Dissertation

# Introducing Extended Reality to Ophthalmic Surgery

Hessam Roodaki

# Technische Universität München

Fakultät für Informatik
Lehrstuhl für Informatikanwendungen in der Medizin

# Introducing Extended Reality
# to Ophthalmic Surgery

## Hessam Roodaki

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.

| | |
|---|---|
| *Vorsitzende(r):* | Prof. Dr. Michael Gerndt |
| *Prüfer der Dissertation:* | Prof. Dr. Nassir Navab |
| | Jun.-Prof. Dr. Christian Hansen |

Die Dissertation wurde am 10.10.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 27.01.2020 angenommen.

# Abstract

Ophthalmic surgical interventions are among the most commonly performed operations in clinics around the world with population aging contributing towards them becoming the most frequent class of interventions. Nevertheless ophthalmic procedures are complex and require significant manual dexterity. Major segments or often the entirety of ophthalmic surgical interventions are performed on human tissues covering merely a few square millimeters under surgical microscopes leading to diminished depth perception and problems with hand-eye coordination. On the other hand, recent developments in medical imaging and computer aided interventions have brought additional imaging modalities such as Optical Coherence Tomography (OCT) and supplementary sources of information into the operating rooms. Although extra intraoperative sources of information promise positive impact on surgical outcomes, they inherently add more complexity to surgical routines. Controlling microscopic devices with information injection capabilities during a medical operation requires extra care and presumably more staff in operating rooms.

Extended Reality (XR) is an effective, efficient and robust approach for information conveyance in complex settings. It has found a foothold in many industries as a cutting-edge tool used in information-rich environments. Medical extended reality systems are already in use in operating rooms and have shown to have a positive effect on quantitative measures defining surgical outcomes. Availability of OCT as a secondary imaging modality along with employment of surgical microscopes make Augmented Reality (AR) and Virtual Reality (VR) suitable information conveyance methods for ophthalmic procedures.

This thesis is an effort to establish the practice of employing extended reality techniques in ophthalmic operating rooms by two contributions. On the one hand, the impact of employing XR on ophthalmic surgical workflows and outcomes is investigated. On the other hand, advancement of XR to applications beyond surgical training and simulation is attempted. To these ends, new computer aiding and guidance solutions using AR and VR are proposed, experiments to assess the significance of such tools are conducted and outcomes of the proposals are compared to conventional methods of representing medical data. Novel approaches introduced in this thesis demonstrate the potential of XR to facilitate complex procedures including but not limited to eye surgery.

# Zusammenfassung

Ophthalmochirurgische Interventionen gehören zu den mitunter am häufigsten durchgeführten Eingriffen in Kliniken weltweit. Die stetige Alterung der Gesellschaft trägt dazu bei, dass sie in naher Zukunft die am häufigsten durchgeführten Eingriffe werden. Doch ophthalmochirurgische Prozeduren sind komplex und erfordern ein erhebliches Maß an manueller Geschicklichkeit. Ein Großteil oder vielmals sogar die Gesamtheit eines ophthalmochirurgischen Eingriffe werden unter einem Operationsmikroskop an menschlichem Gewebe ausgeübt, welches eine Fläche von weniger als einigen Quadratmillimetern umfasst. Die Nutzung des Operationsmikroskops bedingt eine reduzierte Tiefenwahrnehmung und herabgesetzte Hand-Auge-Koordination. Dahingegen haben jüngste Entwicklungen in medizinischer Bildgebung und computergestützer Chirurgie zusätzliche Bildgebungsmodalitäten, darunter die optische Kohärenztomographie (OCT), und ergänzende Informationsquellen hervorgebracht. Obwohl zusätzliche intraoperative Informationsquellen positive Auswirkungen auf chirurgische Resultate versprechen, gehen sie grundsätzlich mit zunehmender Komplexität in der chirurgischen Routine einher. Das Bedienen von mikroskopischen Geräten mit der Fähigkeit, intraoperativ Information zu visualisieren, erfordert besondere Sorgfalt und vermutlich mehr Personal im Operationssaal.

Extended Reality (XR) ist ein effektiver, effizienter und robuster Ansatz zur Übermittlung von Informationen in komplexen Umgebungen. In zahlreichen Industrien hat es sich als innovatives Werkzeug für informationsreiche Umgebungen etabliert. Medizinische Extended Reality Systeme sind bereits in Operationssälen im Einsatz und zeigen einen positiven Effekt auf quantitative Messungen, welche bezeichnend für chirurgische Resultate sind. Die Verfügbarkeit von OCT als sekundärer Bildgebungsmodalität, einhergehend mit der Nutzung von chirurgischen Mikroskopen, machen Augmented Reality (AR) und Virtual Reality (VR) zu geeigneten Methoden für die Informationsübermittlung bei ophthalmochirurgischen Prozeduren.

Diese Arbeit stellt einen Versuch der Etablierung des Einsatzes von Extended-Reality-Verfahren in der ophthalmochirurgischen Praxis mittels zweier Beiträge dar. Einerseits soll die Auswirkung der Anwendung von XR in ophthalmochirurgischen Arbeitsabläufen und auf deren Ergebnisse untersucht werden. Andererseits soll der Versuch einer Ausweitung von XR über chirurgische Simulationen und Training hinaus unternommen werden. Zu diesen Zwecken sollen neue computergestützte und computergeführte Verfahren, aufbauend auf AR und VR, vorgestellt werden. Hierzu werden Experimente durchgeführt, welche die Signifikanz solcher Werkzeuge beurteilen sollen. Die Ergebnisse der Implementation dieser Verfahren sollen mit konventionellen Methoden zur medizinischen Datenpräsentation verglichen werden. Die neuen Ansätze, die in dieser Arbeit vorgestellt werden, zeigen das Potenzial von XR auf, komplexe Prozeduren zu erleichtern; in der Augenchirurgie, allerdings auch darüber hinaus.

# Acknowledgments

# Contents

# Introduction

## 1.1 Outline

This thesis is structured into four chapters and an appendix. The following section briefly describes the main topics covered in each chapter.

**Chapter 1 (Introduction)** is an introduction to ophthalmic surgical procedures. First, a general overview of the current state of surgical interventions in ophthalmology and observed challenges is laid out. Based on this overview, the motivations and aims of the proposals in this thesis are introduced. Then, the fundamentals of ophthalmic surgery including the most frequently performed or challenging surgical procedures of the anterior and posterior segments of the eye are described. To cover the technical aspect of surgical interventions performed on eyes, essential concepts and methodologies in microscopic surgery are then shortly introduced. Finally, basic principles of optical coherence tomography including the latest developments and shortcomings in intraoperative approaches are discussed.

**Chapter 2 (Augmented Reality (AR) for Ophthalmic Procedures)** covers the first line of contributions in this thesis. It starts with discussing the motivation behind employing augmented reality as an assistive tool for ophthalmic surgical interventions. Then, challenges associated with augmented reality in operating rooms and high-precision tasks are expressed. Applications of augmented reality in eye surgeries are then discussed. A brief analysis of the state of the art in both visual and auditory augmented reality follows. Finally, contributions of the author as proposed methods for using augmented reality in ophthalmic operating rooms are portrayed.

**Chapter 3 (Virtual Reality (VR) for Ophthalmic Procedures)** covers the second line of contributions in this thesis. First the incentive behind employing virtual reality for medical high-precision tasks and the challenges associated to that are laid out. Then examples of ophthalmic procedures that benefit from employing virtual reality as an assistive tool are presented. Similar and related studies and proposals for using virtual reality in surgery and in particular ophthalmic procedures are then discussed. Finally, proposals and discussions as contributions of the author in this field are introduced.

**Chapter 4 (Conclusion)** concludes this thesis by summarizing and discussing the main objectives and elements of the work and illustrating a roadmap for future efforts.

## 1.2 Background and Motivation

According to surgical procedure statistics acquired in 2012, cataract surgery has become the most frequent surgical procedure in most European countries due to population aging and the increase in this intervention among the elderly [49]. Estimates from 2009 show that in Germany alone, more than 600,000 cataract operations are performed annually and the corresponding figure worldwide is estimated to be between 6 and 10 million [46]. Cataract surgery is only one example of hundreds of ophthalmic procedures performed annually around the world. Most of the major eye diseases are age-related, in that the prevalence of these sight-threatening diseases dramatically increases above 75 years of age. World estimates in 2013 show that overall 65% of those with visual impairment and 82% of those who are blind are over 50 years of age [16]. With the growth in population and inversion in the population pyramid the need for ophthalmic procedures is projected to ever increase. A survey conducted in 2012 show that despite over 200,000 ophthalmologists worldwide, there is a significant shortfall of ophthalmologists in developing countries. Furthermore, although the number of practitioners is increasing in developed countries, the population aged 60+ is growing at twice the rate of the profession [68]. Hence, the gap in the number of ophthalmic practitioners in general and surgeons in particular to cover the needs of societies in near future and the actual number of available active ophthalmic surgeons is widening.

Eye surgeries generally fall into two major categories. Anterior segment and posterior segment procedures. Surgical interventions performed on the exterior part of the eye including the cornea, the lens and the iris are defined as anterior segment procedures. Similarly, surgical interventions performed inside the eyeball including the retina and vitreous humor are categorized as posterior segment procedures. Surgical interventions performed on the eye are for the most part manual procedures performed under a surgical microscope. The magnification and illumination provided by a surgical microscope is essential for the performance of the surgeon since the operation is limited to small areas. The focus of an anterior segment procedure is on a few centimeters of eye tissue while a posterior segment procedure is often performed on tissues covering only a few millimeters. Recordings of hand movements of an ophthalmic surgeon under realistic conditions show hand tremors with root mean square amplitude of 38 micrometers [84]. On the other hand, another study shows that the majority of interactions between surgical instruments and retinal tissue in a microsurgery are not "felt" by the surgeon [38]. All these complexities make ophthalmic surgery one of the most challenging procedures. The challenges associated with ophthalmic surgery has a negative impact not only on the time and cost of performed operations, but also on the time and cost of training skilled surgeons.

Computer-Assisted Interventions (CAI) are the group of surgical treatments that employ computer algorithms and tools as supports for better surgical outcomes. A better surgical outcome is defined as any combination of more accurate manipulations, less time-consuming surgical phases, less invasive operations, less surprising outcomes, etc. When it comes to complex surgical interventions such as ophthalmic procedures, guidance and assistance from computer methods and tools has the potential of turning lengthy, challenging and error-prone maneuvers into consistent and fluid tasks. Since the increase in the number of ophthalmic procedures performed globally and the shortage in ophthalmic surgeons is

apparent, computer-assisted surgery is a fitting answer. Whether used as training tools or assistive accessories, computer-assisted methods are capable of increasing the public access to ophthalmic interventions in shorter time frames with less cost. Computer-assisted surgical tools come in a multitude of different forms each with capabilities specifically designed to target different aspects of surgical tasks. In this thesis the focus is on techniques used for medical data representation particularly used for conveying information from intraoperative imaging modalities.

Extended Reality (XR) is the term that is used for collectively referring to Augmented Reality (AR), Mixed Reality (MR), Virtual Reality (VR) and other environment-enriching technologies that fall in between. These technologies tend to combine information from multiple often complimentary sources and deliver them as enriched sensory experiences. Since the 1990s, researchers began proposing and partially validating augmented reality concepts and prototypes specifically aimed at the medical field. AR has a huge potential to enable seamless integration of novel technologies into the clinical workflow and to facilitate visualizing novel multimodal imaging within the human body [59]. Ophthalmic surgical theater in particular is a suitable candidate for seamless implementation of visual and auditory extended reality solutions. On the one hand, the presence of a surgical microscope means that the surgical scene is viewed by surgeons through either ocular lenses that could be equipped with Heads-Up-Displays (HUD) or displays that could overlay excessive information. On the other hand, availability of multiple preoperative and intraoperative imaging modalities such as eye fundus images or Optical Coherence Tomography provides the opportunity for extended reality to simplify medical data representation.

Early research on the application of augmented reality for eye surgery and ophthalmic procedures in virtual environments has been conducted particularly in the past decade. Prototypes with both training and assistance in mind have been evaluated. In the following chapters examples of such systems and the state of the art in methods involved are introduced and discussed. Nevertheless the research in this field is still in its infancy with open questions regarding the specific applications, means of integration, impacts, obstacles, outcomes and performance of extended reality techniques in ophthalmic procedures. Difficulties in integration of XR solutions in the ophthalmic operating rooms should be addressed. Different XR approaches targeted for solving similar problems should be evaluated and compared. Most importantly, feasibility of employing XR solutions in realistic setups should be tested. Attempting to answer these questions requires close collaboration with industries providing solutions for operating rooms and physicians active in surgical therapeutics.

The aim of this work is to propose extended reality methodologies in the field of ophthalmic procedures while addressing parts of the mentioned scientific open questions. By trying to overcome the common challenges in ophthalmic surgical interventions using XR, the complexity and complications of surgical therapeutics are hoped to decrease. This in turn would make surgical options more accessible and less costly for the public. Training of surgical staff using approaches introduced in this work has the potential of targeting the need for more practitioners in the field. Contributions of the author to this end are quantitatively evaluated in setups closely resembling real-life operating theaters and discussed in the relevant scientific communities. Techniques introduced here are meant to be applicable in conventional operating setups for ophthalmic procedures and perhaps other surgical interventions.

## 1.3  Essentials of Ophthalmic Surgery

In this chapter, the most common medical conditions of the human eye along with a brief description of the surgical approaches taken to treat or manage these conditions are introduced.
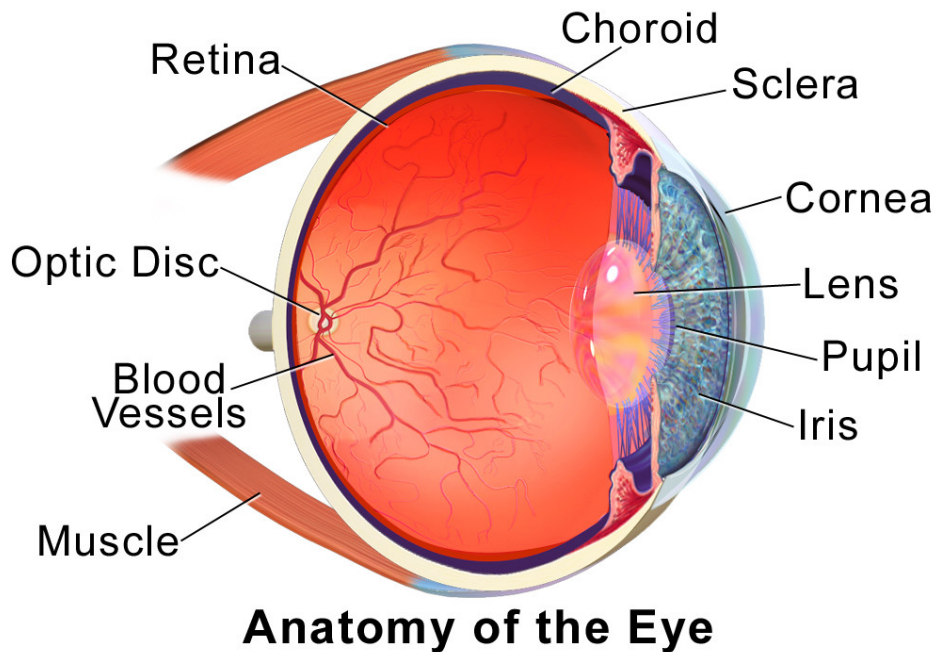


*Source: Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2). DOI:10.15347/wjm/2014.010. ISSN 2002-4436*

**Fig. 1.1.**   Major Anatomical features of the human eye.

The human eye can be divided into two principal parts. The anterior segment of the eye is considered as structures anterior to (situated in front of) the vitreous humor which is the clear liquid between the lens and the retina. The major anatomical features in the anterior segment of the eye are the lens, the iris, the cornea and the sclera. The posterior segment of the eye is likewise defined as the tissues inside the eye including the vitreous humor, the retina and the choroid. As illustrated in Fig. 1.1, performing surgery on the posterior segment of the eye in contrast to the anterior segment requires the surgical instruments to be placed inside the eyeball. Visual access to the posterior side is also only possible through the pupil.

Image formation in the eye starts by rays of light penetrating into the transparent vitreous humor via the pupil. Bending of the light rays happens in the cornea, the lens and also in the aqueous humor which is the clear liquid filling the anterior chamber between the lens and the cornea. In normal eyes, light rays are focused on a small area on the retina called the macula. Since proper image formation on the macula is dependent on the precise structure of the eyeball, a group of medical conditions responsible for compromised vision are caused by structural defects of the eyeball originating from genetic abnormalities or changes over time. Another set of conditions are caused by external damage including trauma or various

infections leading to partial or total, temporary or permanent damage to the tissues in the eye. A third group of diseased are caused by degeneration or decline in the functionalities of specific tissues in the eye usually caused by the process of aging. Some of the medical conditions affecting the human eye could be treated or managed by surgical interventions.



**Fig. 1.2.** Fundus image of a human eye. The area annotated with the yellow marker is the macula, the dark spot in the middle of the macula is fovea and the white area annotated with green is the optic disc.

Various imaging modalities are used for diagnosis of medical conditions associated with the eye. For the anterior segment of the eye, the simplest form of examination could be done by shining intense focused light into the eye. Devices such as slit lamps that shine intense thin sheets of light into the eye and can image the eye using cameras are used for diagnosis of cataract and corneal pathologies. The most frequent and readily available modality used for imaging the retina is fundus imaging using a color camera. Color fundus images are effective at visualization of deviation from normal in retinas. Fig. 1.2 shows a fundus image from a human retina. Optical Coherence Tomography is an imaging modality that employs near-infrared light to form images of transparent or semi-transparent tissues. It is used for both anterior and posterior examination and is excellent at visualizing different layers in

tissues. A combination of different imaging techniques and examination methods are typically used for diagnosis of medical conditions in the eye. Preoperative fundus and OCT images are commonly employed for surgical planning.

