M. Heining, M. Feuerstein, and N. Navab. The virtual mirror: A new interaction paradigm for augmented reality environments. *IEEE Trans. Med. Imag.*, 28(9):1498 – 1510, Sept 2009.

[21] C. Bichlmeier, S. M. Heining, M. Rustaee, and N. Navab. Virtually extended surgical drilling device: Virtual mirror for navigated spine surgery. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2007, 10th International Conference*, Lecture Notes in Computer Science, Brisbane, Australia, October/November 2007. Springer Verlag.

[22] C. Bichlmeier and N. Navab. Virtual window for improved depth perception in medical ar. In *AMIARCS - Virtual Window for Improved Depth Perception in Medical AR*, Copenhagen, Denmark, Oct. 2006. MICCAI Society.

[23] C. Bichlmeier, B. Ockert, S. M. Heining, A. Ahmadi, and N. Navab. Stepping into the Operating Theater: ARAV - Augmented Reality Aided Vertebroplasty. In *Proceedings of the 7th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 165–166, 2008.

[24] C. Bichlmeier, B. Ockert, O. Kutter, M. Rustaee, S. M. Heining, and N. Navab. The visible korean human phantom: Realistic test & development environments for medical augmented reality. In *International Workshop on Augmented environments for Medical Imaging including Augmented Reality in Computer-aided Surgery (AMI-ARCS 2008)*, New York, USA, Sept. 2008. MICCAI Society.

[25] C. Bichlmeier, H. Sandro Michael, R. Mohammad, and N. Nassir. Laparoscopic Virtual Mirror for Understanding Vessel Structure: Evaluation Study by Twelve Surgeons. In *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 125–128, Nara, Japan, Nov. 2007.

[26] C. Bichlmeier, T. Sielhorst, and N. Navab. The tangible virtual mirror: New visualization paradigm for navigated surgery. In *AMIARCS - The Tangible Virtual Mirror: New Visualization Paradigm for Navigated Surgery*, Copenhagen, Denmark, Oct 2006. MICCAI Society.

[27] C. Bichlmeier, F. Wimmer, H. Sandro Michael, and N. Nassir. Contextual Anatomic Mimesis: Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality. In *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 129–138, Nov. 2007.

[28] G. Bingham. Form as information about scale: Perceiving the size of trees. *Journal of Experimental Psychology: Human Perception and Performance*, 19:1139–1161, 1993.

[29] W. Birkfellner, M. Figl, K. Huber, F. Watzinger, F. Wanschitz, J. Hummel, R. Hanel, W. Greimel, P. Homolka, R. Ewers, and H. Bergmann. A head-mounted operating binocular for augmented reality visualization in medicine - design and initial evaluation. *IEEE Trans. Med. Imag.*, 21(8):991–997, 2002.

[30] W. Birkfellner, M. Figl, C. Matula, J. Hummel, H. I. R Hanel, F. Wanschitz, A. Wagner, F. Watzinger, and H. Bergmann. Computer-enhanced stereoscopic vision in a head-mounted operating binocular. *Physics in Medicine and Biology*, 48(3):N49–N57, 2003.

[31] W. Birkfellner, K. Huber, F. Watzinger, M. Figl, F. Wanschitz, R. Hanel, D. Rafolt, R. Ewers, and H. Bergmann. Development of the varioscope ar – a see-through hmd for computer-aided surgery. In *Proc. IEEE and ACM Int'l Symp. on Augmented Reality*, pages 54–59. IEEE Computer Society Press, 2000.

[32] M. Blackwell, C. Nikou, A. M. D. Gioia, and T. Kanade. An image overlay system for medical data visualization. In *Proceedings of the First International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1496 of *Lecture Notes in Computer Science*, pages 232–240, Cambridge, MA, USA, October 1998. Springer-Verlag.

[33] T. Blum, S. M. Heining, O. Kutter, and N. Navab. Advanced training methods using an augmented reality ultrasound simulator. In *8th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2009)*, pages 177–178, Orlando, USA, October 2009.

[34] J. Bluteau, I. Kitahara, Y. Kameda, H. Noma, K. Kogure, and Y. Ohta. Visual support for medical communication by using projector-based augmented reality and thermal markers. In *ICAT '05: Proceedings of the 2005 international conference on Augmented tele-existence*, volume 157, pages 98–105, New York, NY, USA, 2005. ACM.

[35] U. Bockholt, A. Bisler, M. Becker, W. Mueller-Wittig, and G. Voss. Augmented reality for enhancement of endoscopic interventions. In *Proc. IEEE Virtual Reality (VR)*, page 97, 2003.

[36] G. R. Bradski. Computer vision face tracking for use in a perceptual user interface. *Intel Technology Journal*, 2nd Quarter 1998.

[37] M. Brinks. *Aggression gegen Computer. Eine wissenschaftliche Untersuchung eines alltäglichen Phänomens.* Stuttgart: ibidem-Verlag, 2005.

[38] M. J. Brooks. *Shape from shading.* MIT Press Series Of Artificial Intelligence Series, Cambridge, MA, USA, 1989.

[39] S. Bruckner, S. Grimm, A. Kanitsar, and M. Groller. Illustrative context-preserving exploration of volume data. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1559–1569, Nov 2006.

[40] S. Bruckner and M. Gröller. Enhancing Depth-Perception with Flexible Volumetric Halos. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1344–1351, Nov 2007.

[41] E. v. Bruecke. Über asymmetrische strahlenbrechung im menschlichen auge. *Auge. Sber. K. bohm. Gen. Wiss.*, 58:321–328, 1868.

[42] M. R. Burmeister. Popular anatomical museums in nineteenth-century England. 2000.

[43] M. Burns, M. Haidacher, W. Wein, I. Viola, and E. Groeller. Feature emphasis and contextual cutaways for multimodal medical visualization. In *Eurographics / IEEE VGTC Symposium on Visualization*, pages 275–282, May 2007.

[44] O. Cakmakci and J. Rolland. Head-worn displays: A review. *Journal of Display Technology*, 2(3):199–216, September 2006.

[45] J. Cannon, J. Stoll, S. Selha, P. Dupont, R. Howe, and D. Torchiana. Port placement planning in robot-assisted coronary artery bypass. *IEEE Transactions on Robotics and Automation*, 19(5):912–917, 2003.