## 1.3.1 Anterior Segment Interventions

Surgeries performed on the anterior segment of the eye are very common interventions that resolve or manage medical conditions that are among the most prevalent. In almost all cases, an anterior segment surgery is performed under a surgical microscope that provides magnification and illumination for the surgeon. Manipulation of exterior eye tissue is done manually using instruments such as picks, forceps, needles and knives. Here some of the most common conditions and the surgical procedures used to manage them are listed.

**Cataract surgery** is the most performed ophthalmic procedure worldwide. Cataract is a condition that affects the lens of the eye through hardening and clouding and is the leading cause of avoidable blindness worldwide [69]. The most common cause is aging with genetics playing a strong role. The clouding of the lens directly affects visual acuity. Cataract surgery involves removing the damaged lens followed by insertion of an artificial Intraocular Lens (IOL) mimicking the same optical properties. Removal of the damaged lens could be performed by directly removing the lens using an incision into the anterior chamber. More modern methods use a technique called phacoemulsification which uses ultrasonic vibrations to break down the lens into smaller pieces that cold be aspirated out of the anterior chamber. More recently lasers have also been used for breaking down the lens into pieces. The power of IOL lenses to be inserted in place of the natural lens are typically calculated based on measurements of the eye performed preoperatively. IOLs are inserted into the anterior chamber using an instrument called a lens injector.

**Glaucoma surgery** is the procedure dealing with the flow of aqueous humor. Glaucoma is a group of condition affecting the optic nerve of the eye. Any blockage in the natural pathways that let aqueous humor to flow outside freely leads into the build up that increases the Intraocular Pressure (IOP). The increase in pressure over time deteriorates the optic nerve causing tunnel vision and eventually blindness. In glaucoma surgery, either by incisions into the trabecular meshwork or by insertion of shunts, tubes or stents, the natural drainage of the aqueous humor is restored.
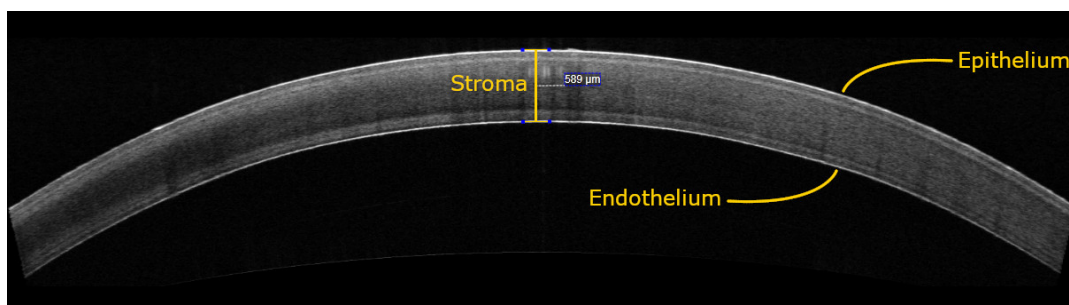


**Fig. 1.3.** Optical coherence tomography of a human cornea. The yellow annotations are indicating the most prominent features of the cornea.

**Corneal refractive surgeries** are a group of procedures that are performed to improve the refractive state of the eye. Refractive surgeries are commonly performed to treat myopia, astigmatism, keratoconus and similar medical conditions affecting the vision by removing portions of corneal thickness. Among the most employed methods of refractive surgery are Radial Keratotomy (RK) which involve direct cut and removal of corneal layers by surgical instruments and usage of high speed (femtosecond) lasers in procedures such as Laser-Assisted Sub-Epithelial Keratectomy (LASEK) and Laser-Assisted In Situ Keratomileusis (LASIK) [54]. Calculation of the amount of cornea to be removed and the exact geometry is performed based on parameters extracted from preoperative acquisitions including OCT images. The latest developed refractive procedure involving femtosecond lasers is called Small Incision Lenticule Extraction (SMILE) which involves the least amount of incisions introduced to the cornea. The visual and refractive outcomes of the SMILE procedure have been shown to be similar to LASIK while there is increasing evidence for the benefits of SMILE over LASIK by leaving the anterior stroma intact and faster recovery [65].
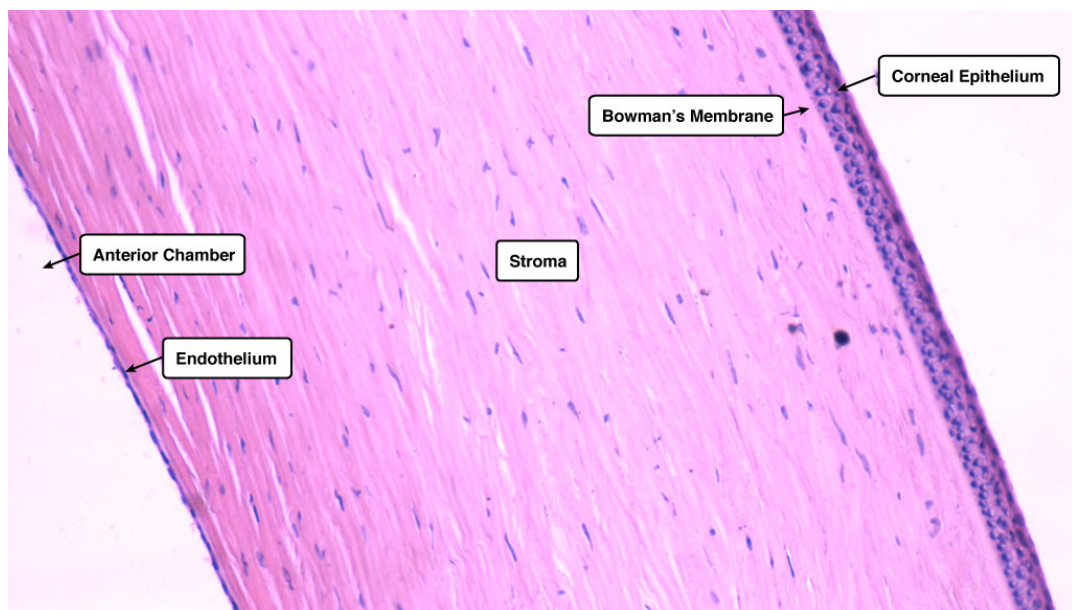


*Source: Yale School of Medicine, Department of Cell Biology, http://medcell.med.yale.edu*

**Fig. 1.4.** Structure of the human cornea in a histology slice. The three main cellular layers of human cornea are the epithelium, the stroma and the endothelium

**Keratoplasty** or corneal transplantation is a procedure in which part or the whole thickness of a patient's cornea is resected and a donor graft is sutured into its place. Corneal pathologies such as keratoconus, corneal ulcers, trauma, clouding, swelling, scarring or thinning of the cornea can be treated or managed by keratoplasty. As illustrated in Fig. 1.3, human cornea has three major layers. The outermost layer called the epithelium, the inner thickness called the stroma and the innermost layer called the endothelium. Inspection of the human cornea using microscopic techniques shows the fine cellular structure of the epithelium and the endothelium (Fig. 1.4). In humans, the corneal epithelium has a mean thickness of $53.4 \mu m$ [66] while the endothelium is a single layer of cells $5 \mu m$ in height [89]. Descemet membrane which lies between the endothelium and the stroma can be as thick as $12 \ \mu m$ [86]. Cellular layers in the cornea play a crucial role in the general function of the eye. For instance,

the main function of the endothelial layer is to control corneal hydration and nutrition by forming a leaky barrier [94]. There are various techniques for corneal transplantation. The traditional technique called Penetrating Keratoplasty (PKP) involves transplantation of the whole thickness of the cornea. PKP is necessary when the damage to the cornea is considered total. However, when certain layers of cornea are healthy, preserving them could be beneficial due to their essential role, less risk of rejection and faster healing. Deep Anterior Lamellar Keratoplasty (DALK) is a corneal transplantation technique in which the epithelium and the stroma is removed while the host Descemet membrane and the endothelium are retained. The donor graft should be prepared to only contain the anterior cornea. Although performing DALK compared to PKP is more complicated and time-consuming, preserving the healthy endothelium of the patient is beneficial [15]. Complexity of a surgical task such as DALK comes from the fact that intraoperative continuous imaging and tracking of cellular layers only a few micrometers in thickness and manually manipulating them using instruments is rather challenging. There are also other keratoplasty methods that involve removal of endothelium and the Descemet membrane and preserving the stroma and the epithelium (Fig. 1.5).
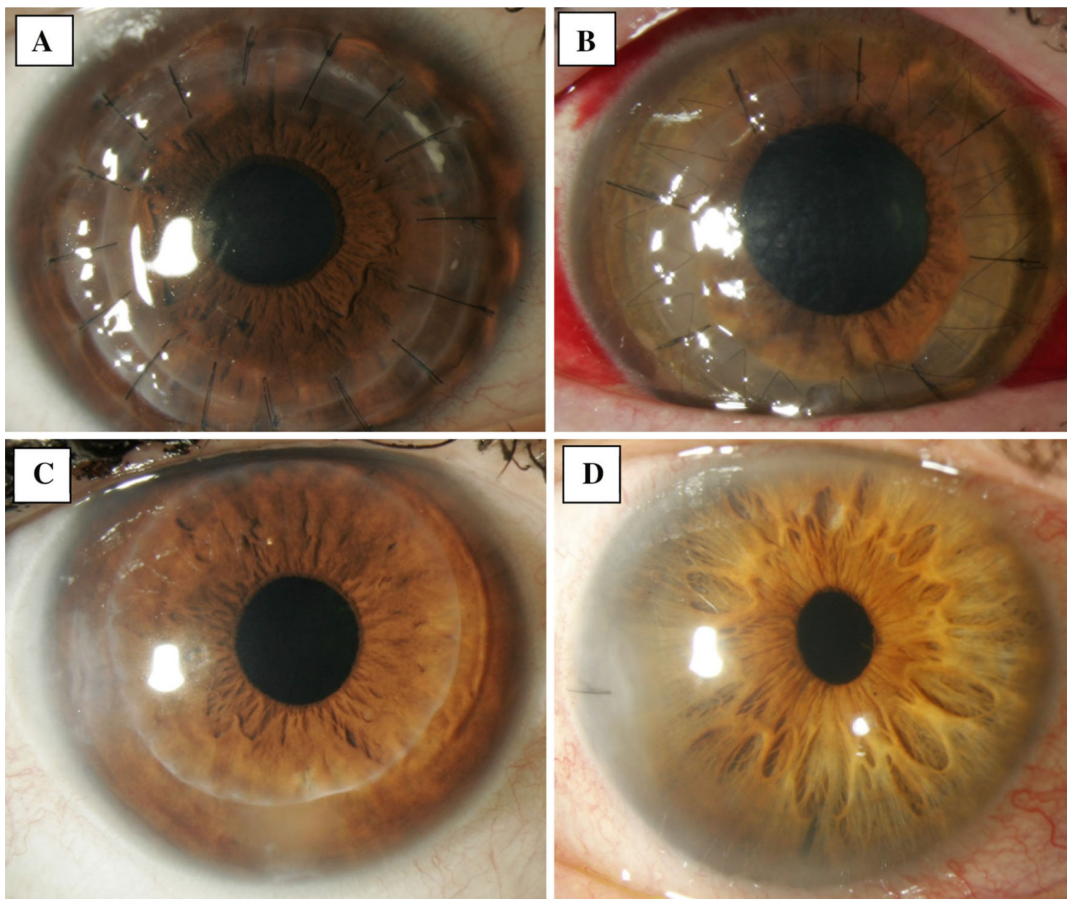


*Source: [15] G. E. Boynton et al., Current surgery reports, 3.2 2015, © Springer.*

**Fig. 1.5.** Photographs after: (A) Penetrating keratoplasty, (B) Deep anterior lamellar keratoplasty, (C) Descemet stripping automated keratoplasty, (D) Descemet membrane endothelial keratoplasty.

## 1.3.2 Posterior Segment Interventions

Surgeries performed on the posterior segment of the eye are inherently more complex. They typically involve insertion of trocars into the sclera of the eye which provide portals for the surgeon to have access to the interior of the eyeball. Since illumination inside the eyeball should be artificially provided for the surgeon to be able to see the tissues, a trocar typically provides access for a handheld illumination probe (endoilluminator) or a light fixture known as a chandelier illumination probe (chandelier endoilluminator). Another trocar provides access for a handheld surgical instrument for manipulation of posterior tissues. A third portal could provide irrigation with fluids such as sterile Balanced Saline Solution (BSS). In some cases where two instruments are needed simultaneously, an extra fourth trocar is also put in place. The only possible way for the surgeon to have visual access to the interior of the eye is through the pupil. Since the area which a typical posterior segment intervention is focused on is only a few square millimeters, a microscope is needed to provide the magnification necessary for the surgeon. Some of the most common posterior segment surgical treatments and the medical conditions associated to them are listed below.
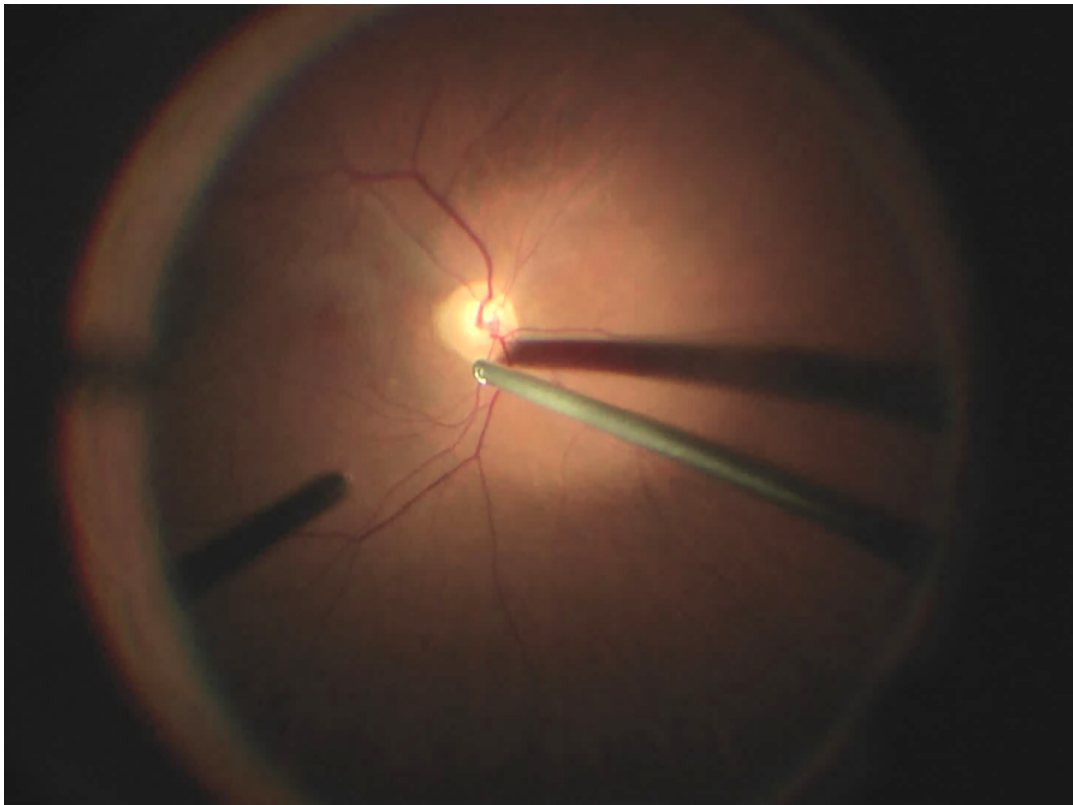


**Fig. 1.6.** Surgeon's microscopic view of a vitrectomy procedure. The vitrectomy probe (right) has a small opening at the tip for cutting and aspiration. The endoillumination probe (left) provides the necessary light.

**Vitrectomy** is a surgical treatment in which the vitreous humor inside the eye is partially or totally removed and substituted with another fluid of choice. Such an intervention is commonly performed using a transconjunctival sutureless technique [35]. A vitrectomy probe and fluid irrigation are introduced via trocars. The vitrectomy probe has integrated micro-cutters and aspiration combined into a small cannula. Since the vitreous humor has

gel-like consistency with collagen protein strands connecting it to the lens capsule and the optic nerve on both ends, cutters are needed for effective removal of the humor. For total removal of the vitreous humor, the surgeon moves the vitrectomy probe around in the eye in a circular motion while controlling the cut and aspiration rate via foot control pedals. Devices such as trocars, syringes, endoillumination probes and vitrectomy probes that come in cylindrical shapes are measured in gauge system. Most ophthalmic surgery techniques are introduced using 20-gauge devices which are among the largest in ophthalmic surgery. Later efforts are usually in the direction of bringing instruments and devices with smaller 23-, 25- and even 27-gauge into market. Using 25- or 23-gauge vitrectomy techniques instead of 20-gauge may shorten operating time, improve patient comfort, and recovery speed [96]. Vitrectomy is performed to improve vision when external particles such as debris or floaters make the vitreous humor cloudy. Damaged blood vessels of the retina can lead to blood leaking into the vitreous (vitreous hemorrhage) which in turn leads to deteriorated vision. Another example of complications in connection to the vitreous humor is infections inside the eye. Vitrectomy is also beneficial when shrinkage in the vitreous, typically caused by the process of aging, lifts the retina causing tears or detachment. In such scenarios the vitreous humor is substituted with fluids such as silicone oil or heavy liquids to keep the retina in place [5]. Vitrectomy could also be performed as a step before another surgery so that the surgeon has access to the retina and macula.
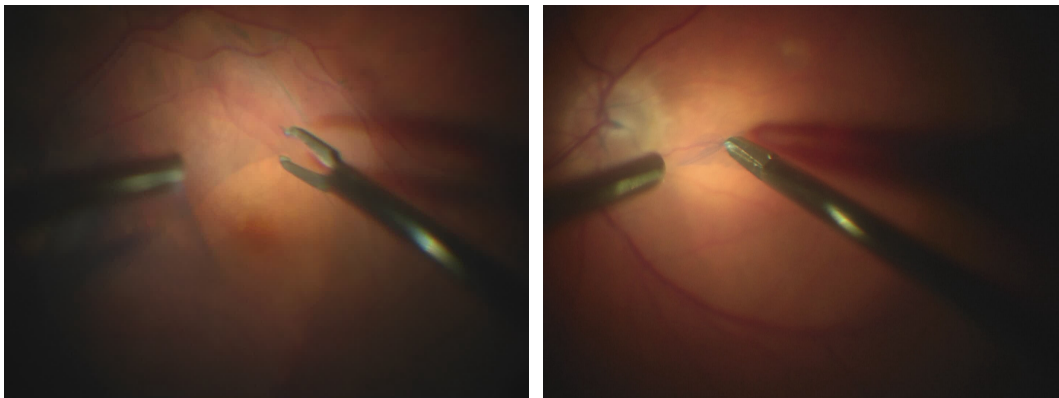


**Fig. 1.7.**   Microscopic view of an epiretinal membrane peeling procedure. The forceps is manually brought as close as possible to the retina without touching, then closed to grasp and remove the membrane.