[46] A. M. Chiu, D. Dey, M. Drangova, W. D. Boyd, and T. M. Peters. 3-d image guidance for minimally invasive robotic coronary artery bypass. *The Heart Surgery Forum*, 3(3):224–231, 2000.

[47] A. Chu, W.-Y. Chan, J. Guo, W.-M. Pang, and P.-A. Heng. Perception-aware depth cueing for illustrative vascular visualization. In *BMEI '08: Proceedings of the 2008 International Conference on BioMedical Engineering and Informatics*, volume 1, pages 341–346, Washington, DC, USA, 2008. IEEE Computer Society.

[48] K. Cleary and C. Nguyen. State of the art. in surgical robotics: clinical applications and technology challenges. *Comput. Aided Surg.*, 6(6):312–328, 2001.

[49] J. E. Cutting and P. M. Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *W. Epstein & S. Rogers (Eds.), Perception of Space and Motion*, pages 69–117, 1995.

[50] T. Darrell, G. Gordon, J. Woodfill, and M. Harville. A virtual mirror interface using real-time robust face tracking. In *Proc. Third IEEE International Conference on Automatic Face and Gesture Recognition*, page 616, Washington, DC, USA, 1998. IEEE Comput. Soc.

[51] M. Das, F. Sauer, U. Schoepf, A. Khamene, S. Vogt, S. Schaller, R. Kikinis, E. vanSonnenberg, and S. Silverman. Augmented reality visualization for ct-guided interventions: System description, feasibility, and initial evaluation in an abdominal phantom. *Radiology*, 240(1):230, 2006.

[52] L. Dieker, M. Hynes, C. B. Stapleton, and C. E. Hughes. Virtual classrooms: Star simulator, teacher preparation in effective classroom management. In *Symposium on Applied Learning Technology*, Jan 2007.

[53] J. Diepstraten, D. Weiskopf, and T. Ertl. Transparency in interactive technical illustrations. In *Computer Graphics Forum*, volume 21, pages 317–325, Sept 2002.

[54] D. H. Douglas and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 10(2):112–122, Dec. 1973.

[55] M. V. During, M. Poggesi, and G. Didi-Huberman. *Encyclopedia Anatomica: Museo La Specola Florence*. Taschen, 2004.

[56] P. J. Edwards, L. G. Johnson, D. J. Hawkes, M. R. Fenlon, A. J. Strong, and M. J. Gleeson. Clinical experience and perception in stereo augmented reality surgical navigation. In *Proceesings of Medical Imaging and Augmented Reality: Second International Workshop, MIAR 2004, Beijing, China, August 19-20, 2004.*, pages 369–376, 2004.

[57] F. W. E.H.Roth, C.Hoffmann. Stereo - Test Düsseldorf. Ein neuer Test zur quantitativen Messung der Stereopsis. In *Z. prakt. Augenheilkd.*, volume 19, pages 273–278, 1998.

[58] V. p. Einthoven, W. Stereoscopie durch Farbendifferenz. *Albrecht von Graefes Archiv fur Opthamologie*, 31:211–238, 1885.

[59] K. Engel, M. Hadwiger, J. M. Kniss, A. E. Lefohn, C. R. Salama, and D. Weiskopf. Real-time volume graphics. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Course Notes*, page 29, New York, NY, USA, 2004. ACM Press.

[60] W. Epstein. The influence of assumed size on apperent distance. *American Journal of Psychology*, 76:257–265, 1963.

[61] W. Epstein. Nonrelational judgements of size and distances. *American Journal of Psychology*, 78:120–123, 1965.

[62] M. C. Escher. *The graphic work of M.C.Escher*. Oldbourne Press, Lond. :, 1961.

[63] C. K. Eveland, K. G. Konolige, and R. C. Bolles. Background modeling for segmentation of video-rate stereo sequences. In *CVPR*, pages 266–271, 1998.

[64] M. Feuerstein. *Augmented Reality in Laparoscopic Surgery – New Concepts for Intraoperative Multimodal Imaging*. PhD thesis, Technische Universität München, 2007.

[65] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab. Registration-free laparoscope augmentation for intra-operative liver resection planning. In *SPIE Medical Imaging*, volume 6509, San Diego, California, USA, February 2007.

[66] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab. Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection. *IEEE Trans. Med. Imag.*, 27(3):355–369, March 2008.

[67] M. Fiala. Magic mirror system with hand-held and wearable augmentations. In *IEEE Virtual Reality Conference*, pages 251–254. IEEE Computer Society, 2007.

[68] G. Fichtinger, A. Deguet, K. Masamune, E. Balogh, G. Fischer, H. Mathieu, R. H. Taylor, L. M. Fayad, and S. J. Zinreich. Needle insertion in ct scanner with image overlay - cadaver studies. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 3217 of *Lecture Notes in Computer Science*, pages 795–803. Springer-Verlag, 2004.

[69] G. Fichtinger, A. Deguet, K. Masamune, E. Balogh, G. S. Fischer, H. Mathieu, R. H. Taylor, S. J. Zinreich, and L. M. Fayad. Image overlay guidance for needle insertion in ct scanner. *IEEE Transactions on Biomedical Engineering*, 52(8):1415–1424, 2005.

[70] M. Figl, W. Birkfellner, J. Hummel, R. Hanel, P. Homolka, F. Watzinger, F. Wanschitz, R. Ewers, and H. Bergmann. Current status of the varioscope ar, a head-mounted operating microscope for computer-aided surgery. *IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, page 20, 2001.

[71] M. Figl, W. Birkfellner, F. Watzinger, F. Wanschitz, J. Hummel, R. Hanel, R. Ewers, and H. Bergmann. Pc-based control unit for a head mounted operating microscope for augmented reality visualization in surgical navigation. In *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 2489 of *Lecture Notes in Computer Science*, pages 44–51. Springer-Verlag, 2002.

[72] M. Figl, C. Ede, J. Hummel, F. Wanschitz, R. Ewers, H. Bergmann, and W. Birkfellner. A fully automated calibration method for an optical see-through head-mounted operating microscope with variable zoom and focus. *IEEE Trans. Med. Imag.*, 24(11):1492–1499, 2005.

[73] G. S. Fischer, A. Deguet, D. Schlattman, R. Taylor, L. Fayad, S. J. Zinreich, and G. Fichtinger. Mri image overlay: Applications to arthrography needle insertion. In *Medicine Meets Virtual Reality (MMVR) 14*, 2006.