**Epiretinal Membrane Peeling (Membranectomy)** is a procedure performed via various types of forceps, picks or scrapers to remove the thin Epiretinal Membrane (ERM) around the macula. ERM is a semi-translucent tissue formed on the surface of retina and is observed to be in the order of 60 micrometers in thickness [95]. Formation of ERM does not have a clearly identifiable cause and in most cases has no major impact on visual acuity. However, when formation of ERM causes traction on the retina or the vitreous to partially separates from the retina, it leads into blurry vision, reduced visual acuity, difficulty using both eyes together or other visual problems [34]. The excess traction on the retina caused by the formation of ERM could also lead into a tear in the macular structure causing a so called macular hole. Removing the ERM from the retina eases the tension leading to closure in the formed hole. Epiretinal membrane peeling is always done after a vitrectomy procedure. A pair of forceps, a pick or a retinal scraper is introduced via the trocars into the eyeball. Often a dye is injected into the

eye which stains the ERM for better visibility. Then the surgeon attempts to manually remove the ERM piece by piece by grasping it and pulling away or by scraping it away from the retina. The removed tissue is then aspirated out of the eye using a vitrectomy probe. One of the complexities associated with ERM peeling is the unnecessary touch of the retina. Collision of the instrument with the retina could cause hemorrhage or trauma to the photoreceptors. Since the ERM is not clearly visible and is thin, grasping and removing it involves micro-forces that are imperceptible to human touch.

**Subretinal Injection** is an intervention to deliver pharmaceuticals into specific layers of retina as opposed to intravitreal injections in which pharmaceuticals are injected into the vitreous humor. It has recently gained attention in the ophthalmology community amid recent advances in gene and stem-cell therapy. The idea behind a subretinal injection is to deliver a custom gene vector, stem-cells or drugs specific to a patient into a particular layer of retina where the active elements of the delivered fluid trigger or enhance cell generation for treatment of visual impairments. Subretinal injection provides better and safer effects in gene and cell therapies and might be considered as a potential delivery option for personalized medical care with specific targets in the subretinal space [62]. Since delivery of fluids into specific layers of retina essentially requires cellular level sensing and manipulation, most introduced techniques are still in their infancy with many active trials around the globe. Some trials use preoperative planning with optical coherence tomography and autofluorescence (AF) for cellular level tracking of the induced alterations [97]. Using robots for performing intraocular surgeries has had various applications in the past but now subretinal injection has become one of the major applications of robotic surgeries. A safe and viable robotic solution for intraocular surgery would enable precise, feasible and minimally traumatic delivery of gene therapy or cell therapy to the retina [27]. In current robotic approaches for subretinal injection, the robot head is controlled by the surgeon via a controller in a master-slave fashion. Another approach to ease the procedure would be to fuse information from multiple imaging modalities such as microscopy, OCT and autofluorescence to navigate the surgeon into the target area in the retina. Other than assistance in injection, developing methods that can quantify the amount and distribution of the deposited fluids during and after injection are of great interest.
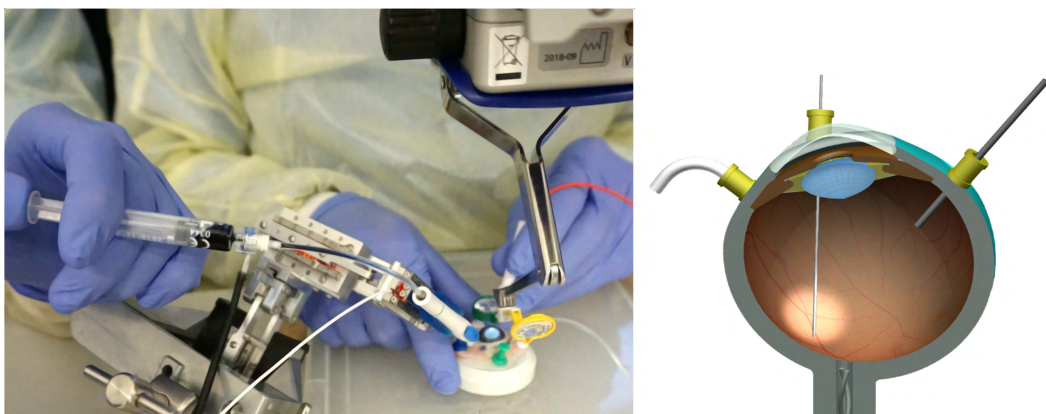


*Source: [100] M. Zhou et al., IEEE Access, 7 2019, © 2019 IEEE.*

**Fig. 1.8.** A trial robot-assisted subretinal injection procedure performed on an *ex vivo* porcine eye. The surgeon controls the illumination probe and the robot via a controller, when the needle reaches the target layer, the assistant starts the injection process.

## 1.3.3 Microscopic Surgery

Ophthalmic procedures are among the interventions that are typically performed under surgical microscopes due to the small size of the target surgical area. Microscopes used for ophthalmic procedures come in either standing form factor or as a mountable module usually mounted on the ceilings of operating rooms. The head of the microscope can freely move around the base to be positioned above the head of the patient who is in lying position. An ophthalmic surgical microscope in its simplest form comprises of an array of lenses that divert the light from the patient eye into the surgeon's eyes. For the surgeon to have proper depth perception, two optical pathways for the left and right eye are necessary. The illumination of the anterior segment of the eye could be accomplished via the same optical pathway or from a separate dedicated one. In interventions such as cataract surgery, having a so called coaxial illumination that shares the same optical pathway as the surgeon's view point has benefits. Coaxial illumination provides reflection back from the retina into the microscope which helps the surgeon to better visualize the cornea and the structures in the anterior capsule of the eye. Various types of light sources such as Xenon surgical lights or LEDs are used for illumination.

Although the basic functionality of a surgical microscope is to provide magnification and illumination, current line of microscopes found in most operating rooms provide much more than that. An integrated video camera in the microscope is usually used for the documentation of procedures. Having a pair of cameras for stereo imaging can provide a digital view of the scene to the surgeon via 3D displays. One benefit of a digital view is that it can be shared between multiple surgeons in the operating room. It can also be digitally enhanced or overlaid with additional information. Studies have shown that some patients develop retinopathy after undergoing a microscopic surgery due to exposure to the surgical light (light toxicity) [44]. Using cameras instead of direct optical view for the surgeon has the benefit that because of the high sensitivity of camera sensors, the light intensity required for the surgical scene to be considered well-lit becomes much less. Clinical trials have shown that with real-time digital processing and automated brightness control surgeons could operate comfortably at an endoillumination level equivalent to 10% of the maximum output [1].



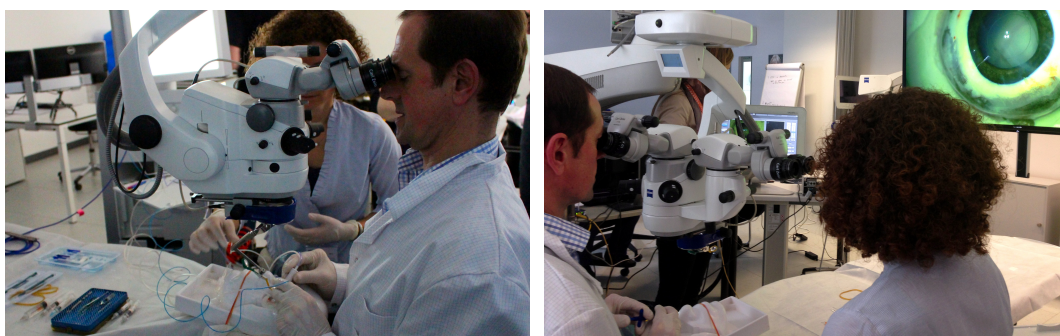**Fig. 1.9.** Ophthalmic microsurgery simulation on an *ex vivo* porcine eye from multiple views. The microscope provides views for a surgeon and an assistant through eyepieces and digital view via a display.

Performing surgery with a microscope with three-dimensional digital view of the surgical field on a display is called surgery in heads-up position. Among the advantages of heads-up surgery

over the traditional eyepiece view approach are the possibility of digital signal enhancement, improved depth of field, digital filtering and high dynamic field. In a clinical study with volunteer surgeons performing vitrectomies in a heads-up setup, 91.7% of the volunteers preferred the ergonomics of the heads-up technique while the speed and ease of microscopic manipulations were considered to be similar compared to the conventional method [25]. In clinical routine, usage of 4K large three-dimensional displays with the benefit of digital image processing has the potential of becoming a game changer.

Some microscopes have displays integrated into their eyepiece that can be used for overlaying information from secondary imaging modalities or preoperative data. Optical see-through displays in the eye-piece of the microscope eliminates many hassles associated with geometric calibration. These displays can be geometrically calibrated to the microscope focal field during the manufacturing time. Augmented reality solutions based on eye-piece optical see-through displays are among the most feasible medical AR solutions.

### 1.3.4 Optical Coherence Tomography (OCT)

Optical Coherence Tomography (OCT) is a tomographic imaging technique that in its basic form is based on time-coherence interferometry. In standard OCT scheme, a low time-coherence light source is used in a standard Michelson interferometer [32]. As depicted in Fig. 1.10, a Michelson interferometer consists of a light source, a beam splitter and a reference mirror. The light source being short in coherence length is key in having the ability to sense a sample. The system is used to capture optical measurements from a sample that are recorded on an optical detector. To do so, a source beam is passed through the beam splitter which divides it into two coherent beams. One beam called the reference beam gets reflected by the reference mirror back to the splitter and directly to the detector and the other called the sample beam passes through the sample. Scattered photons from the sample that fall back to the beam splitter are also redirected to the detector. The optical detector captures the interference pattern formed by the two beams. The interference pattern is only captured from the photons coming from the sample depth that is as far from the beam splitter that the reference mirror is from the beam splitter. Hence, only a single point in the sample is optically measured at a time. The measurements can be turned into a brightness intensity value showing the variation of tissue materials and boundaries of the measured points. By moving the reference mirror closer or further away from the beam splitter, it is possible to capture a line of measurements along the depth of the sample. Such measurements similar to ultrasound imaging are called Amplitude Scans (A-Scans). By moving the sample in XY direction or by deflecting the sample beam using added mirrors and other optical elements, it is possible to form two-dimensional images called Brightness Scans (B-scans) and three-dimensional volumes. This means that the resolution of OCT in axial and lateral directions are decoupled and independent. Furthermore, OCT essentially is a non-contact sensing method that uses only light as a means of probing samples. If the wavelength of the light source is chosen to be in higher than ultraviolet, then it could be used as a non-invasive modality for imaging tissues. Typically near-infrared is used for medical purposes. This choice comes from the need to operate in a spectral range in which the penetration of light into tissue is adequate. Because of the short mean scattering length of photons in tissue at wavelengths in the blue and ultraviolet, OCT imaging with light sources that emit in these spectral regions would
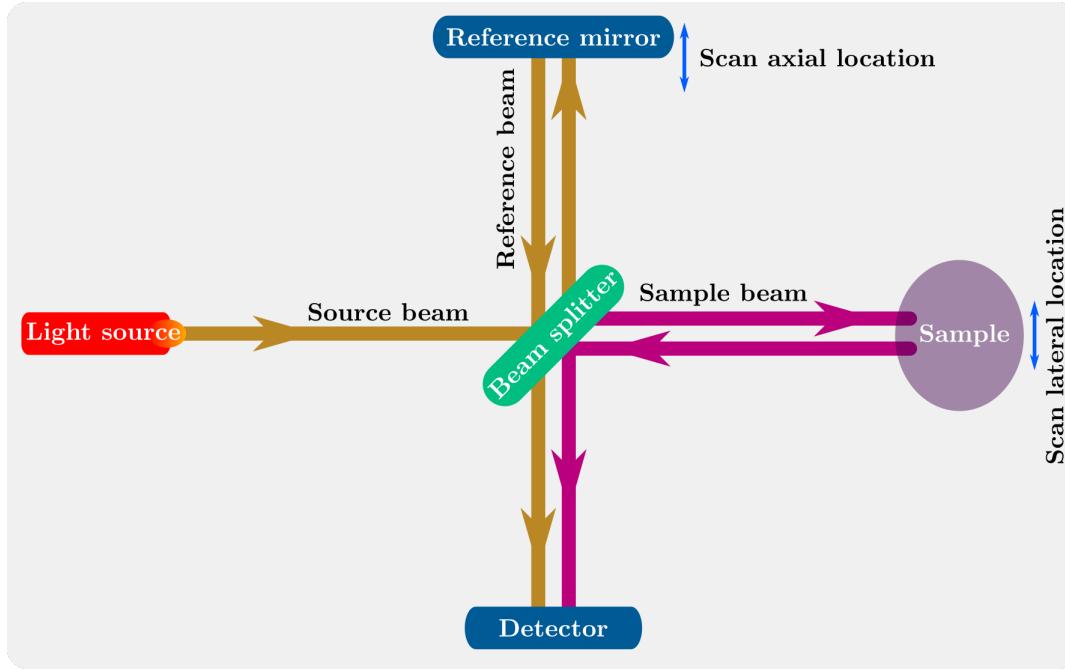
**Fig. 1.10.** Basic OCT scheme. The beam splitter splits the incoming source beam into reference and sample beams. The photons coming back to the beam splitter are then redirected to the detector.

be limited to superficial layers less than a few hundred micrometers thick. At wavelengths greater than 2500 nm, vibrational absorption by water limits imaging to similar depths. In studies carried out to date, the deepest penetration has been achieved using sources that emit at wavelengths between 1200 and 1800 nm [78]. Since optical coherence tomography is limited to imaging tissues that are semi-transparent, ophthalmology has been and continues to be the main medical application for OCT.

The optical coherence tomography concept depicted in Fig. 1.10 is Time-Domain OCT (TD-OCT). In time-domain OCT, depth of the tissue is scanned using time delays. The basic principle behind imaging using time-domain optical coherence tomography by use of the interference law and introducing a time delay $\tau$ could be mathematically described [32] as

$$\overline{I_D}(t,\tau) = \langle I_S(t) \rangle + \langle I_R(t) \rangle + 2\sqrt{\langle I_S(t) \rangle \langle I_R(t) \rangle}\ |\gamma_{SR}(\tau)|\ \cos(\alpha_{SR} - \delta_{SR}(\tau)), \qquad (1.1)$$

where $\overline{I_D}$ is the average intensity read at the detector, $I_S$ and $I_R$ are the intensity of the two interfering sample and reference beams, $|\gamma_{SR}(\tau)|$ is the degree of coherence between source and reference beams, $\alpha_{SR}$ is a constant phase, $\delta_{SR}(\tau)$ is the phase delay, $\tau$ is the time delay and the angle brackets are ensemble averages. The time delay is defined as

$$\tau = (\Delta z/c), \qquad (1.2)$$

where $c$ is the speed of light and $\Delta z$ is the length difference between the beams correlated with the position of the reference mirror. For modeling the detector intensity when dense tissues with different levels of attenuation and backscattering are probed, modified versions of the introduced mathematical terms should be used. Some tissue scattering models express the interference signal as a convolution. In the more general models in which a two-dimensional

cross-sectional image is built from multiple axial scans the optical transfer function of the optics that scan the beam and focus it into the tissue must be considered [78].

A distinct characteristic of OCT B-scans is the prominent speckle noise. The insidious noise comes in form of highlight points degrading the quality of OCT images. Speckle noise reduces the contrast of OCT images and makes the tissue structures and boundaries hard to resolve. The origin of the speckle noise could be traced to two major sources. The primary source is the multiple backscattering of the beam inside and outside of the desired sample volume. Although the OCT sample beam is focused into a small region of interest for imaging, the backscattering of the beam happens from multiple spots in that region and even other spots that are illuminated via secondary illumination. The secondary source is random delays of the forward-propagating and returning beam that is inherent to any imaging modality that employs penetrating waves. Since there are multiple and random scatterers in the sample volume, backscatter waves that reach the same point on the detector could come from different forward-propagating beams, out of phase and within an interval of time less than the coherence time of the source [79]. There are also multiple minor sources of speckle noise such as the motion of the sample volume or the detector, reflectivity of the reference mirror, etc. Since there are many sources of speckle noise, finding a statistical distribution describing the noise occurrence and directly using it for noise suppression is not feasible. There has been a great interest and effort in OCT noise suppression in the science community leading to a multitude of different methods such as machine learning approaches, wavelet or curvelet thresholding, non-local means filters, anisotropic diffusion and many more [18].
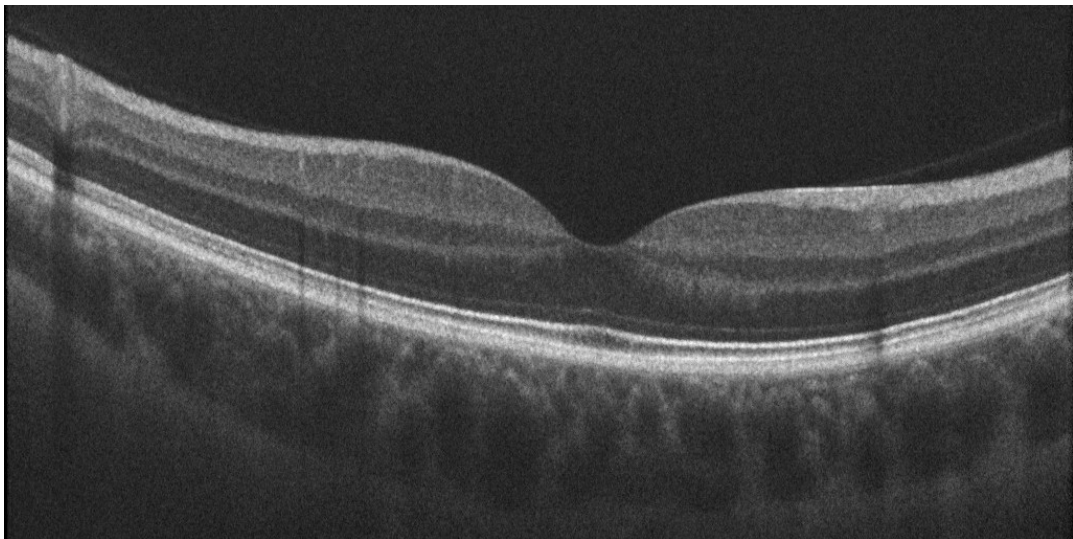


**Fig. 1.11.** *In vivo* spectral domain OCT image from a human retina. Retinal layers are distinctively visible in OCT images.