[74] J. Fischer, D. Bartz, and W. Straßer. Occlusion Handling for Medical Augmented Reality using a Volumetric Phantom Model. In *Proc. of the ACM symposium on Virtual reality software and technology - VRST '04*, pages 174–177. ACM Press, November 2004.

[75] A. François, E. Kang, and U. Malesci. A handheld virtual mirror. In *Proc. ACM SIGGRAPH 2002 conference abstracts and applications on - SIGGRAPH '02*, page 140. ACM Press, 2002.

[76] H. Freeman. Computer processing of line drawing images. *ACM Computing Surveys*, 6(1):57–97, Mar. 1974.

[77] W. Freysinger, M. J. Truppe, A. R. Gunkel, and W. F. Thumfart. Stereotaktische telepräsenz in der hals-, nasen- und ohrenchirurgie. *HNO*, 50:424–432, May 2002.

[78] E. M. Friets, J. W. Strohbehn, J. F. Hatch, and D. W. Roberts. A frameless stereotaxic operating microscope for neurosurgery. *IEEE Transactions on Biomedical Engineering*, 36(6), 1989.

[79] H. Fuchs, S. A., E. Pisano, W. Garrett, G. Hirota, M. Livingston, M. Whitton, and S. Pizer. Towards performing ultrasound-guided needle biopsies from within a head-mounted display. pages 591–600, 1996.

[80] H. Fuchs, M. A. Livingston, R. Raskar, D. Colucci, K. Keller, A. State, J. R. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer. Augmented reality visualization for laparoscopic surgery. In *Proceedings of the First International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1496 of

Lecture Notes in Computer Science, pages 934–943, Cambridge, MA, USA, October 1998. Springer-Verlag.

[81] W. Fuchs. Experimentelle Untersuchungen über das simultane Hintereinandersehen auf derselben Sehrichtung. *Zeitschrift für Psychologie*, 91:145–235, 1923.

[82] C. Furmanski, R. Azuma, and M. Daily. Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information. In *Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR)*, page 215, Washington, DC, USA, 2002. IEEE Computer Society.

[83] F. Galasso and J. Lasenby. *Advances in Visual Computing*, volume 5358, chapter Shape from Texture Via Fourier Analysis, pages 803–814. Springer Berlin / Heidelberg, 2008.

[84] S. C. I. M. J. Gerbino, W. and C. M. de Weert. Transparent layer constancy. *Journal of Experimental Psychology: Human Perception and Performance.*, 16(1):3–20, 1990.

[85] J. Gibson. *The perception of the visual world.* Greenwood Pub Group, 1950.

[86] J. J. Gibson. *The senses considered as perceptual systems.* Greenwood Pub Group Inc, 1966.

[87] S. Grange, T. Fong, and C. Baur. M/ORIS: A medical/operating room interaction system. In *ACM International Conference on Multimodal Interfaces*, State College, PA, Oct. 2004.

[88] W. E. L. Grimson, G. J. Ettinger, T. Kapur, M. E. Leventon, W. M. Wells, and R. Kikinis. Utilizing segmented MRI data in image-guided surgery. *International Journal of Pattern Recognition and Artificial Intelligence*, 1996.

[89] W. E. L. Grimson, T. Lozano-Perez, W. M. Wells, III, G. J. Ettinger, S. J. White, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. *IEEE Trans. Med. Imag.*, 15(2):129–140, 1996.

[90] J. Grosjean and S. Coquillart. The magic mirror: A metaphor for assisting the exploration of virtual worlds. *Proc. of Spring Conference on Computer Graphics (Budmerice, Slovakia, April 1999)*, pages 125–129, 1999.

[91] J. Hajnal, D. Hawkes, and D. Hill. *Medical Image Registration.* CRC Press, 2001.

[92] M. Haller, S. Drab, and W. Hartmann. A real-time shadow approach for an augmented reality application using shadow volumes. In *VRST '03: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 56–65, New York, NY, USA, 2003. ACM.

[93] R. A. Hart, B. L. Hansen, M. Shea, F. Hsu, and G. J. Anderson. Pedicle screw placement in the thoracic spine: a comparison of image-guided and manual techniques in cadavers. *Spine*, 30(12):326–31, 2005.

[94] J.-M. Hasenfratz, M. Lapierre, N. Holzschuch, and F. Sillion. A survey of real-time soft shadows algorithms. *Computer Graphics Forum*, 22(4):753–774, 2003. State-of-the-Art Report.

[95] J. Hatch and J. S. D.W.Roberts. Reference-display system for the integration of ct scanning and the operating microscope. In *Proceedings of the Eleventh Annual Northeast Bioengineering Conference*, 1985.

[96] H. Hedegaard, M. R. Dahl, and K. Grønbæk. Ekgar: Augmented reality support for learning of electrocardiography (ecg) analysis. In C. Montgomerie and J. Seale, editors, *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2007*, pages 3544–3549, Vancouver, Canada, June 2007. AACE.

[97] S.-M. Heining, P. Stefan, L. Omary, S. Wiesner, T. Sielhorst, N. Navab, F. Sauer, E. Euler, W. Mutschler, and J. Traub. Evaluation of an in-situ visualization system for navigated trauma surgery. *Journal of Biomechanics*, 39(Supplement 1):209, 2006.

[98] E. Hering. *Grundzüge der Lehre vom Lichtsinn.* Springer (Berlin), 1820.

[99] W. A. Hoff and T. L. Vincent. Analysis of head pose accuracy in augmented reality. *IEEE Trans. Visualization and Computer Graphics*, 6(4):319 – 334, 2000.

[100] K. Homma and E.-I. Takenaka. An image processing method for feature extraction of space-occupying lesions. *Journal of Nuclear Medicine*, 26:1472–1477, 1985.

[101] C. E. Hughes, C. B. Stapleton, D. E. Hughes, and E. M. Smith. Mixed reality in education, entertainment, and training. *IEEE Comput. Graph. Appl.*, 25(6):24–30, 2005.

[102] V. Interrante, H. Fuchs, and S. Pizer. Enhancing transparent skin surfaces with ridge and valley lines. In *VIS '95: Proceedings of the 6th conference on Visualization '95*, page 52, Washington, DC, USA, 1995. IEEE Computer Society.

[103] A. Ionescu. A mouse in the O.R. In *Ambidextrous Magazine*. Stanford University Journal of Design, 2006.

[104] J. Itten. *The art of color : the subjective experience and objective rationale of color / Johannes Itten ; translated by Ernst van Haagen.* John Wiley & Sons, 1973.