Spectral domain OCT (SD-OCT) is a superior imaging method that borrows from spectroscopy techniques. The basics of imaging in SD-OCT is similar to TD-OCT however instead of a detector an array of detectors is used to detect the spectral variation of the incoming signal. The illumination source used in SD-OCT has a broad spectral bandwidth. By using an optical dispersive element, the incoming sample beam is distributed into a photo detector array similar to a camera charge coupled device (CCD). Since the information from various depths is

encoded into the spectral domain, an inverse Fourier transform of the spectral measurements will produce a line scan profile similar to that obtained from TD-OCT [98]. One of the most important advantages of introducing Fourier domain detection to OCT imaging is the substantial improvement in the quality of cross-sectional images acquired at high speeds. The increase in axial resolution could be obtained with SD-OCT because of the lack of fundamental relationship between imaging sensitivity and axial resolution in the case of Fourier domain detection [19].

Time encoded, or more commonly known as Swept Source OCT (SS-OCT) is another method of OCT imaging. In SS-OCT, the illumination source has narrow instantaneous line width but is rapidly swept in wavelength and the spectral interference pattern is detected on a single or small number of photoreceivers as a function of time [23]. Hence, the spectral components are not encoded by spatial separation but are encoded in time. The entire depth-resolved structure of the sample at the position of the focal spot is encoded in the spectral interference pattern and its spectral frequency content. The simplified mechanism used in SS-OCT contributes to high-speed data acquisition that is twice as fast as that achieved by SD-OCT and it results in a clearer image. The depth of tissue that OCT is capable of imaging is governed by the wavelength of the light source used. The median wavelength of the light source in typical SD-OCT devices is approximately 840 nm. In SS-OCT, the median wavelength is commonly 1050 nm. Swept source optical coherence tomography has enabled the visualization of the whole thickness of the choroid and structures beneath a retinal hemorrhage or the retinal pigment epithelium (RPE) [45].
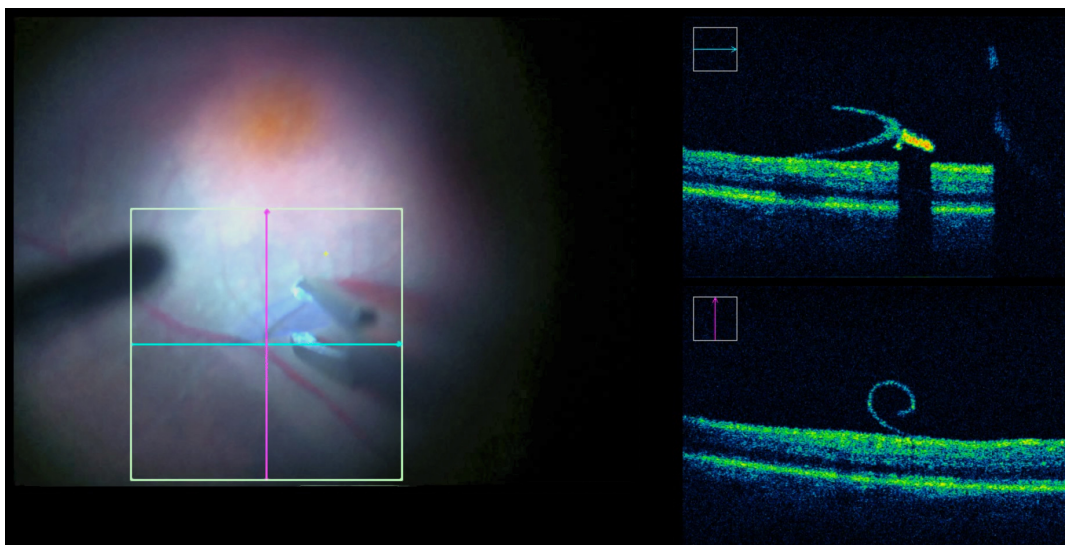


**Fig. 1.12.** Application of intraoperative OCT during an epiretinal membrane peeling procedure. The images on the right side are OCT B-scans captured at locations marked with cyan and magenta arrows. Note the rolled membrane and the total shadowing effect under the instrument visible in OCT.

OCT has become the imaging modality that is widely used in ophthalmology. By improvements regarding the hardware required for OCT and evidence supporting the utility, OCT has found its way into the operating rooms. Intraoperative OCT first came about as handheld devices that could be used from time to time during the surgery when need be. However the feasibility of using a handheld device during a fast-paced microscopic operation is questionable. With

the advent of surgical microscopes equipped with OCT capture capabilities, real-time and continuous OCT acquisition is a reality. During surgical maneuvers in anterior segment and posterior segment surgery, intraoperative OCT can provide rapid visualization of the area of interest and can provide the surgeon with information regarding instrument-tissue interactions [30]. Using OCT during ophthalmic interventions has the potential of introducing major improvements in surgical outcomes. However, obstacles in realization of that are the fact that handling intraoperative OCT devices are difficult, the information they provide is hard to rapidly interpret and the way that information is presented is irrelevant to the surgical flow. In recent years two major clinical studies, PIONEER [28] and DISCOVER [29] have shown demonstrated the feasibility of real-time intraoperative OCT with a microscope-integrated OCT system for ophthalmic surgery. But time burden and difficulties associated with image acquisition were also considerable.
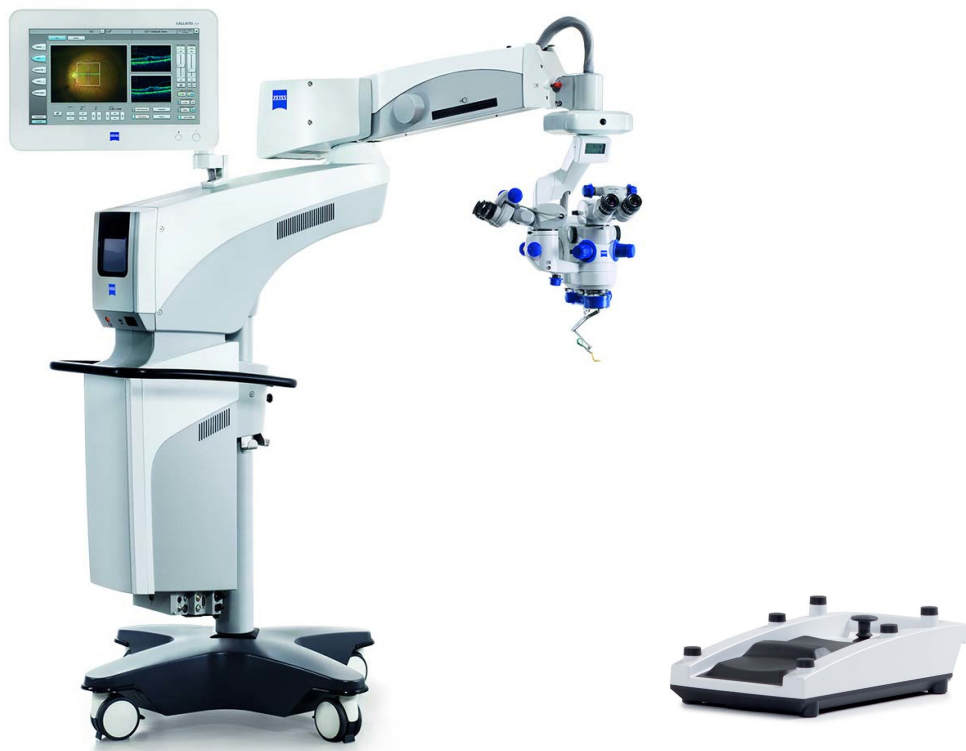


**Fig. 1.13.** Zeiss OPMI Lumera 700 with RESCAN 700 surgical microscope equipped with intraoperative OCT. The foot control pedal is used by the surgeon for manipulation of the microscope and OCT parameters.

One example of a surgical microscope with integrated OCT functionality is OPMI LUMERA 700 with RESCAN 700 from Carl Zeiss Meditec. RESCAN 700 is a spectral domain OCT module that is fully integrated into the standard LUMERA 700 ophthalmic microscope. It comes with a foot control pedal for manipulation of the OCT parameters such as the capture location. Using RESCAN 700 multiple OCT B-scans can be continuously captured and visualized from the surgical scene or dense three-dimensional volumes can be captured in offline mode. OCT acquisitions are visualized on a separate assistant computer attached to the microscope. It is also possible to visualize OCT acquisitions during the operation using overlay techniques. One of the ocular lenses of the microscope is equipped with a display that can overlay images for

the surgeon. Hence, the OCT is visible over the surgical scene close to the location where it is captured. The presence of such display is a perfect opportunity for using augmented reality for providing information to the surgeon.

# Augmented Reality (AR) for Ophthalmic Procedures

## 2.1 Motivation and Challenges

Using augmented reality for medical purposes and specifically surgical interventions has been an interest in the research community since a few decades. Arguably the primary motivation behind the push for augmented reality systems has been the need of visualizing medical data and the patient within the same physical space [83]. With the recent advances in medical imaging modalities, miniaturization of imaging devices and advent of systems capable of rapid acquisition, more and more data is generated in the operating rooms. On the other hand, the progress in less invasive and faster procedures, not only justifies the need for more intraoperative online data, but also requires new methodologies for transferring that data to the surgical staff.

In surgical interventions such as ophthalmology, the benefits of augmented reality systems are more apparent. Secondary modalities such as optical coherence tomography are becoming commonplace in ophthalmic operating rooms. Online continuous data acquisition using OCT can benefit the surgeon by continuously providing information regarding instrument-tissue interactions and manipulation results. OCT could also be used for navigation towards locations marked preoperatively. However, the current techniques used for providing OCT images to the surgical staff is neither convenient nor efficient. Since ophthalmic procedures are performed under a surgical microscope or via heads-up displays, the view of the surgeon can more easily be manipulated compared to conventional open surgeries. Many operating microscopes already incorporate fully calibrated ocular displays that overlay OCT images to the surgical scene. Using the available technology, it is possible to push the limits of what is the current gold standard for information conveyance in ophthalmic operating rooms. Augmented reality could be used to provide information that is extracted from multiple modalities or views in a way that requires the least effort from the surgeon to interpret.

Another motivation behind employing augmented reality in surgery and specially in microscopic surgery is in preserving the surgeon's focus. While more information could benefit surgeons in decision making, it could also be a distraction from the surgical field. This point becomes more apparent when circumstances of a microscopic surgery is considered. The center of attention in a microscopic surgery is a few millimeters of tissue where the manipulation occurs via manual movements. Depth perception under a surgical microscope is to some extent compromised and making sure instrument movements are precise and concise is difficult. Under such conditions paying attention to extra information from various sources including secondary imaging modalities and trying to interpret the provided information quickly becomes a burden. Utilizing augmented reality has the benefit of providing an

immersive environment where the surgeon subconsciously receives the necessary information with the least amount of mental effort. AR has the potential to seamlessly integrate into a microscopic surgical workflow with the least amount of alteration. Keeping the focus of the surgeon where it should be is achievable with proper AR system designs.

Although AR is a feasible method of choice for facilitating complexities associated with ophthalmic microscopic surgeries, it does not come without challenges and complexities of its own. Proposition of any AR solution should be complemented with mechanisms for suppression of complications affiliated with it. One of the notable challenges of AR in microsurgery is the accuracy requirements. Since common surgical tasks performed under microscopes are inherently precise and limited to a small space, AR methods proposed for such tasks are also required to have the same amount of precision. For an ophthalmic procedure the required accuracy for an AR solution is in the order of tens of microns based on what is known about instrument-tissue interactions. Displays used for presenting excess information are required to be precisely calibrated to the focal plane of the microscope at all times. Maintaining depth perception in proposed solutions could be a challenge specially when monocular displays are used. Typical phases of surgery in an ophthalmic intervention are only a few minutes. Hence, the efficiency of conveying information in such short time cycles is crucial. Methods of choice for providing fused multimodal data are required to be compatible with the procedure they are trying to support. Feasibility of a proposed solution under unpredictable situations is also of considerable importance. Even though ophthalmic procedures and the methods involved are to some extent standard, there are always unpredictable situations where ordinary treatment of the data is not effective. Last but not least, AR methods that rely solely on enriching the surgical environment via visual methods are limiting. Vision is the primary source of information for an ophthalmic surgeon. Interfering with a surgeon's vision could potentially be troublesome.

Even when an effective AR solution is proposed, there are several obstacles that can interfere with successful deployment of such solutions in the real world. Firstly, proper evaluation of the efficacy and precision of an AR system is a complex task. Bringing about the conditions of an ophthalmic surgery where it is possible to adequately test a proposed solution is a challenge. When a solution has shown its merits, active and appropriate tutorial of the surgical staff is central to the success of the proposed system. In most cases it is necessary to design the solution in such a way that the minimum amount of training is sufficient for the surgical staff to be able to interact with the system. It should be noted that adoption of medical AR solutions require a cultural change in the medical and computer science community. Education of medical staff using AR solutions that adhere to their basic needs is a compelling way to further integrate AR into the surgical theater.

Although gaining the mentioned benefits of augmented reality for ophthalmic procedures while overcoming the obstacles may seem hard to achieve, the author believes AR is the evident method of choice for attacking the current complexities of ophthalmic procedures. This is specially true when it comes to certain ophthalmic procedures where the availability of intraoperative OCT is believed to be revolutionary but the current trends in employing intraoperative OCT adds so much complexity to interventions that its potency is compromised.

## 2.2  Applications

Augmented reality could be employed in various procedures with different levels of impact. However surgical procedures that benefit the most from AR are the ones with information from peripheral sources having a pivotal role in the outcome. In ophthalmic interventions OCT is a peripheral source of information. Assistive computer algorithms with the ability to extract information from video streams or other available data streams are also peripheral sources of information. Presentation of information in a straightforward manner is in this context inferior to AR since it requires the surgeon to deliberately pay attention to subjects away from the surgical scene. For the same reason, it is preferable to extract the most relevant information form the available ones and present them in a minimalistic fashion. During the course of this research project, the following set of ophthalmic procedures have proven to be the most benefiting from AR.

Vitrectomy is the basis for many posterior segment procedures. Although a vitrectomy procedure is considered a routine intervention that typically takes only a few minutes, it is not devoid of challenges. The vitrectomy probe used for this procedure not only cuts the protein fibers in the vitreous humor into smaller pieces, it also aspirates the humor out of the eye, while the irrigation channel provides substitute fluids back into the eye. The surgeon hence regulates the pressure inside the eye by balancing the irrigation and vacuum rates. This is typically done by foot movements on a control pedal. At the same time, the vitrectomy probe should be moved in a circular motion to cover the intended area while kept away from the retina to avoid unintentional touches or tear damage. Perceiving depth under a ill-lit hazy visual field while controlling other parameters of the system via foot is laborious. OCT can potentially help with depth perception. However not only controlling the scan location requires the surgeon to use a second foot control pedal, it also does not provide much information below the vitrectomy probe because of the total shadowing effect caused by the metallic instrument. Providing OCT images on a second display can also be distracting and troublesome. A well-designed AR system can extract depth information from OCT images and provide them in such a way that does not affect the normal workflow of the procedure.

Epiretinal membrane peeling is similarly affected by problems associated with depth perception. In a common peeling procedure a pair of forceps is used for grabbing and removing of the epiretinal membrane. Unnecessary touch of the retina by the forceps tips can cause microhemorrhages leading to loss of vision. Information regarding the distance of the instrument to the retina can be continuously extracted from OCT images captured from the surgical scene. Augmented reality has the potential to provide such peripheral information with minimum distraction using overlays over OCT images.

Subretinal injection procedure can benefit from AR since one of the major challenges associated with it is the proper placement of the needle tip in the appropriate retinal layer. guidance and navigation towards the depth within the accepted tolerance of the choice can be done via AR techniques. The role of intraoperative OCT in guidance towards the desired depth is evident. However, since surgeons rely on their visual feedback of the surgical field to determine the force they are applying on the needle [38], it is necessary to provide information from OCT without obstructing the surgical view. Ophthalmic surgeons are exposed to the concept of

guidance via sound as a result of devices that convey information regarding surgical parameters via sound. As an example, phacoemulsification devices provide information regarding the intraocular pressure of the eye with sound tones. Hence, auditory augmented reality is a good candidate to be used when guidance is in mind. Information captured in OCT images could be extracted and provided aurally so that the surgical view is not obstructed.