[105] K. Jacobs, J.-D. Nahmias, C. Angus, A. Reche, C. Loscos, and A. Steed. Automatic generation of consistent shadows for augmented reality. In *GI '05: Proceedings of Graphics Interface 2005*, volume 112, pages 113–120, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005. Canadian Human-Computer Communications Society.

[106] N. John and I. Lim. Cybermedicine tools for communication and learning. In *Proc. of Theory and Practice of Computer Graphics*, pages 165–170, 2007.

[107] L. Johnson, P. Edwards, and D. Hawkes. Surface transparency makes stereo overlays unpredictable: the implications for augmented reality. In *Stud Health Technol Inform. 94*, pages 131–136, Department of Surgery, Duke University Medical Center, Durham, NC 27710, USA, 2003.

[108] L. G. Johnson, P. Edwards, and D. Hawkes. Surface transparency makes stereo overlays unpredictable: The implications for augmented reality. In J. D. Westwood, editor, *Medicine Meets Virtual Reality (MMVR)*, volume 94 of *Studies in Health Technology and Informatics*, pages 131–136. IOS Press, 2002.

[109] C. Juan, F. Beatrice, and J. Cano. An augmented reality system for learning the interior of the human body. In *Proc. of Eighth IEEE International Conference on Advanced Learning Technologies*, 2008.

[110] D. Kalkofen, E. Mendez, and D. Schmalstieg. Interactive Focus and Context Visualization for Augmented Reality. In *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 191–200, Nov. 2007.

[111] D. Kalkofen, E. Mendez, and D. Schmalstieg. Comprehensible visualization for augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 15(2):193–204, 2009.

[112] A. R. Kancherla, J. P. Rolland, D. L. Wright, and G. Burdea. A novel virtual reality tool for teaching dynamic 3d anatomy. In *CVRMed '95: Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, pages 163–169, London, UK, 1995. Springer-Verlag.

[113] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *IWAR '99: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, page 85, Washington, DC, USA, 1999. IEEE Computer Society.

[114] J. Kelso, K. Holt, and A. Flatt. The role of proprioception in the perception and control of human movement: Toward a theoretical reassessment. *Perception & Psychophysics*, 28(1):45–52, July 1980.

[115] S. Kerschbaumer. Diploma thesis: Design, development and evaluation of a multimodal user interface for medical in-situ visualization. Master's thesis, Technische Universitaet Muenchen, Department of Computer Science, Chair for Computer Aided Medical Procedures and Augmented Reality, 2009. 1.

[116] D. Kersten, P. Mamassian, and D. Knill. Moving cast shadows and the perception of relative depth. Technical Report 6, Max-Planck-Institut fuer biologische Kybernetik, 1994.

[117] A. Khamene, S. Vogt, F. Azar, T. Sielhorst, and F. Sauer. Local 3D reconstruction and augmented reality visualization of freehand ultrasound for needle biopsy procedures. In R. E. Ellis and T. M. Peters, editors, *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Lecture Notes in Computer Science, pages 344–355. Springer-Verlag, 2003.

[118] A. King, P. Edwards, C. Maurer, Jr., D. de Cunha, R. Gaston, M. Clarkson, D. Hill, D. Hawkes, M. Fenlon, A. Strong, T. Cox, and M. Gleeson. Stereo augmented reality in the surgical microscope. *Presence: Teleoperators and Virtual Env.*, 9(4):360–368, 2000.

[119] A. P. King, P. J. Edwards, C. R. Maurer, Jr., D. A. de Cunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. J. Strong, C. L. Chandler, A. Richards, and M. J. Gleeson. A system for microscope-assisted guided interventions. *Stereotact Funct Neurosurg*, (72):107–111, 1999.

[120] A. P. King, P. J. Edwards, C. R. Maurer, Jr., D. A. de Cunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. J. Strong, C. L. Chandler, A. Richards, and M. J. Gleeson. Design and evaluation of a system for microscope-assisted guided interventions. *IEEE Trans. Med. Imag.*, 19(11):1082–1093, 2000.

[121] M. Kipot. Master thesis: Improving depth perception for video-based, intraoperative augmented reality systems. Master's thesis, Technische Universitaet Muenchen, Department of Computer Science, Chair for Computer Aided Medical Procedures and Augmented Reality, 2009. 1.

[122] M. Kölsch. *Vision Based Hand Gesture Interfaces for Wearable Computing and Virtual Environments*. PhD thesis, University of California Santa Barbara, 2004.

[123] K. Koffka. *Principles of Gestalt psychology*. Routledge, London, 1935.

[124] I. Kohler. Experiments with goggles. *Scientific American*, 206:63 – 72., 1962.

[125] K. Kojima, A. Sato, F. Taya, Y. Kameda, and Y. Ohta. Naviview: visual assistance by virtual mirrors at blind intersection. In *Proceedings of IEEE Intelligent Transportation Systems*, pages 592– 597, 2005.

[126] J. Krueger, J. Schneider, and R. Westermann. Clearview: An interactive context preserving hotspot visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):941–948, 2006.

[127] J. Krueger and R. Westermann. Acceleration techniques for gpu-based volume rendering. In *VIS '03: Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 38, Washington, DC, USA, 2003. IEEE Computer Society.

[128] O. Kutter, A. Aichert, C. Bichlmeier, S. M. Heining, B. Ockert, E. Euler, and N. Navab. Real-time Volume Rendering for High Quality Visualization in Augmented Reality. In *International Workshop on Augmented environments for Medical Imaging*

including Augmented Reality in Computer-aided Surgery (AMI-ARCS 2008), New York, USA, Sept. 2008. MICCAI Society.

[129] F.-J. Lenz and A. K. Nischelwitzer. My inner body - book, interactive three-dimensional information brokering with augmented reality. In *2nd Usability Symposium, TechGate Vienna*, 2006.

[130] V. Lepetit, P. Lagger, and P. Fua. Randomized trees for real-time keypoint recognition. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 775 – 781, June 2005.

[131] M. Lerotic, A. J. Chung, G. Mylonas, and G.-Z. Yang. pq -space based non-photorealistic rendering for augmented reality. In N. Ayache, S. Ourselin, and A. Maeder, editors, *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4792 of *LNCS*, pages 102–109, Brisbane, Australia, October 2007. MICCAI Society, Springer-Verlag Berlin Heidelberg.

[132] M. Levoy, H. Fuchs, S. Pizer, J. Rosenman, E. Chaney, G. Sherouse, V. Interrante, and J. Kiel. Volume rendering in radiation treatment planning. In *Proceedings of the First Conference on Visualization in Biomedical Computing*, pages 4–10, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.

[133] H. Liao, N. Hata, S. Nakajima, M. Iwahara, I. Sakuma, and T. Dohi. Surgical navigation by autostereoscopic image overlay of integral videography. *Information Technology in Biomedicine, IEEE Transactions on*, 8(2):114–121, 2004.

[134] W. Lorensen, H. Cline, C. Nafis, R. Kikinis, D. Altobelli, L. Gleason, G. Co, and N. Schenectady. Enhancing reality in the operating room. In *IEEE Conference on Visualization*, pages 410–415, 1993.

[135] K. Low, A. Ilie, G. Welch, and A. Lastra. Combining head-mounted and projector-based displays for surgical training. *Virtual Reality, 2003. Proceedings. IEEE*, pages 110–117, 2003.

[136] C. Ludwig and F. Mortimer. *History and bibliography of anatomic illustration.* Hafner Pub. Co. New York, 1962.

[137] A. Lund. Measuring usability with the USE questionnaire. *Usability Interface*, 8(2), October 2001.

[138] L. W. MacDonald. Tutorial: Using color effectively in computer graphics. *IEEE Comput. Graph. Appl.*, 19(4):20–35, 1999.

[139] P. Mamassian, D. Knill, and D. Kersten. The perception of cast shadows. *Trends in Cognitive Sciences*, 2(8):288–295, 1998.

[140] J. Marmurek, C. Wedlake, U. Pardasani, R. Eagleson, and T. Peters. Image-guided laser projection for port placement in minimally invasive surgery. In *Medicine Meets Virtual Reality 14: Accelerating Change in Healthcare: Next Medical Toolkit*, volume 119 of *Studies in Health Technology and Informatics*, pages 367–372, 2005.

[141] Y. Masutani, M. Iwahara, O. Samuta, Y. Nishi, N. Suzuki, M. Suzuki, T. Dohi, H. Iseki, and K. Takakura. Development of integral photography-based enhanced reality visualization system for surgical support. *Proc. of ISCAS*, 95:16–17, 1995.

[142] U. Matern, K. Kehl, C. Giebmeyer, and M. Faist. Monitor-position in laparoscopic surgery. *Source: Human Factors and Ergonomics Society Annual Meeting Proceedings, Medical Systems and Rehabilitation*, 5:1506–1510, 2005.

[143] F. Mattatall. A very real art. *CMAJ*, 164(7):1027–1028, 2001.

[144] C. R. Maurer, F. Sauer, B. Hu, B. Bascle, B. Geiger, F. Wenzel, F. Recchi, T. Rohlfing, C. Brown, R. Bakos, R. Maciunas, and A. Bani-Hashemi. Augmented-reality visualization of brain structures with stereo and kinetic depth cues: system description and initial evaluation with head phantom. In *Proc. SPIE Vol. 4319, p. 445-456, Medical Imaging 2001: Visualization, Display, and Image-Guided Procedures, Seong K. Mun; Ed.*, pages 445–456, May 2001.

[145] M. Meissner, J. Huang, D. Bartz, K. Mueller, and R. Crawfis. A practical evaluation of popular volume rendering algorithms. In *VVS '00: Proceedings of the 2000 IEEE symposium on Volume visualization*, pages 81–90, New York, NY, USA, 2000. ACM Press.

[146] J. Mellor. Enhanced reality visualization in a surgical environment. Technical Report 1544, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, January 1995.

[147] E. Méndez, D. Kalkofen, and D. Schmalstieg. Interactive context-driven visualization tools for augmented reality. In *Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR)*, pages 209–218, 2006.

[148] F. Metelli. The perception of transparency. *Scientific American*, 230(4):90–98, April 1974.

[149] N. Michalke. Trying on clothes in a magic mirror. *Fraunhofer Magazin*, 1:28, 2009.

[150] P. Milgram and H. Colquhoun. A framework for relating head-mounted displays to mixed reality. In *Proceedings of the human factors and ergonomics society*, pages 123–134, 1999.

[151] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. In *IEICE Trans. Information Systems*, pages 1321–1329, 1994.

[152] M. Mine, F. B. Jr., and C. Sequin. Moving objects in space: Exploiting virtual proprioception in virtual-environment interaction. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 19–26, 1997.

[153] P. Moon and D. E. Spencer. Geometrical formulation of classical color harmony. *Journal of the Optical Society of America*, 34:46–60, 1944.

[154] N. Navab, C. Bichlmeier, and T. Sielhorst. Application no. pct/ep2007/003206: Virtual penetrating mirror device for visualizing of virtual objects within an augmented reality environment, April 2006.

[155] N. Navab, M. Feuerstein, and C. Bichlmeier. Laparoscopic virtual mirror - new interaction paradigm for monitor based augmented reality. In *Virtual Reality*, pages 43–50, Charlotte, North Carolina, USA, March 2007.

[156] N. Navab, J. Traub, T. Sielhorst, M. Feuerstein, and C. Bichlmeier. Action- and workflow-driven augmented reality for computer-aided medical procedures. *IEEE Computer Graphics and Applications*, 27(5):10–14, September/October 2007.

[157] M. Nicolaou, A. James, B. Lo, A. Darzi, and Y. Guang-Zhong. Invisible shadow for navigation and planning in minimal invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 3750, pages 25–32, 2005.

[158] A. Nishikawa et al. Face mouse: A novel human-machine interface for controlling the position of a laparoscope. *IEEE Transactions on Robotics and Automation*, 19(5):825–841, 2003.

[159] D. Norman. Emotion and design: attractive things work better. *Interacion Magazine*, 9:36–42, 2002.

[160] A. O'Connor, A. Rostom, V. Fiset, J. Tetroe, V. Entwistle, H. Llewellyn-Thomas, M. Holmes-Rovner, M. Barry, and J. Jones. Decision aids for patients facing health treatment or screening decisions systematic review. *BMJ*, 319:731–734, 1999.

[161] K. M. P. Thompson and R. Stone. Chromostereopsis: A multicomponent depth effect? *Displays*, 14:227–234, 1993.

[162] S. Pardhy, C. Shankwitz, and M. Donath. A virtual mirror for assisting drivers. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 255–260, 2000.