Undoubtedly both visual and auditory augmented reality has applications other than the ones mentioned in ophthalmic procedures. One could think of interventions such as deep anterior lamellar keratoplasty where similarities between the air injection phase with subretinal injection makes AR a good facilitator. AR solutions proposed in this work could potentially be applied to similar procedures in ophthalmology and beyond.

## 2.3 Related Work

Augmented reality for medical interventions has been tried out in many clinical studies and is a part of limited operating rooms around the world. But when it comes to ophthalmology, medical AR is neither prominently found in the surgical theater nor yet strongly studied in the scientific literature. However, there are several studies on related topics such as the application of augmented reality in microscopic surgery and training in ophthalmology. Here, the work in the scientific community that is most relevant to the topic under discussion is presented. Since the techniques proposed in this work are based on visual and auditory enrichment of the surgical environment, these two methods are covered separately.

### 2.3.1 Visual AR

In the report by Shuhaiber [82] of the history and current knowledge of augmented reality in the field of surgery in 2004, there is no mention of AR in ophthalmology. The state of AR at the beginning of the millennium according to that report is in a preliminary stage with the lack of understanding of the requirements in registration and ergonomics keeping it limited. The closest application to ophthalmology is neurosurgery with interactive image-guided surgery being the main focus with heads-up displays in an outside of surgical microscopes. Edwards *et al.* in [26] present one of the early examples of stereo augmented reality systems for a surgical microscope that injects overlays from preoperative images into the field of view of the microscope. In such a system, the fundamental elements are considered to be the calibration to establish correspondence between the three-dimensional world and the two-dimensional image overlays, the segmentation of the preoperative image data to determine the structure of the interest for the surgeon, the registration to determine the transformation that relates the preoperative image data to the patient, the tracking that keeps the transformations in previous steps relevant even though all the elements involved in the operating room are mobile, and finally the visualization that consists of the set of algorithms that render the final overlays and the hardware that supports the rendering. Errors in all of these elements and 3D perception are considered as issues of such systems.

Berger *et al.* in [7] initiates an study towards the design and implementation of an ophthalmic augmented reality environment. This investigation is the first of its kind for ophthalmology.

The engineering considerations laid out in this research are practical methods for overlaying information from one imaging modality in ophthalmology onto another imaging modality with robustness against moderate changes in eye position and illumination. The similarity metric proposed for registration of anatomical targets such as vessels extracted from monochromatic or angiographic images onto fundus images is the Hausdorff distance. To cover a larger area to be robust against eye movements, multiple images from each modality are montaged using the proposed similarity metric. Finally, a Zeiss operating microscope with image injection capabilities is used to overlay the scene with the computed feature image in real-time. In [8] the authors use similar techniques to overlay stored diagnostic data onto the real-time slit lamp fundus view. Previously acquired fundus photographs and angiography images are digitized. A slit lamp image acquisition device equipped with a CCD camera, a frame grabber and a computer allows for real-time acquisition and digitization of slit lamp fundus images that are synchronous with posterior segment examination. The proposed method also includes a custom-developed video injector that uses a miniature cathode ray tube display allowed for real-time superposition of angiographic images to the fundus view. Registration and tracking algorithms are in place to compensate for movements. The feasibility of this approach is demonstrated in 5 human subjects. The computer vision algorithms provide robust tracking, registration and image overlay of previously stored photographic and angiographic images directly onto the real-time fundus view. Accurate tracking is demonstrated with updates at 3 to 5 Hz. Such a system could guide treatment for diseases related to the macular region by providing fused information that is not visible in fundus images but are available in processed images such as eye angiograms. Providing overlays using augmented reality for this application is superior to conventional methods due to the accuracy requirements.
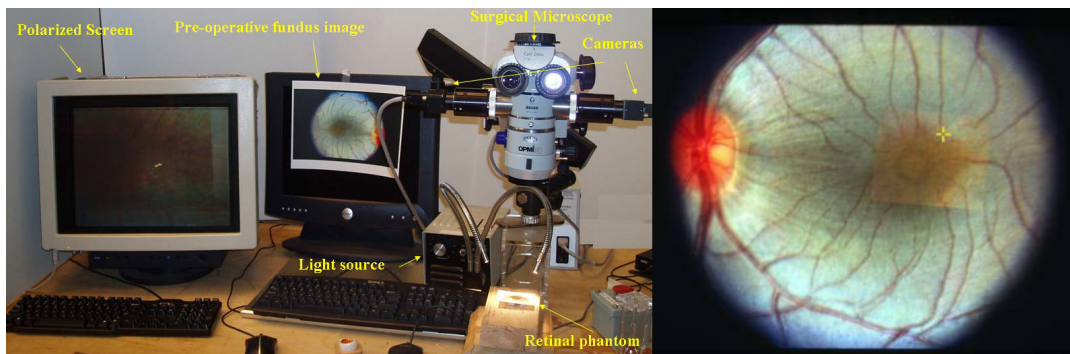


*Source: [33] I. N. Fleming et al., 2008 IEEE Workshop on Applications of Computer Vision, © 2008 IEEE.*

**Fig. 2.1.** Augmented reality example for fundus / microscope fusion. Experimental system setup (left) and overlay of a fundus image with a microscopic image with an anatomical target marked in yellow (right).

Fleming *et al.* in [33] propose a solution for providing interventional targets using preoperative processed images applicable for surgical interventions such as membrane peeling. The key element of their solution in accordance to similar methods is augmenting the surgeon's view with additional information with tracking and registration in place. Their proposal consists of three main steps. First, optical coherence tomography images of patients with fundus images are acquired using a diagnostic device. The acquired OCT and fundus images are aligned and registered using the dual bootstrap iterative closest point algorithm [85]. For that, one or more pathological targets in the OCT en face image and the fundus image are marked by the user of the system. In the second step, the preoperative fundus image marked with

pathological targets is registered to the intraoperative fundus view of the surgical microscope. A region-based tracking algorithm is used to compensate for the movement of the anatomical targets in stereo. The final microscopic view together with target overlay is displayed to the surgeon in 3D using a polarizing screen. The proposed system has the potential to improve the vitreoretinal surgical experience by improving the positioning accuracy of instruments.

To evaluate the efficiency of providing intraoperative guidance using vision-based methods, Palma *et al.* have designed and performed a set of experiments [61] in which manual high-precision tasks are performed by novices and an experienced vitreoretinal surgeon. The designed task performed by the subjects is to trace a circle $500\mu m$ in diameter three times using a hand-held micromanipulator while holding the tip of the instrument $500\mu m$ above a rubber surface. The planar error in tracing is provided to the user as feedback via visual cues. These visual cues are presented using a two-dimensional monitor, a three-dimensional monitor or a custom-built microscope image injection mechanism using a monocular display. Results from their study shows that both for novice users and experienced surgeons providing guidance via augmented reality significantly reduces error compared to monitors ($p < 0.05$). These preliminary results suggest that visual cues displayed via monocular augmented reality display within the surgical microscope enable more accurate micromanipulation than visual cues displayed on 2D or 3D monitors.

In another study [10] by Bhadri *et al.*, relevant to digital ophthalmic systems that are AR-ready, a stereoscopic camera-based three-dimensional viewing workstation for ophthalmic surgery is evaluated. For that, surgeons are engaged in anterior and posterior segment trial surgeries which involved designated standardized tasks on porcine eyes after training on prosthetic plastic eyes. The results from that study suggests that three- dimensional systems may be a viewing system in ophthalmic surgery with improved ergonomics with respect to traditional microscopic viewing. Such systems inherently support AR solutions by providing a digital representation of the surgical field. Since they provide better ergonomics for the surgeon, it is assumed that future surgical solutions are tailored towards possessing AR capabilities.

A study by Leitritz *et al.* [50] is conducted to evaluate the value and usability of augmented reality for training of inexperienced ophthalmology examiners. In this study, thirty seven medical students are randomly assigned to undergo training of binocular indirect ophthalmoscopy either using augmented reality or in a conventional way. To evaluate their training performance, they answer questionnaires and examine a real person using conventional ophthalmoscopy afterwards. Results from the questionnaires show no differences in performance. However, the median ophthalmoscopy training score for the conventional ophthalmoscopy group is 1.2 and shows a significant difference ($P < 0.0033$) to the group undergoing training using augmented reality methods with median of 2. The superior performance of AR in training demonstrated in this study indicates that augmented reality not only is suitable for providing intraoperative guidance, it could be utilized for the training of the next generation of ophthalmic specialists.

With optical coherence tomography in the operating rooms, overlaying of targets detected in OCT on the microscopic view is a natural next step in AR solutions. Balicki *et al.* have proposed a guidance solution [4] for epiretinal membrane peeling by localizing difficult to identify anatomical features on the retina using video stereo-microscopy and intraocular OCT.

Their solution consists of a Zeiss surgical microscope equipped with two cameras covering the left and right light pathways and a 3D monitor with shutter glasses for visualization. A 4 Degrees of Freedom (DOF) transformation between the current field of view and an internal planar map of the retina is continuously estimated. Templates of $30 \times 30$ pixels are matched with Normalized Cross Correlation (NCC) as a similarity metric for compensating the motion of the retina. For tracking OCT, a near-infrared camera is used to determine the location of the scan. For tracking the instrument an external algorithm is used. The tracking performance is evaluated in a realistic water-filled eye phantom. Using the system, a surgeon can identify targets found in the OCT image on the surface of the retina with the accuracy of around $100\mu m \pm 100\mu m$. Systems developed based on extraction of anatomical features in OCT and presenting them on microscopic views using augmented reality methods are helpful for surgical procedures that involve manipulation of tissues that are visible in OCT but not in the microscopic view. Registration and tracking errors are typical bottlenecks of these systems.
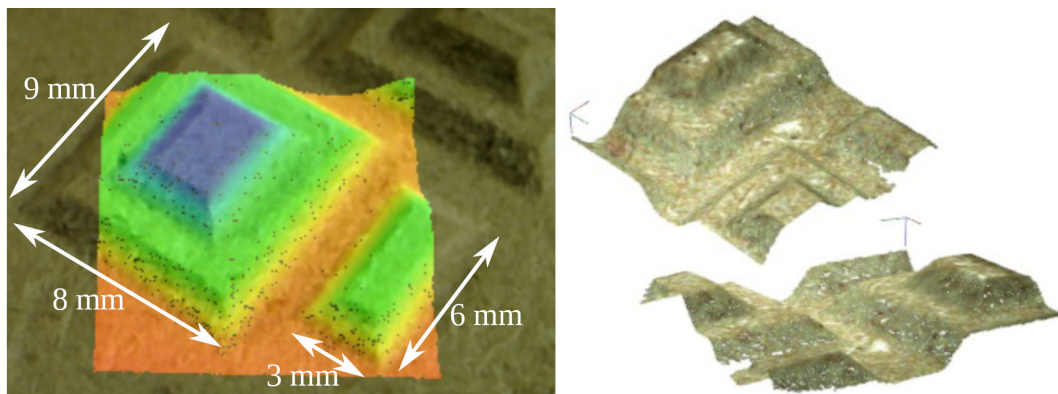
**Fig. 2.2.** Augmented reality example for OCT / microscope fusion. Overlay of a colored depth map onto the camera image (left) and textured OCT point cloud from two different perspectives (right).

In another study related to optical coherence tomography and stereo camera images, Bergmeier *et al.* investigate [9] a system for combining OCT with microscopic view. OCT provides superior depth perception compared to stereo microscopy. However, OCT does not provide the surgeon with a familiar view of the surgical scene. Hence, the aim of their work is to process OCT images and fuse them with the stereo microscopic view in a way that is perceptually better than conventional methods. In their solution, surface reconstruction is performed on the stereo image feed to for a 3D model of the tissue surface. Surface information is also extracted from denoised OCT volumes. These two three-dimensional models are then registered together using the Iterative Closest Point (ICP) method. In one approach, depth information extracted from OCT volumes is then overlaid into the microscopic view using color-coding techniques. In another approach texture information from the microscopic view is projected into the OCT point cloud. Their investigation is a basis for integration of OCT data into stereo imaging. Measurement and monitoring of surgical maneuvers and depth perception could be improved with such a system. On the other hand, processing OCT and video information based on detection of surface structures has the downside of negligence towards subcutaneous structures. In applications such as subretinal injection, benefits of having OCT when only depth information from surface structures is presented is not realized.

## 2.3.2 Auditory AR

When it comes to data display in an operating room, there are different methodologies with advantages and shortcoming that could be employed for different applications. In a review [11] by Bitterman, various technologies and displays are described and compared. According to the review, most of the physiological information presented in operating rooms is visual and displayed on computer screens and the most popular and almost the only acoustic display available in any clinical setting is the pulse oximeter which supplies continuous information about heart rate and arterial hemoglobin saturation. Sonification of data which is defined as the use of non-speech sound to convey information is developed with the aim of attracting the attention of the surgical staff only when changes in the physiological parameters happen in contrast to auditory alarms. Auditory icons which are non-speech sounds that have immediate natural associations with a state or object and "Earcons" which are sounds that are learned to be associated to phenomena are also techniques for conveying information to the surgical staff. However, unlike sonification, auditory icons and earcons require focused attention to be interpreted. Hence, simple alarm sounds, auditory icons, earcons and intricate sonification techniques all have benefits of their own and should be used in accordance to their merits. Auditory displays as opposed to visual displays are omnidirectional and therefore immediately available to the staff around the operating room with no line of sight problem. Auditory displays are not space-restricted and are advantageous in equipment-congested environments such as operating rooms. Among the downsides of auditory displays are difficulties with presenting some data streams that are predominantly perceived visually and also the fact that background noise can disrupt information transmission. In a survey [13] by Black *et al.*, auditory display in image-guided interventions are investigated. Covered auditory display methodologies include alerts, auditory icons and parameter mapping with sonification. Results from the reviewed literature show that in most cases systems with auditory display are found to be beneficial. Advantages include improved recognition of the presence of or distance to anatomical risk structures, reduced complication rate, improved placement accuracy, improved resection volume similarity, improved orientation and reduced workload. Drawbacks seen in the reviewed literature include increased task time and increased amount of non-target tissue removed during volumetric resection. None of the reviewed literature report negative subjective perception of the implemented auditory display. However, despite apparent benefits of augmenting or replacing certain aspects of image-guided interventions with sound information, investigations have been sparse and there is a need for intensified development and comprehensive evaluation of novel auditory displays that reach beyond simple alerts and alarms.

One of the early applications for auditory displays in surgery is the proposal by Vickers *et al.* [91] in which audio is proposed for controlling and monitoring of the equipment in minimally invasive surgeries (laparoscopy). In such surgical sessions, the surgeon focuses on the visual feedback from the video control units providing the surgical view. Hence, showing extra information visually is either obstructing the field of view of the surgeon or is redirecting the focus of the surgeon from the surgical scene. Their proposed solution to avoid these problems is research and development of control and display systems for laparoscopic equipment with surgeon being able to control the parameters of the equipment using speech and monitor the parameters using auditory renderings. Similar to laparoscopy, surgeries performed under a surgical microscope can also benefit from auditory displays. In [93], Wang

*et al.* propose a solution for micromanipulation training using multisensory cues. Audio signals are used to transmit information about trajectory error relative to a desired movement or contact with virtual objects. Their solution provides a more immersive experience for training purposes.

For another application, Giller *et al.* in [37] have proposed a system for sonification of Electroencephalogram (EEG) signal for epilepsy surgery. Their method includes several transformation of the signal from EEG recordings into the human hearing range which could be reviewed by neurosurgeons. Short segments of high pitched, high amplitude signals are frequent in tissue that eventually developed seizures but are rare in normal tissue when their method is applied to EEG recordings coming from contacts placed directly on the patient brain. This method of sonification is a tool for the surgeon to perceive information that is under conventional circumstances only available as visual cues in complex data.

A common practical use case for sonification in surgery is navigation for reaching a certain target or avoiding certain structures. Dixon *et al.* have studied the role of auditory icons as proximity alerts in endoscopic surgery [21]. They conclude that the development of auditory icons for proximity alerts during their trials better informed the surgeon while limiting distraction. In case of neurosurgery, Plazak *et al.* in [64] have developed a navigation system that sonifies distance information between a surgical probe and the location of the anatomy of interest. Their auditory augmented reality system consists of five different types of sonification: Sine tone frequency matching, Sine tone pitch mapping, Pulsed tone sonification, Signal-to-noise sonification and Binaural beat sonification. In their controlled experiment, 15 participants are tasked with navigation of a surgical tool to randomly placed three-dimensional locations. Their performance is measured when they are provided with visual only, audiovisual or audio only guidance. The conclusion of the study is that audiovisual feedback results in greater accuracy compared to visual only. Combining auditory distance cues with existing visual information may result in greater accuracy when locating a given target in a 3D volume. Combining auditory and visual information also reduces the perceived difficulty in locating a target within a 3D volume.

One of the best demonstrations of the benefits of auditory augmented reality for navigated surgery is the work by Hansen *et al.* [40]. They propose an auditory display system for open liver surgery with support for guiding the tracked instrument towards and remaining on a predefined resection line. The instrument in the hand of the surgeon is tracked by a commercial navigation system. The distance between the surgical instrument tip and the nearest point on the planned resection line is the input to the function that determines the auditory feedback. The area around the resection line is divided into safe, warning and outside regions and the location of the instrument tip relative to these regions is associated with different tones that provide feedback. A combination of different sound manipulation techniques including change in pitch, tone length and inter-onset interval are incorporated in the tone generating function used for feedback. The clinically oriented tests that are performed in the context of this work revealed that auditory feedback could be a beneficial extension to surgical navigation systems. The auditory display significantly reduces the time that surgeons looked on the navigation system screen while reducing the mean distance to the planned resection line. Among the downsides of the proposed solution are the lack of aesthetically pleasing sounds that could be tolerated to hear over a time range long enough for surgical

tasks and the more time it takes for subjects to complete a resection task when auditory feedback is available. However, one contributing factor for the increase in task completion time could be that the auditory display alerted the surgeon when leaving the safe margin. This auditory notification may have caused a more cautious and accurate marking of the resection line than without having this information. A second contributing factor could be the lack of extensive training. All in all, the proposed method demonstrates the benefits of auditory AR guidance systems in navigates surgery.