[163] A. Park, G. Lee, F. J. Seagull, N. Meenaghan, and D. Dexter. Patients benefit while surgeons suffer: An impending epidemic. *Journal of the American College of Surgeons*, In Press, Corrected Proof:–, 2009.

[164] J. Park, M. Chung, S. Hwang, Y. Lee, D. Har, and H. Park. Visible korean human: Improved serially sectioned images of the entire body. *IEEE Trans. Med. Imaging*, 24(3):352–360, March 2005.

[165] P. Paul, O. Fleig, and P. Jannin. Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery: Methods and performance evaluation. *IEEE Trans. Med. Imag.*, 24(11):1500–1511, 2005.

[166] P. F. Peter Eisert and J. Rurainsky. Virtual mirror: Real-time tracking of shoes in augmented reality environments. In *Proc. of the International Conference on Image Processing ICIP 2007*, San Antonio, Texas, USA, September 2007.

[167] T. M. Peters. Image-guided surgery: From x-rays to virtual reality. *Computer Methods in Biomechanics and Biomedical Engineering*, 2000.

[168] T. M. Peters. Image-guidance for surgical procedures. *Phys. Med. Biol.*, 51:R505–R540, 2006.

[169] B. Peuchot, A. Tanguy, and M. Eude. Virtual reality as an operative tool during scoliosis surgery. In *Proceedings of Computer Vision, Virtual Reality and Robotics in Medicine, First International Conference, CVRMed'95*, pages 549–554, 1995.

[170] Z. P.G. and G. R.J. *Psychology and Life (17th Edition).* Allyn and Bacon, 2004.

[171] B. T. Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975.

[172] B. Pinna and M. Tanca. Perceptual organization reconsidered in the light of the watercolor illusion: The problem of perception of holes and the object-hole effect. *Journal of Vision*, 8(7):1–15, 5 2008.

[173] A. Pommert and K. H. Hoehne. Validation of medical volume visualization: a literature review. In *Computer Assisted Radiology and Surgery (CARS 2003)*, volume 1256, pages 571–576, 2003.

[174] U. Ramer. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1:244–256, 1972.

[175] W. T. Reeves, D. H. Salesin, and R. L. Cook. Rendering antialiased shadows with depth maps. *SIGGRAPH Comput. Graph.*, 21(4):283–291, 1987.

[176] B. Reichert. *Anatomie in vivo 1: Palpieren und verstehen im Bereich der Extremitäten.* Thieme, Stuttgart, October 2005.

[177] B. Reichert. *Anatomie in Vivo 2: Palpieren und verstehen im Bereich Rumpf und Kopf.* Thieme, Stuttgart, August 2007.

[178] B. A. Rifkin, M. J. Ackerman, and J. Folkenberg. *Human anatomy : depicting the body from the Renaissance to today.* London : Thames & Hudson, 2006.

[179] F. Ritter, H. Sonnet, K. Hartmann, and T. Strothotte. Illustrative shadows: integrating 3d and 2d information displays. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 166–173, New York, NY, USA, 2003. ACM.

[180] D. Roberts, J. Strohbehn, J. Hatch, W. Murray, and H. Kettenberger. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *J Neurosurg*, 65(4):545–9, 1986.

[181] J. Rolland, D. Wright, , and A. Kancherla. Towards a novel augmented-reality tool to visualize dynamic 3-d anatomy. In *Proceedings of Medicine Meets Virtual Reality*, pages 337–348, San Diego, California, 1997.

[182] J. P. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence*, 9:287–309, 2000.

[183] J. P. Rolland, F. Hamza-Lup, L. Davis, J. Daly, Y. Ha, and G. Martin. Development of a training tool for endotracheal intubation: Distributed augmented reality. In *Proceedings of Medicine Meets Virtual Reality*, volume 11, pages 288–294, Newport Beach, California, 2003.

[184] M. Rosenthal, A. State, J. Lee, G. Hirota, J. Ackerman, E. D. P. Kurtis Keller, M. Jiroutek, K. Muller, and H. Fuchs. Augmented reality guidance for needle biopsies: An initial randomized, controlled trial in phantoms. *Medical Image Analysis*, 6(3):313–320, 2002.

[185] R. Roy-Camille, G. Saillant, and C. Mazel. Internal fixation of the lumbar spine with pedicle screw plating. *Clin. Orthop.*, 203:7–17, 1986.

[186] E. Samset, E. Gjesteland, and M. Sæter. 3D graphical user interface for computer assisted surgery. In H. U. Lemke, M. W. Vannier, K. Inamura, A. G. Farman, K. Doi, and J. H. C. Reiber, editors, *CARS*, volume 1256 of *International Congress Series*, pages 414–418. Elsevier, 2003.

[187] A. P. Santhanam, C. Imielinska, P. Davenport, P. Kupelian, and J. P. Rolland. Modeling real-time 3-d lung deformations for medical visualization. *IEEE Transactions on Information Technology in Biomedicine*, 12(2):257–270, 2008.

[188] M. Sappol. *Dream Anatomy.* National Institutes of Health, 2006.

[189] F. Sauer, A. Khamene, B. Bascle, and G. J. Rubino. A head-mounted display system for augmented reality image guidance: Towards clinical evaluation for imri-guided neurosurgery. In *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages 707–716, London, UK, 2001. Springer-Verlag.

[190] F. Sauer, A. Khamene, B. Bascle, L. Schimmang, F. Wenzel, , and S. Vogt. Augmented reality visualization of ultrasound images: systemdescription, calibration, and features. In *Proc. IEEE and ACM Int'l Symp. on Augmented Reality*, pages 30–39, 2001.

[191] F. Sauer, A. Khamene, B. Bascle, S. Vogt, and G. J. Rubinob. Augmented reality visualization in imri operating room: System description and pre-clinical testing. In *Proceedings of SPIE, Medical Imaging*, volume 4681, pages 446–454, 2002.

[192] F. Sauer, A. Khamene, and S. Vogt. An augmented reality navigation system with a single-camera tracker: System design and needle biopsy phantom trial. In *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 2489, pages 116–124, 2002.

[193] F. Sauer, U. J. Schoepf, A. Khamene, S. Vogt, M. Das, and S. G. Silverman. Augmented reality system for ct-guided interventions: System description and initial phantom trials. In *Medical Imaging: Visualization, Image-Guided Procedures, and Display*, 2003.

[194] F. Sauer, F. Wenzel, S. Vogt, Y. Tao, Y. Genc, and A. Bani-Hashemi. Augmented workspace: designing an ar testbed. In *Proc. IEEE and ACM Int'l Symp. on Augmented Reality*, pages 47–53, 2000.

[195] I. S.B., M. W.C., P. E.R., L. G. D., and S. R.J. Features and uses of high-fidelity medical simulations that lead to effective learning: A beme systematic review. *Medical Teacher*, 27:10–29, Jan 2005.

[196] M. SC. An experimental comparison of three- versus four-surface phenomenal transparency. *Percept Psychophys.*, 35:325–32., 1984.

[197] B. Schwald, H. Seibert, and T. Weller. A flexible tracking concept applied to medical scenarios using an ar window. In *Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR)*, page 261, Washington, DC, USA, 2002. IEEE Computer Society.

[198] T. Sielhorst. High accuracy tracking for medical augmented reality. Master's thesis, Technische Universität München, October 2003.

[199] T. Sielhorst. *New Methods for Medical Augmented Reality.* PhD thesis, Technische Universität München (TUM), Chair for Computer Aided Medical Procedures, March 2008.

[200] T. Sielhorst, C. Bichlmeier, S. M. Heining, and N. Navab. Depth perception a major issue in medical ar: Evaluation study by twenty surgeons. In R. Larsen, M. Nielsen, and J. Sporring, editors, *Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Lecture Notes in Computer Science, 2006.

[201] T. Sielhorst, T. Blum, and N. Navab. Synchronizing 3d movements for quantitative comparison and simultaneous visualization of actions. In *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, pages 38–47, Vienna, Austria, Oct. 2005.

[202] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays: A literature review of augmented reality. *IEEE/OSA Journal of Display Technology; Special Issue on Medical Displays*, 4(4):451–467, December 2008.

[203] T. Sielhorst, M. Feuerstein, J. Traub, O. Kutter, and N. Navab. Campar: A software framework guaranteeing quality for medical augmented reality. *International Journal of Computer Assisted Radiology and Surgery*, 1(Supplement 1):29–30, June 2006.

[204] T. Sielhorst, T. Obst, R. Burgkart, R. Riener, and N. Navab. An augmented reality delivery simulator for medical training. In *International Workshop on Augmented Environments for Medical Imaging - MICCAI Satellite Workshop*, 2004.

[205] T. Sielhorst, J. Traub, and N. Navab. The ar apprenticeship: Replication and omnidirectional viewing of subtle movements. In *ISMAR '04: Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 290–291, Washington, DC, USA, 2004. IEEE Computer Society.

[206] C. B. Stapleton and A. Hirumi. Pedagogical implications for online instruction. submitted for publication 2010.

[207] A. State, D. T. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, and H. Fuchs. Case study: observing a volume rendered fetus within a pregnant patient. In *Proceedings of the conference on Visualization '94*, pages 364–368. IEEE Computer Society Press, 1994.

[208] A. State, M. Livingston, W. Garret, G. Hirota, M. Whitton, E. Pisano, and H. Fuchs. Technologies for augmented reality systems: Realizing ultrasound-guided need biopsies. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 439–446. ACM Press, 1996.

[209] D. Stoyanov, G. P. Mylonas, M. Lerotic, A. J. Chung, and G.-Z. Yang. Intraoperative visualizations: Perceptual fidelity and human factors. *J. Display Technol.*, 4(4):491–501, 2008.

[210] I. Sutherland. The ultimate display. In *Proceedings of IFIP Congress*, 1965.

[211] I. Sutherland. A head-mounted three dimensional display. pages 757–764, 1968.

[212] J. E. I. Swan, M. A. Livingston, H. S. Smallman, D. Brown, Y. Baillot, J. L. Gabbard, and D. Hix. A perceptual matching technique for depth judgments in optical, see-through augmented reality. In *VR '06: Proceedings of the IEEE Conference on Virtual Reality*, pages 19–26, Washington, DC, USA, 2006. IEEE Computer Society.

[213] P. Taylor. The notebooks of leonardo da vinci. *New York: The New American Library.*

[214] R. Thomas, N. John, and I. Lim. A mixed reality anatomy teaching tool. In *Proc. of Theory and Practice of Computer Graphics, Teesside*, pages 165–170, 2006.

[215] R. Thomas, N. John, and I. Lim. Anatomy education using rapid prototyping. In *Proc. of Theory and Practice of Computer Graphics, Bangor*, pages 251–258, 2007.

[216] J. K. A. K. Todd, A. Oomes. The perception of doubly curved surfaces from anisotropic textures. *Psychological Science*, 15:40–46, 2004.

[217] J. T. Todd, L. Thaler, and T. M. Dijkstr. The effects of field of view on the perception of 3d slant from texture. *Vision Research*, 45(12):1501–1517, 2005.

[218] M. Tosas. *Visual Articulated Hand Tracking for Interactive Surfaces*. PhD thesis, University of Nottingham, 2004.

[219] J. Traub. *New Concepts for Design and Workflow Driven Evaluation of Computer Assisted Surgery Solutions*. PhD thesis, Technische Universität München (TUM), Chair for Computer Aided Medical Procedures, November 2008.

[220] J. Traub, M. Feuerstein, M. Bauer, E. U. Schirmbeck, H. Najafi, R. Bauernschmitt, and G. Klinker. Augmented reality for port placement and navigation in robotically assisted minimally invasive cardiovascular surgery. In *Computer Assisted Radiology and Surgery*, pages 735–740, Chicago, USA, June 2004.

[221] J. Traub, P. Stefan, S.-M. M. Heining, C. R. Tobias Sielhorst, E. Euler, and N. Navab. Hybrid navigation interface for orthopedic and trauma surgery. In *Proceedings of MICCAI 2006*, LNCS, pages 373–380, Copenhagen, Denmark, Oct. 2006. MICCAI Society, Springer.

[222] J. Traub, P. Stefan, S.-M. M. Heining, C. R. Tobias Sielhorst, E. Euler, and N. Navab. Stereoscopic augmented reality navigation for trauma surgery: cadaver experiment and usability study. *International Journal of Computer Assisted Radiology and Surgery*, 1(Supplement 1):30–31, June 2006.