**Fig. 2.3.** Simulated liver surgery with auditory feedback. The participating surgeon uses the tracked instrument to mark the planned resection line on the liver phantom. The surgeon mostly observes the liver phantom when auditory feedback is available (left) and repeatedly glances at the navigation system screen when no audio feedback is present (right).

Black *et al.* in [12] present a novel auditory display mechanism for interventional procedures that contributes towards increasing awareness of the uncertainty present in numerical navigation information. Information coming from navigation systems is uncertain due to various sources of potential errors, including from soft-tissue motion estimation processes and instrument tracking hardware. In the proposed method, mean distance to target is mapped to the sound's stereo position. Positive distances are heard in the left ear (target left of instrument) and negative distances in the right ear (target right of instrument). When the instrument reaches the target, the sound is heard centered in both ears. For encoding the variance, two methods are proposed based on different psychoacoustic principles. The first method maps the variance to the frequency modulation of the base sound by adding a vibrato component. The second method uses white noise audio signal to encode the variance. The preliminary qualitative evaluation of the proposed auditory display system with medical imaging professionals shows that users are able to perceive the level of uncertainty present in the navigation information. Ideally, this could increase the safety of critical navigation procedures, particularly since information conveyed via auditory displays does not require the need for shifting the gaze away from the patient. Needle placement is another common application for auditory augmented reality to provide continuous feedback. Black *et al.* in [14] propose an auditory synthesis model using pitch comparison and stereo panning parameter mapping to augment or replace visual feedback for navigated needle placement. With their system, using combined audiovisual display, participants show similar task completion times and report similar subjective workload and accuracy while viewing the screen less compared to

using the conventional visual method. The auditory feedback leads to higher task completion times and subjective workload compared to both combined and visual feedback. Auditory AR can be a supplement to visual AR as an assisting tool.

## 2.4 Contributions

Based on the previous work in the direction of employing augmented reality for surgical procedures, there seems to be more effort into using AR in conventional open surgery, laparoscopic surgery and neurosurgery compared to ophthalmic microscopic procedures. Although the groundwork for employing AR in ophthalmic procedures has been laid out in the previous decade, utilization of such systems in clinical settings are substantially less compared to other surgical interventions. Part of this absence in the opinion of the author could be attributed to the fact that a considerable number of AR proposals in ophthalmology are translations of AR from other disciplines. Despite ophthalmic procedures accommodating many aspects of other interventions to some extent, there are fundamental dissimilarities that prevent AR solutions designed for other interventions manifest the same success in ophthalmology. One principal difference between ophthalmic procedures and interventions such as neurosurgery, which is usually referred to as similar to eye surgery, is that an ophthalmic procedure is considerably shorter in duration. Any AR solution that requires significant calibration, preparation or activation time is irrelevant to eye surgery. On the other hand, imaging modalities in an ophthalmic operating room including optical coherence tomography have micrometer resolution. Even solutions that are proposed for microscopic neurosurgery do not reach such levels in precision. Last but not least ergonomics, the operating room structure and the interactions between the surgical staff in an ophthalmic operating room is not comparable to other interventions.

When it comes to AR techniques that are specifically proposed for ophthalmic procedures, one downside that can be observed is absence of a systematic approach. Design and implementation of a proposed technique to be used in an ophthalmic procedure should incorporate all design concerns and restrictions associated with an ophthalmic operating room and the equipment at hand as well as the limitations posed by the way the surgical staff can interact with the system. The same way in a laparoscopic intervention obstructing the view of the surgeon is last resort for AR since that is the sole visual access of the surgeon to the tissue under manipulation, in eye surgery providing haptic feedback to surgeons hands is the last resort since even slightest involuntary movements of the hand can cause harm. In modern operating ophthalmic operating rooms, the surgeon has their head positioned fixed in front of the microscope, both hands occupied with the surgical instruments and feet on a foot control pedal. Any AR solution proposed for such an environment should consider these limitations. In contrast to other surgical interventions, AR solutions for ophthalmic interventions should require the least amount of interaction since the surgeon has no means of interaction and almost no downtime.

In the sections below, the contributions of the author towards an augmented reality solution that adhere to fundamental requirements and needs of ophthalmic procedures are listed. The aim of each work is to cover a specific aspect of one or several ophthalmic procedures by proposing an assistive AR system. Combining the discussed methods has the potential of

becoming a system that can address information presentation in more scenarios and cover more surgical phases.

## 2.4.1 Introducing Augmented Reality to Optical Coherence Tomography in Ophthalmic Microsurgery (IEEE ISMAR 2015)

Despite the recent research and development in augmented reality for ophthalmic procedures which exploits the availability of intraoperative optical coherence tomography, the efforts are restrained to using OCT for augmenting the microscopic view of the surgeon. Additional information is typically extracted from secondary sources such as OCT and the surgeons view of the surgical scene is visually augmented with the extracted information. In the work presented here, the authors aim at augmenting the OCT view of the surgical scene. By doing so, the concentration of the research into the applicability of AR in ophthalmic procedures is shifted from the microscopic view to the OCT view. The main motivations behind this approach are summarized below.

The first motivation behind introducing AR to the OCT representation of the surgical field is the fact that intraoperative OCT is a continuous and complex source of information. Since the main focus of the surgeon is on the microscopic view, the information that is perceived from OCT images is superficial. However, there is additional information in OCT images that requires mental effort or time to be extracted and identified. Extracting the underlying information in OCT images and augmenting them with such information removes the burden of interpretation from the surgeon. Augmented reality is a suitable solution for this purpose. Secondly, employing AR to augment the microscopic view of the surgical field in a straightforward fashion has the downside of being obstructive and distracting. Embedding excess information in OCT images rather than the microscopic view has the benefit that surgeons can actively choose to receive the information when they desire by looking at the OCT representation of the surgical field and discard the augmented information when they are focused on the microscopic view. Having control over the surgical scene without requiring to interact with the system is reassuring. Last but not least, since the content of OCT images are in the level of micrometers, the information that can be provided with augmented reality in OCT view is also micron level. Providing information about structures and tissues that are a few micrometers in size in the microscopic view that covers a few millimeters could be out of proportion and hard to perceive.

The focus of this contribution is on posterior segment procedures such as vitrectomy and epiretinal membrane peeling where one or more instruments are introduced to the posterior segment of the eye. In such procedures the instruments are typically maneuvered around to cover a region or manipulate delicate tissue while unnecessary touch of certain regions such as the macula should be avoided. Since perception of depth when looking through a surgical microscope is problematic, OCT is an advantageous source of information. However, the metallic nature of surgical instruments means that no structure below the top segment of the instrument is visible. The proposed system extracts geometrical information regarding the instrument cross-section in OCT images and the distance of the bottommost part of

the instrument to the surface of the retina and augments the OCT images with instrument cross-section shapes and color coding. By providing this information the surgeon is able to correctly perceive the distance of their instrument to the surface of retina and avoid unnecessary touch. The approach in this work is to extract information about the cross-section of the instrument based on two assumptions. The instrument or each instrument segment in multi-segment instruments is assumed to have a cylindrical or half-cylindrical profile and The OCT engine is assumed to provide two OCT scans at the tip of the instrument, one along and over the instrument and one perpendicular to it. For the second assumption to hold true throughout the surgical phase, an instrument tracker based on the microscopic view is employed. The quantitative evaluation of the proposed AR solution shows that the maximum error in the estimation of the distance to retina is 9.76 micrometers which is the equivalent of 5 pixels in the OCT images of the test dataset. The proposed novel method for augmenting OCT images has the potential to not only help in training surgeons but also provide additional information in real clinical settings. Since commercial OCT-equipped surgical microscopes internally have calibration between the OCT engine and the optics of the microscope, employing the proposed AR system requires no additional calibration or preparation step. For the corresponding full publication the reader is referred to appendix C of this manuscript.
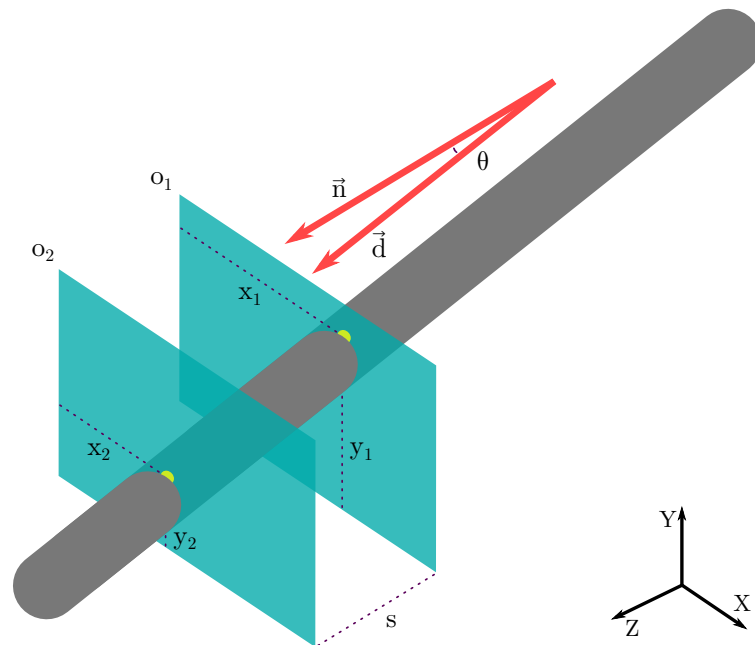


**Fig. 2.4.** Instrument cross-section calculation scheme based on consecutive parallel OCT images.

One of the limitations of the proposed method is the complication with keeping the OCT scan location and orientation exactly on the instrument and at the tip. Reliance of the system on this requirement brings about two problems. Firstly, it could interfere with the desire of the surgeon to move the OCT scan location even though in almost all cases the point of interest for the surgeon is the tip of the instrument. Secondly, errors in tracking the instrument tip, shaft and pose directly affect the quality and precision of the visual augmentation of OCT images. As a complement to the discussed literature by the author, the following extension is proposed to remove or loosen the reliance on instrument trackers by using a different scan geometry

and calculations. By having OCT scans that are parallel rather than perpendicular in scan geometry, it is possible to extract information about the cross-section of the instrument even if the instrument is in an arbitrary orientation relative to the OCT scan locations. Consecutive parallel OCT images are captured over the instrument at a location that is either selected by the surgeon or is automatically selected based on a rough estimation of the instrument location using an approximate tracker. The concept is illustrated in Fig. 2.4. $O_1$ and $O_2$ are two parallel OCT B-scans acquired with a known spacing of $s$. The reference coordinate system is defined with respect to the OCT image planes. Hence, the normal vector of the OCT image planes is $\vec{n} = (0,\ 0,\ 1)$. By detection of the topmost point of the instrument reflection in the OCT images, the three-dimensional instrument vector $\vec{d}$ can be defined as $\vec{d} = (x_2 - x_1,\ y_2 - y_1,\ s)$, where $x_i$ and $y_i$ are the coordinates of the topmost point of the instrument reflection in the $i^{th}$ OCT plane. $\theta$ is the angle between $\vec{d}$ and the OCT planes normal vector $\vec{n}$. The elliptical cross-section of the cylindrical surgical instrument can be calculated using the instrument angle, direction vector and shadow width in the OCT images.
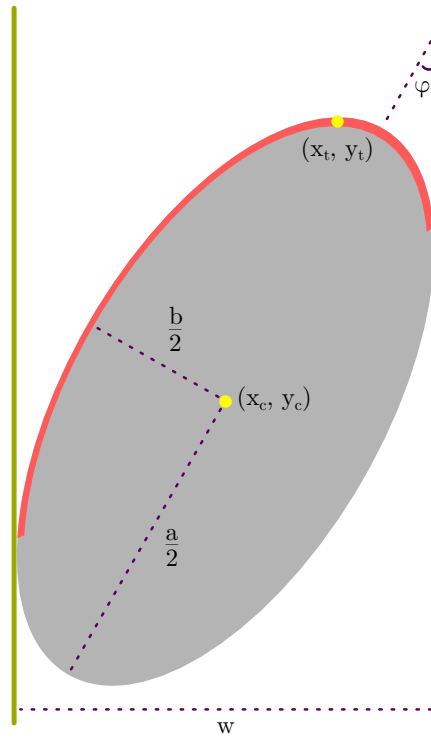


**Fig. 2.5.** Sketch of an elliptical cross-section of a cylindrical surgical instrument in an arbitrary orientation. The upper red segment is the instrument reflection visible in the corresponding OCT B-scan.

To calculate the elliptical cross-section of a cylindrical surgical instrument in one OCT image, $5$ parameters have to be estimated. (Fig. 2.5) The coordinate of the ellipse center point in the OCT image $(x_c,\ y_c)$, two ellipse axes lengths $(a,\ b)$ and the ellipse rotation angle $(\varphi)$. The rotation of the ellipse depends on the instrument direction vector relative to the OCT planes normal vector and is calculated as

$$\vec{m} = \vec{d} \times \vec{n}, \quad \varphi = \tan^{-1} \frac{\vec{m}_y}{\vec{m}_x} \tag{2.1}$$

where $\times$ is the vector cross product and $\vec{m}_x$ and $\vec{m}_y$ are the $x$ and $y$ elements of the $\vec{m}$ vector respectively. In order to calculate the ellipse axes, the instrument angle ($\theta$) relative to the OCT planes normal vector is required which can be obtained using the plane-vector angle formulation

$$\theta = \cos^{-1} \frac{|\vec{n}.\vec{d}|}{\|\vec{n}\|.\|\vec{d}\|} \tag{2.2}$$

where . denotes vector dot product, $|.|$ is the abstract and $\|.\|$ is the vector Euclidean norm. The axes of the cross-section ellipse are then defined using $\theta$ and the radius of the instrument cylinder ($r$) as follows

$$a = 2r\sec\theta, \quad b = 2r \tag{2.3}$$

Calculation of the instrument radius is possible since the width of the elliptical cross-section bounding-box (shadow width) is known ($w$).

$$w = \sqrt{a^2 \sin^2 \varphi + b^2 \cos^2 \varphi} \tag{2.4}$$

By substituting $a$ and $b$, the radius can be defined as

$$r = \frac{w}{2\sqrt{\frac{\sin^2 \varphi}{\cos^2 \theta} + \cos^2 \varphi}} \tag{2.5}$$

which leads to the determination of the cross-sectional ellipse axes. To calculate the instrument cross-section center point in the OCT image ($x_c$, $y_c$), the topmost point of the estimated cross-section along with the instrument reflection topmost point ($x_t$, $y_t$) can be used. By setting the top point of the estimated ellipse to the instrument reflection topmost point, the center points also coincide.
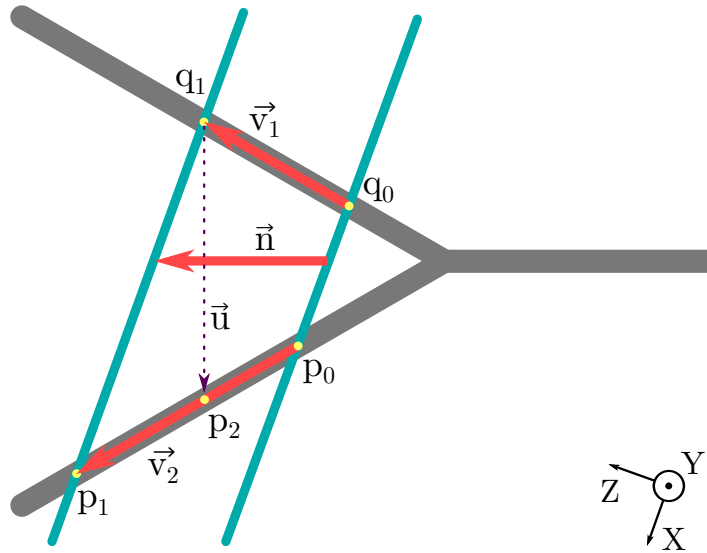


**Fig. 2.6.** Estimation of the tilt of a multi-segment cylindrical surgical instrument in an arbitrary orientation. First, calculation of the intersection of the instrument tips with a hypothetical OCT plane perpendicular to the instrument shaft is necessary ($q_1$ and $p2$ points).

Estimation of the cross-section in multi-segment instruments (e.g. forceps) needs further calculations. Such instruments in their most basic form have cylindrical shafts, but cylindrical

tips cut into halves along the axis of the shaft cylinder. Hence, the cross-sections to estimate are half-ellipses. The instrument casts two shadows in each OCT image. Each shadow in one OCT image along with its corresponding shadow in the other OCT image can be considered independent from the other pair of shadows for the cross-section estimation. In contrast to instruments with a single tip (e.g. vitrectomy probe), a pair of forceps can also be tilted along the shaft axis. This instrument tilt angle ($\psi$) must be taken into account for accurate cross-section estimation. The instrument tilt angle can be determined using the difference in the position of the tips reflection topmost points in one OCT image. But this estimation of the instrument tilt is only accurate if the OCT image is acquired perpendicular to the instrument shaft. But since this requirement is not guaranteed to be met, for a more accurate estimation of the tilt, the position of one tip reflection topmost point must be estimated on a hypothetical OCT image that is perpendicular to the instrument shaft and is on the reflection topmost point of the other tip of the instrument. The concept is illustrated in fig.2.6.
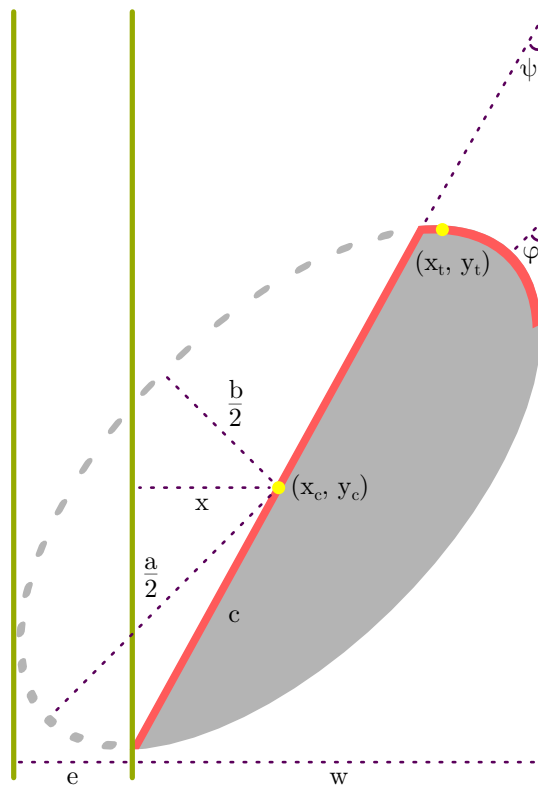


**Fig. 2.7.** Sketch of a half-elliptical cross-section of a multi-part cylindrical surgical instrument in an arbitrary orientation. The upper red segment is the instrument reflection visible in the corresponding OCT B-scan.

For this end, ray-plane intersection formulation is employed

$$\vec{n} = \hat{v_1} + \hat{v_2}, \quad p_2 = p_0 + t.\vec{v_2}, \quad t = \frac{\vec{n}.(q_1 - p_0)}{\vec{n}.\vec{v_2}} \tag{2.6}$$

in which as demonstrated in fig.2.6, $p_0$, $q_0$ and $p_1$, $q_1$ are the instrument tips reflection topmost points in the two consecutive OCT B-scans, $\vec{v_j}$ vectors are the instrument tip direction vectors and the $\vec{n}$ vector denotes the instrument shaft direction vector which is also the normal vector of the hypothetical OCT plane. $p_2$ is the instrument tip reflection topmost point if the

OCT image was acquired at the hypothetical plane. The tilt angle is then calculated as the angle between the $\vec{u}$ vector and the XZ plane.

$$\vec{u} = p_2 - q_1, \quad \psi = \sin^{-1} \frac{\vec{u}.(0,\ 1,\ 0)}{\|\vec{u}\|} \tag{2.7}$$

For each of the instrument tips an independent elliptical cross-section is to be estimated. The ellipses are then cut into halves according to each ellipse rotation angle and the tilt of the instrument. Estimation of the ellipses rotation angles and the ellipses center points is done as for the case of instruments with a single tip. But the estimation of the axes of the ellipses is done with further considerations. The shadow width of each tip is extended to a value which would be covered if the tip was fully cylindrical (fig.2.7). The width of each of the elliptical cross-sections bounding-box relative to the ellipse axes is defined as

$$w_j + e_j = \sqrt{a_j^2 \sin^2(\varphi_j) + b_j^2 \cos^2(\varphi_j)}, \quad j = 1,\ 2 \tag{2.8}$$

where $j$ is the tip index. As illustrated in fig.2.7, the extended shadow length $(w_j + e_j)$ can be calculated based on each ellipse axes.

$$
\begin{aligned}
w_j + e_j &= 2(e_j + x_j) \\
x_j &= c_j \sin(\psi) \\
c_j &= \frac{ab}{2\sqrt{a_j^2 \sin^2(\varphi_j - \psi) + b_j^2 \cos^2(\varphi_j - \psi)}}, \quad j = 1,\ 2
\end{aligned}
\tag{2.9}
$$

By substituting $w_j + e_j$ from equation 2.9 and $a_j$ and $b_j$ form equation 2.3 in the set of equations 2.8, the following set of equations can be formed. The positive roots of these equations are the radius of each instrument tip.

$$
\begin{aligned}
&\left(\frac{\sin^2 \varphi_j}{\cos^2 \theta_j} + \cos^2 \varphi_j - \frac{\sin^2 \psi}{\sin^2(\varphi_j - \psi) + \cos^2 \theta_j \cos^2(\varphi_j - \psi)}\right) r_j^2 \\
&+ \frac{2w_j|\sin \psi|}{\cos \theta_j \sqrt{\frac{\sin^2(\varphi_j - \psi)}{\cos^2 \theta_j} + \cos^2(\varphi_j - \psi)}} r_j - w_j^2 = 0, \quad j = 1,\ 2
\end{aligned}
\tag{2.10}
$$

In either case of an instrument with a single tip or an instrument with two tips, for the estimation of the distance between the instrument tip and the retina, the bottommost point of the estimated cross-section ellipse is needed. This point is straightforwardly calculated based on the instrument tip reflection topmost point and the height of the estimated cross-section ellipse or half-ellipse axis-aligned bounding box. The difference between the mean of the retina surface height detected immediately before and after the shadow region and the cross-section bottommost point is the desired distance. Since the pixel size of the OCT device is known, the actual distance is reported in micrometers. The augmentation of the OCT view of the surgical scene is done as per the original method. Fig. 2.8 shows examples of augmented OCT images taken near the tip of vitreoretinal instruments in a simulated posterior setup. The top row images are OCT acquisitions from a 23-gauge vitreoretinal probe and the bottom row images are OCT acquisitions of a 23-gauge diamonized end-gripping forceps. The four OCT images are taken at different locations with different distances between the instrument tip and the retina and with different angles between the OCT planes normal vector and the instrument shaft vector. The distance between the bottommost point of the estimated instrument cross-section

and the mean of the retina surface height detected immediately before and after the shadow region is reported in micrometers. The color of the cross-section shifts towards red as the instrument gets closer to the retinal surface.
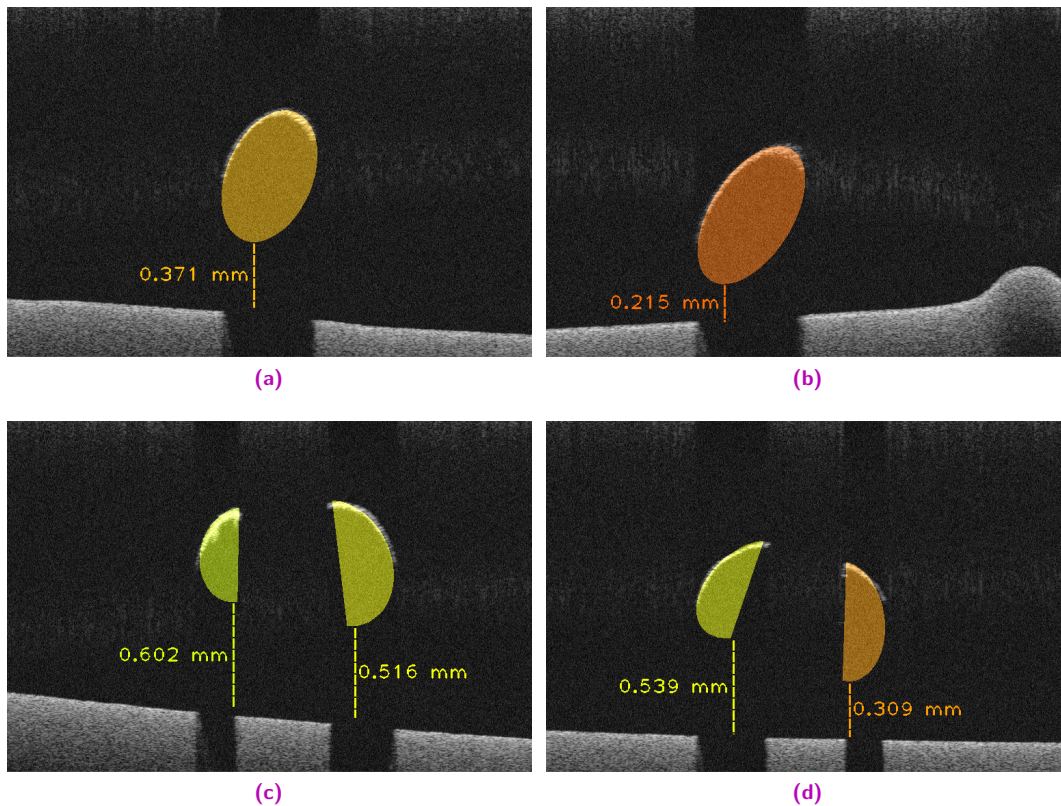


**Fig. 2.8.** Examples of augmented OCT images in a simulated setup with the proposed extended AR method. (a) and (b) are OCT images of a 23-gauge vitreoretinal probe and (c) and (d) are OCT images of a 23-gauge diamonized end-gripping forceps.

Although extracting knowledge regarding the geometrical profile of instruments in OCT images and using AR to present such knowledge seems trivial at first glance, this type of information is precisely the sort that is not immediately available to surgeons. The mental effort necessary to perceive where the bottom segment of the instrument is relative to the tissue under manipulation makes OCT images either distracting or impractical for this purpose. On the other hand, the choice of the mechanism by which information such as the one discussed is presented to the surgical staff plays a major role in the usability and accessibility of the proposed system. Conveying peripheral assistive cues and guidance with conventional methods is evidently subpar. However, employing augmented reality without care could also lead into distraction and inattentional blindness [22]. Visually augmenting OCT images rather than the surgical view in the opinion of the author has the benefit of embedding guidance and decision support knowledge in an imaging modality that is considered secondary to the surgical flow. A multitude of different cues and signals extracted from OCT images or other data sources could be presented with the same technique via visual augmentation of OCT images. This opens a new path for AR in ophthalmic procedures.

## 2.4.2 SonifEye: Sonification of Visual Information using Physical Modeling Sound Synthesis (IEEE TVCG)

Most of the research and development in augmented reality for ophthalmic procedures is focused around visual AR. This is the case not only as a result of visual AR being generally the primary and predominant approach in augmented reality, but also because surgical microscopes provide a suitable and convenient platform for visual AR. In the work presented here, the authors aim at using auditory augmented reality for guidance in high-precision tasks. Ophthalmic procedures can directly benefit from such an approach since reaching targets and manipulating them while avoiding other elements under a surgical microscope falls in the category of high precision tasks that require guidance. The main motivations behind this approach are summarized below.

Vision is the primary means of gathering information concerning interventional tasks for surgeons. The most vital signals in an operating room, instrument-tissue interactions, outcomes of decisions and results of actions are all perceived primarily visually. Although this essentially means that using visual channels for conveying information could be the most effective, On the other end of the spectrum this also means that interfering with surgeons' visual perception of the outside world could potentially be the most distracting and unproductive way. Visually augmenting the surgical scene with cues that obstruct crucial information regarding the surgical task at hand is not only hazardous, it could also lead to loss of confidence and trust in an AR system. This is particularly true in case of interventional disciplines such as ophthalmic procedures where the surgeon has very limited means of interacting with the AR system. Ophthalmic surgical microscopes equipped with intraoperative OCT already use overlays to show OCT over the surgical view. Polluting the remaining microscopic view with additional in the opinion of the author is hardly constructive whereas auditory augmented reality fits well into the workflow of ophthalmic surgeries. Another motivation to use aural methods is that even if visual presentation of information is not disruptive and done with care, it requires active comprehension of information and change of focus. Experiments on visual and auditory reaction times of medical students show that auditory reaction time is faster than visual reaction time [42] meaning auditory perception requires less cognitive effort. Hence, an auditory AR system has the potential to present data with less cognitive load on the medical staff. Moreover, research into cross-modal influences from the auditory modality on the visual modality [92] show that perceptual organization in the auditory modality can have an effect on perceptibility in the visual modality. The detectability of a visual stimulus can be enhanced by a synchronously presented abrupt tone. Hence, the multimodal nature of combining the visual stimuli from the surgical field with the aural stimuli from the auditory AR system can enhance the perception of the surgeon from events and cues present in intraoperative data streams.

The focus of this contribution is on presenting information that is extracted from visual sensors via computer vision techniques with sound. Even though the concept of sonification, *i.e.* the transformation of data relations into perceived relations in an acoustic signal, has been around for a long time, the novelty of the proposed method is in three fronts. First, the usage of sonification specifically for guidance and navigation in high-precision tasks has not been explored extensively. The present work explores the possibility of guidance via auditory

augmented reality in tasks that similar to ophthalmic procedures are performed under a microscope. Secondly, the possibility of extracting visual information and presenting it aurally with the aim of discarding the visual source of information is investigated. In ophthalmic procedures that could translate into extracting the most relevant information from OCT images to be rendered aurally without the surgeon having the possibility of looking at them. Lastly, many investigations into sonification neglect aesthetics and human psychoacoustic factors. The work at hand uses physical modeling sound synthesis to generate aural feedback. Physical modeling simulates vibrations and resonances in sound generating matters to generate audio. This leads to synthesis of audio patches that closely resemble sounds generated by natural phenomena. Contrary to sounds sampled from real sources, the parameters governing the quantitative and qualitative characteristics of the generated sounds can be changed on the fly. All of these properties could be exploited so that the final auditory augmented reality solution is pleasant for the users, comfortable to be used for long periods of time and yet informative for the employed purpose.

For the purpose of demonstrating the potential of guidance with auditory AR using sound synthesis in high-precision tasks, three examples are considered in the manuscript. Maintaining a certain pose with a miniature instrument such as a needle, applying a specific amount of pressure to a surface using an instrument and gently touching a highly viscous material. To guide a user towards maintaining an instrument in a certain pose, synthesized sounds resembling plucking and dampening a bounded string in intervals is used. The pitch of the string is set to two fixed values based on the current pose of the instrument relative to the desired pose. When the recorded angle is less than the desired one, the pitch of the plucked string is low and when the recorded angle is higher than the one desired, the pitch is high. In spite of most common sonification systems, in the proposed method plucking intervals are inversely proportioned to the distance to the desired angle. Hence, perfect pose results in total silence. For informing the user about a touch event, the sound of resonating a bounded elastic membrane when struck by hand is used. The recorded acceleration on impact is mapped directly into the sound volume and is inversely proportional to the elasticity of the virtual membrane. Representation of pressure applied to a real object is done by synthesizing the sound of tapping on a wooden square plate in intervals. As the applied pressure rises, the material which the virtual plate is composed of gradually turns from wood to metal. The target pressure is implicitly defined to be when the material of the plate is in between wood and metal. The methodology assumes there are computer vision and image processing algorithms in place for extraction of information regarding the state of the task from the relevant imaging modalities. In case of an ophthalmic procedure, continuous volumetric OCT acquisitions could be used for the estimation of the pose of an instrument or stereo cameras could be used to determine the distance between an instrument and the surface of a viscous material. Two experiments using a digitizer board and an ophthalmic surgical microscope equipped with OCT are conducted to evaluate the sonification methods. Results from the experiments indicate the effectiveness of using such sonification systems for extremely focus demanding tasks. For the corresponding full publication the reader is referred to appendix D of this manuscript.

The main tool used in this work to generate sounds using physical modeling is Modalys. Fig. 2.9 shows two instances of pipelines constructed in Max visual programming environment which uses Modalys. Fig. 2.9 (a) is a pipeline for generating sounds that resemble tapping a square plate in intervals while two pickup microphones are attached to it in arbitrary locations.
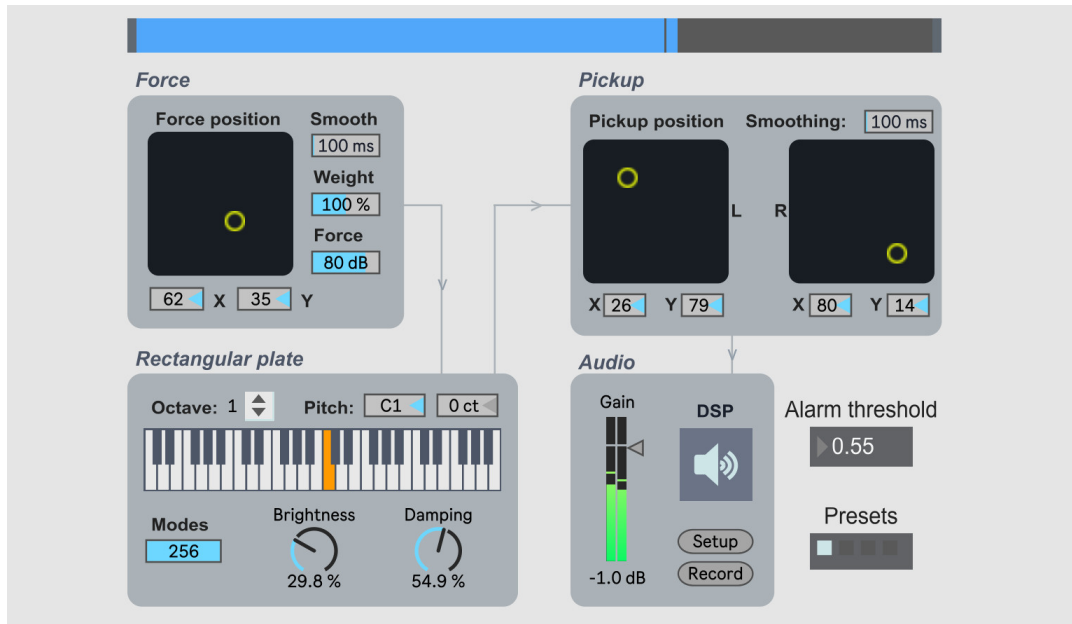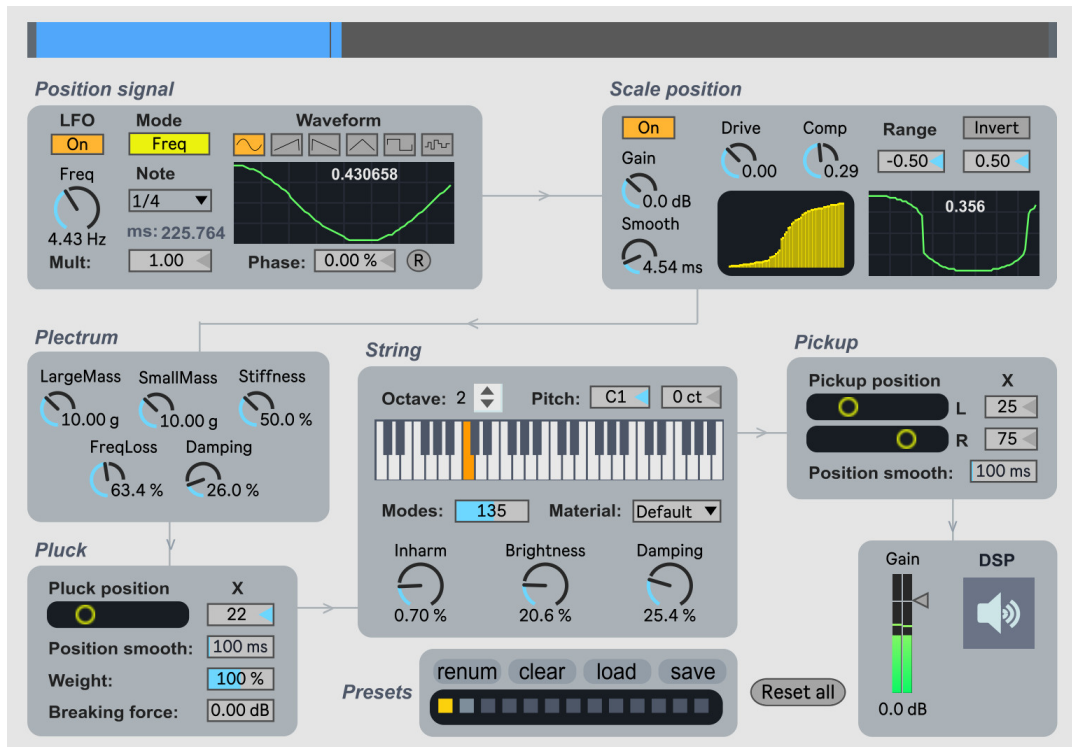
**Fig. 2.9.** Schematics of physical modeling sound synthesis pipelines using Modalys in Max visual programming environment. (a) shows a pipeline for synthesizing sounds generated by tapping on a rectangular plate and (b) is a pipeline for synthesizing sounds generated by plucking a string.