[223] M. Truppe, F. Pongracz, O. Ploder, A. Wagner, and R. Ewers. Interventional video tomography. In *Lasers in Surgery: Advanced Characterization, Therapeutics, and Systems*, volume 2395, pages 150–152, San Jose, CA, 1995. SPIE.

[224] A. Tsiaras and B. Werth. *The Architecture and Design of Man and Woman: The Marvel of the Human Body, Revealed*. Doubleday, firsts edition, 2004.

[225] unknown. *Anatomie Modèle Femme*. 1937.

[226] Y. Uralsky. Efficient soft-edged shadows using pixel shader branching. In *GPU Gems 2*. Ed. NVIDIA Corporation, 2005.

[227] M. A. van Veelen, C. J. Snijders, E. van Leeuwen, R. H. M. Goossens, and G. Kazemier. Improvement of foot pedals used during surgery based on new ergonomic guidelines. *Surgical Endoscopy*, 17(7):1086–1091, 2003.

[228] G. P. Viganò, S. Mottura, L. Greci, M. Sacco, and C. R. Boer. Internal fixation of the lumbar spine with pedicle screw plating. *International Journal of Computer Integrated Manufacturing*, 17:653–660, 2004.

[229] S. Vogt, F. Wacker, A. Khamene, D. R. Elgort, T. Sielhorst, H. Niemann, J. Duerk, J. Lewin, and F. Sauer. Augmented reality system for mr-guided interventions: Phantom studies and first animal test. In *Proceedings of SPIE, Volume 5367, Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, 2004.

[230] G. von Hagens. Körperwelten: press release, 2009.

[231] J. J. Vos. Depth in colour, a history of a chapter in physiologie optique amusante. *Clinical & Experimental Optometry*, 91:139–147, 2008.

[232] J. P. Wachs, H. I. Stern, Y. Edan, M. Gillam, J. Handler, C. Feied, and M. Smith. A gesture-based tool for sterile browsing of radiology images. *Journal of the American Medical Informatics Association*, 15(3):321–323, 2008.

[233] F. K. Wacker, S. Vogt, A. Khamene, J. A. Jesberger, S. G. Nour, D. R. Elgort, F. Sauer, J. L. Duerk, and J. S. Lewin. An augmented reality system for mr image - guided needle biopsy: Initial results in a swine model. *Radiology*, 238(2):497–504, 2006.

[234] A. Wagner, O. Ploder, G. Enislidis, M. Truppe, and R. Ewers. Virtual image guided navigation in tumor surgery – technical innovation. *Journal of Cranio-Maxillofacial Surgery*, 23(5):271 – 273, 1995.

[235] A. Wagner, O. Ploder, G. Enislidis, M. Truppe, and R. Ewers. Image-guided surgery. *International Journal of Oral and Maxillofacial Surgery*, 25(2):147 – 151, 1996.

[236] A. Wagner, M. Rasse, W. Millesi, and R. Ewers. Virtual reality for orthognathic surgery: The augmented reality environment concept. *Journal of Oral and Maxillofacial Surgery*, 55(5):456 – 462, 1997.

[237] B. Wallisch, W. Meyer, A. Kanitsar, and M. E. Gröller. Information highlighting by color dependent depth perception with chromo-stereoscopy. Technical Report TR-186-2-02-08, Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11/186, A-1040 Vienna, Austria, July 2002. human contact: technical-report@cg.tuwien.ac.at.

[238] L. Wang, J. Giesen, K. T. McDonnell, P. Zolliker, and K. Mueller. Color design for illustrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1739–1754, 2008.

[239] F. Wanschitz, W. Birkfellner, M. Figl, S. Patruta, A. Wagner, F. Watzinger, K. Yerit, K. Schicho, R. Hanel, F. Kainberger, H. Imhof, H. Bergmann, and R. Ewers. Computer-enhanced stereoscopic vision in a head-mounted display for oral implant surgery. *Clinical Oral Implants Research*, 13(6), 2002.

[240] T. Wendler, J. Traub, S. Ziegler, and N. Navab. Navigated three dimensional beta probe for optimal cancer resection. In R. Larsen, M. Nielsen, and J. Sporring, editors, *Proceedings of MICCAI 2006*, volume 4190 of *LNCS*, pages 561–569, Copenhagen, Denmark, Oct 2006. MICCAI Society, Springer.

[241] L. Williams. Casting curved shadows on curved surfaces. In *In Computer Graphics (SIGGRAPH '78 Proceedings*, pages 270–274, 1978.

[242] F. Wimmer. Diploma thesis: Focus and context visualization for medical augmented reality. Master's thesis, Technische Universitaet Muenchen, Department of Computer Science, Chair for Computer Aided Medical Procedures and Augmented Reality, 2007. 1.

[243] F. Wimmer, C. Bichlmeier, S. M. Heining, and N. Navab. Creating a vision channel for observing deep-seated anatomy in medical augmented reality. In *Proceedings of Bildverarbeitung fuer die Medizin (BVM 2008)*, Munich, Germany, April 2008. BVM.

[244] D. Winkel, A. Vleeming, and O. G. Meijer. *Anatomie in vivo für den Bewegungsapparat*. Elsevier, München, Oct 2003.

[245] M. P. B. A. T. L. Winn B., Strang N.C. Reversals of the colour-depth illusion explained by ocular chromatic aberration. *Vision Research*, 35 (10):2675–2684, 1995.

[246] D. L. Wright, J. P. Rolland, and A. R. Kancherla. Using virtual reality to teach radiographic positioning. *Radiologic Technology*, 66(4):233–8, 1995.

[247] P. Wucherer. Diploma thesis: Multimodal consultation system for patient education in plastic surgery. Master's thesis, Technische Universitaet Muenchen, Department of Computer Science, Chair for Computer Aided Medical Procedures and Augmented Reality, 2009. 1.

[248] P. Wucherer, C. Bichlmeier, M. Eder, L. Kovacs, and N. Navab. Multimodal medical consultation for improved patient education. In *Proceedings of BVM 2010*. Springer, Mar. 2010.

[249] L. W. Yaggy and J. J. West. *Yaggy's Anatomical Study*. Western Publishing House, 1886.

[250] S.-X. Zhang, P.-A. Heng, Z.-J. Liu, L.-W. Tan, M.-G. Qiu, Q.-Y. Li, R.-X. Liao, K. Li, G.-Y. Cui, Y.-L. Guo, and Y.-M. Xie. Chinese visible human data sets and their applications. In *HCI (12)*, pages 530–535, 2007.