The force applied to the plate by tapping, the fundamental frequency (pitch) of the plate, the roughness of the tapping element, the material of the tapping element and the plate, etc. are all among the parameters that could be fixed or controlled by input parameters. In case of the described sonification schemes, some of these parameters are set by the information

extracted from imaging modalities. Fig. 2.9 (b) is a pipeline for generating sounds that resemble plucking a bounded string in intervals when two pickup microphones are attached to it. Among the parameters defining the quality of the generated sound in this case are the fundamental frequency (pitch) of the string, the pluck location along the string, the pluck force, the material of the string and the location of the pickup microphones on the string. How these parameters should be defined based on the extracted information from imaging devices is a matter of qualitative measurement and trial. In the presented study a pair of headphones was used for rendering the aural augmented reality system. When such a system is intended to be used in an operating room, there are constraints on the methodologies that could be used for that purpose. Loudspeakers installed away from the sterilized area is an option. Bone induction headphones could also be employed, however the sound rendering would be restricted to the users of the headphone. The presented work introduces a new paradigm for AR in ophthalmic procedures.

# Virtual Reality (VR) for Ophthalmic Procedures

<div style="text-align: right">3</div>

## 3.1 Motivation and Challenges

Augmented reality has many merits when it comes to surgical interventions including the possibility of accessing the real intact view simply by turning off augmented elements. However, taking one step further and employing virtual reality for medical interventions has benefits of its own. Medical procedures in virtual environments can potentially improve surgical outcomes by removing inconsequential features of the surgery at each phase. Relevant information at each step of the procedure can be carefully selected based on previous trials and the virtual representation of the surgical field could be constructed based on the selected information in the most basic form. This could lead to improved comprehension of the surgical field and less imposed distractions from the regions of interest. VR can provide a more immersive experience to users. Although virtual representations are typically simpler in form, their components are curated with the aim of presenting the best portrayal of the scene. Data acquired from visual sensors or other sources could be presented in forms that are more relevant to the surgical task at hand using virtual reality. All of these features, if carefully used, have the potential of improving general user experience as well as task outcomes.

One interesting application of VR systems is training. Conventional training methods require access to a real or as close as possible to real setups. For medical training this translates to high costs, regulatory issues and imposed risks. A VR solution can replicate the same environment a surgeon perceives in a real surgery and mimic the same interactions they experience when performing a surgical procedure in an entirely virtual environment. Even more beneficial is that scenarios such as emergencies, complications and unforeseeable issues could be rehearsed over and over again without added risk or cost. Such techniques can also be used for studying the effects of new methodologies or technologies on the outcomes of surgical tasks. These benefits can explain the popularity of VR solutions both in the literature and in training and educational environments. For medical staff trained by VR systems interaction with similar frameworks in real clinical environment could be natural. This could be exploited for better integration of VR frameworks in clinical workflows.

Since ophthalmic procedures are performed using magnification provided by surgical microscopes, a virtual reality solution could help with the perception of the surgeon form the miniature elements of the surgery and the microinteractions between tissues and instruments. A virtual depiction of the surgical field could illustrate components of the surgery in larger dimensions with virtual elements that are not present in the real scene. Inputs from sources such as the surgical microscope, the phacoemulsification device, OCT, preoperative angiographic images, etc. could be incorporated into building a virtual model of the surgical

scene which encapsulates all the information relevant to each phase of the surgery in the rendering. In the opinion of the author, continuous OCT volumes captured from the eye and visualized for the purpose of supporting a surgeon is on its own a virtual representation of the surgical scene. Renderings of the elements of an ophthalmic surgery visible in OCT images such as tissues and instruments are not representative of the original element in terms of visual characteristics. They are rather virtual representations of the scene formed based on the physics behind the interaction of near-infrared light with them. Any solution that provides the means for a sole OCT reproduction of the ophthalmic procedure scene to be used for a surgical procedure or a specific phase of an intervention is categorized as a VR solution. Adding other virtual elements on top of the OCT depiction of the scene only riches the rendering and makes it more representative of the entire available information at disposal at any time.

On the other end of the spectrum, it is challenging to make the virtual rendering of a surgical procedure as close as possible to the real scene or relevant. Although the aim of a VR system is not necessarily replicating the exact features of a scene, fundamental aspects of the intervention should remain intact. Deviations from the actual features could be counterintuitive if not set in place with caution. Design of a VR solution that is intended to be used in a clinical setting should accommodate factors and nuances that in anyway carry information for the surgeon. Modes of interaction with a system that is virtual in essence require study and trials. Training the surgical staff regarding the VR system plays an crucial role in the success of a proposed solution.

There are multiple technical challenges associated with virtual reality solutions specially if they are meant to be used in medical settings. surgeons in conventional surgical settings rely on the visual and haptic feedback from their maneuvers when manipulating tissues using instruments. a VR solution is typically removing their direct visual or other sensory access to the patient. Hence, the virtual reconstruction of the region of interest should have the minim delay in terms of mirroring events and illustrating alterations in the scene in real-time. The computational processing necessary for formation and conveyance of the virtual rendering should be optimized to the extent that without compromising on the level of detail and accuracy, the entire snapshot of the scene is perceived as instantaneous. On the other hand, the choice of the technology involved in the presentation of the VR system plays a major role in the success of the proposal. Using monitors and in particular 3D monitors is a proper choice both because of the ease of use and ergonomics. Head-mounted displays (HMD) are also reasonable even though for ophthalmic procedures in particular, because of the space confinement imposed my the surgical microscope employing them could lead into minor problems. Another technical challenge that should be handled is the presence of a fallback solution under unforeseen circumstances. Failure of the VR system for any reason effectively means information blindness for the surgeon. There should always be methods in place for the surgeon to have the possibility of continuing the procedure using conventional techniques. In case of ophthalmic procedures, proposing a VR solution that presents the delicacy involved in a proper manner while avoiding noise such as hand tremors is difficult. Small movements captured in the real scene or minor signals visible in data streams could either be futile or due to actual delicate events and interactions. Distinguishing the two requires suitable signal processing, image processing and computer vision techniques. Fallback solutions for microscopic surgeries are easier to handle since the ocular view of the microscope is practically available at all times.

## 3.2 Applications

Performing ophthalmic procedures in a virtual environment is in most cases in practice for training purposes. When it comes to clinical implementation of a solution that enables surgeons to perform at least portions of a surgical task in a virtual environment, procedures that heavily rely on the information from OCT are the first to come to mind. Such interventions are mostly complex tasks that require precise and real-time information about the position, orientation and movements of instruments and tissues particularly in depth. High axial resolution of OCT makes it a better source of information compared to the stereo microscopic view for perceiving how deep the tip of an instrument is or how far apart tissues and instruments are. To showcase the potential of VR techniques in real clinical settings, the focus of this work is on the OCT representation of the surgical field with applications in depth-sensitive procedures.

Deep Anterior Lamellar Keratoplasty (DALK) is a good example for a surgical task that is high-precision and depth-sensitive. Among many surgical techniques used for DALK, the big-bubble technique proposed by Anwar and Teichmann [3] is one that is successfully and regularly exercised. The defining phase of the big-bubble technique is the insertion of the needle into the stroma of the cornea to a depth of more than 80% of the stroma. The aim is to inject air or saline solution so that the mass of the stroma is detached from the healthy Descemet and endothelial layer. The stroma is then removed and the graft is put in place of it and held with stitches. If the needle does not reach the desired depth, the formation of the big-bubble and detachment does not occur and the stroma should be manually and layer by layer removed. If on the other hand the needle perforates through the endothelial layer, the DALK procedure fails and a fallback to Penetrating Keratoplasty (PKP) is needed. Since the entire human cornea is around 550 micrometers, the task of guiding the surgeon to the appropriate depth is better done with OCT rather than stereo microscopy. However, even with OCT there are many difficulties affecting outcomes of a DALK procedures. For one, OCT is conventional devices is captured along fixed lines that do not necessarily coincide with the location of the instrument tip. On the other hand, even OCT does not have the resolution required to visualize corneal layers such as the DM and the endothelial layer. And when movement and instrument-tissue interactions are involved, artifact in OCT imaging such as the total shadowing effect underneath instruments affect the surgeons' perception. A VR solution that employs continuous 3D volumetric OCT scanning with virtual elements depicting the instrument, corneal layers, targets and so on can be of great help for DALK.

Subretinal injection is also depth-sensitive to the extent that not only OCT is considered necessary for it, most of the publications in the science community proposing methods and techniques for it rely on robotic solutions. When the fine needle used for injection enters the retina, identifying the tip and the penetration depth is not trivial. However in most cases, the aim of a subretinal intervention is to reach an exact layer of the retina to deposit the therapeutic agent. In the opinion of the author, VR systems that extract all the necessary information such as segmented retinal layers visible in the OCT, the location of the needle using instrument trackers or the amount of the injected fluid from the injection device and present them by virtual elements in a constructed environment are a viable solution to many problems faced by surgeons performing subretinal injections. Of course the same VR techniques could be applied to many other phases of ophthalmic interventions.

## 3.3 Related Work

Technologies enabling virtual reality for medical interventions have been around since the later decades of the 20th century. In the 1990s, research teams including at the University of North Carolina and the US Department of Defense developed the concept of surgeons of the future equipped with virtual reality headsets and rehearsing robotic or manual procedures using advanced computer generated images [57]. The kind of transformation that was envisioned VR would bring to surgical interventions revolves around surgical practice, preoperative planning and assistance in robotic interventions [48]. The idea of using VR systems for surgical education and training has notably been considered to have tremendous promise for the future in early 2000s publications [39]. Today, virtual reality solutions providing training, simulation and planning for a variety of surgical interventions are found in abundance in literature. VR training systems for arthroscopic Surgeries [41], VR simulators for laparoscopic surgery [99], surgical planning, training, and teaching of liver surgery using virtual reality [51, 67] are among the examples that are studied and employed to improve interventional efficiency.

Ophthalmic training systems are a step forward from conventional wet laboratories. They provide the means for trainees to perform risk-free surgical tasks and augment and accelerate surgical training [36]. Throughout the last two decades many VR solutions have been proposed for training of ophthalmic surgeons each targeting specific procedures. One of the early examples of such solutions is the ORIS VR system [58] used for simulation of retrobulbar injection. ORIS is designed to permit hazard free and physically realistic practice of an invasive procedure prior to patient contact. Another example from the early days of VR in ophthalmic procedure simulation is the work by Sagar *et al.* [76] which proposes a surgical virtual environment containing a detailed model of the eye that can be used for simulation of various procedures. These early proposals mostly handle problems such as realism, render speed and such that rooted in the limited technology of the time. Neumann *et al.* [60] have proposed a VR system for simulating vitrectomy procedures with a 3D mouse and stereo glasses. Trainees can virtually perform different aspects of a vitrectomy probe with virtual instruments such as picks, blades, suction cutters, drainage needles and lasers. Verma *et al.* [90] take a more general approach by designing and testing a VR system that consists of an operating work station, computer models and a magnetic position tracking system. Their solution can provide the realism necessary for learning the skills required for performing the initial steps of a vitreoretinal procedure. In another work [17] by Choi *et al.*, a low-cost cataract surgery simulator for trainees to practice phacoemulsification procedures with computer-generated models in virtual environments is presented. Trainees can exercise their skills of cornea incision, capsulorhexis and phacoemulsification with their proposed system which consists of a computer with a conventional monitor and a pair of Phantom Omni haptic devices. One early example of a VR simulator that incorporates optical coherence tomography in the simulation process is the work by Kozak *et al.* [47]. They propose a VR simulation system that integrates OCT scans from real patients for practicing vitreoretinal surgery using different pathological scenarios. Their customized peeling algorithm provides full controllability and is developed to simulate the peeling of the epiretinal membrane and the inner limiting membrane. Including OCT in the simulation scenarios offered for training could potentially benefit the next generation of surgeons who are more exposed to avant-garde devices capable of acquiring intraoperative OCT. Adopting sophisticated simulation technology

with feedback provided by the system allows for objective, structured training to a high level of skill [55]. Considering the gap in the literature regarding VR systems that provide OCT view and the role of OCT in modern operating rooms, more research and development in this direction is expected.
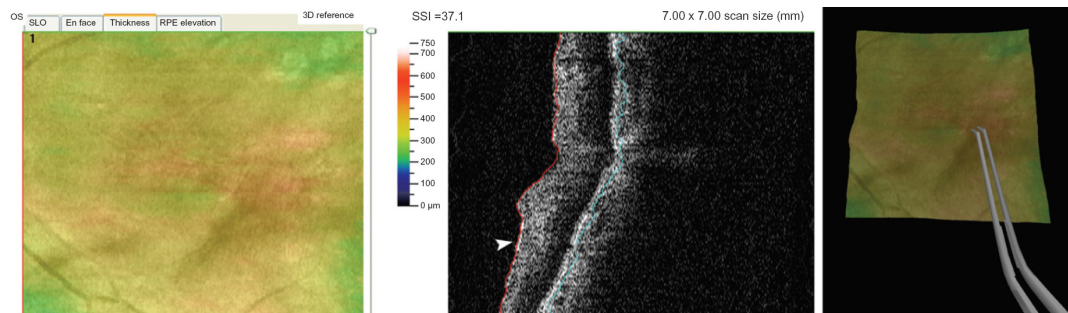


Source: [47] I. Kozak et al., Clinical ophthalmology (Auckland, NZ), 8 2014, © Dove Medical Press Limited.

**Fig. 3.1.** Example of a VR system for training with integrated OCT. The left image is a posterior pole image of an eye with ERM, the middle image is an OCT B-scan of the same eye with the arrow pointing to the ERM and the right image is the view of the simulator with an operator practicing peeling using virtual forceps.

There has been many studies comparing the outcomes of using virtual reality training simulators versus conventional training schemes such as wet laboratories. A first clinical assessment of a VR system designed for training retinal photocoagulation [63] concludes that employing virtual reality educational techniques reduces training time in addition to increasing safety for patients. Another study on the influence of virtual reality training on phacoemulsification performance [6] deduces that medical residents who are trained using VR simulators have a shorter learning curve, shorter phacoemulsification times, lower percentage powers and fewer intraoperative complications. Investigations into the benefits of using VR for teaching cataract surgery [81] show the potential of such systems to be part of the initial training of new cataract surgeons. Clinically relevant cataract surgical skills of novices and surgeons alike can be learned and improved by proficiency-based training on a virtual reality simulator [87]. And that performance on the simulators for training cataract surgery is significantly and highly correlated to real-life surgical performance [88]. A study specifically investigating surgical simulators for used for exercising capsulorhexis [56] strongly suggests that virtual reality surgical simulation training reduces the rate of errant capsulorhexes and surgical training using VR simulation should be strongly considered in ophthalmology resident surgical education to reduce the unnecessary risk of complications for live patients. Last but not least, an effort into developing a virtual reality training curriculum for ophthalmology [77] concludes that structured and supervised VR training can offer a significant level of skills transfer to novice ophthalmic surgeons. All in all, through two decades of investigations and development, the potential of virtual reality for training ophthalmic surgeons is evident.

Compared to solutions that target employing virtual reality for preoperative simulation and training, proposals that aim at guidance and navigation during ophthalmic procedures are scarce. One of the early tries at this end is the work by Dewan *et al.* [20] which provides vision-based assistance for ophthalmic microsurgeries. Their proposal does not provide virtual rendering of the scene but rather introduces the concept of virtual fixtures based on visual information for applications in retinal microsurgery. Such a system provides human-machine

cooperative guidance entirely based on a visual reconstruction of the surrounding environment by constraining the movement of the instrument. Restraint in movement could be introduced to guide the surgeon towards a target or assisting them in aligning an instrument along a preferred orientation and the like. There are also some efforts in the direction of rendering OCT data using VR techniques [70, 80]. However the aim of these studies are data fusion, analysis and interactive inspection of preoperative OCT volumes without investigating the possibility and implications of utilizing the same techniques for intraoperative OCT volumes during procedures. Considering the availability of intraoperative OCT data and the promising prospects of using VR for data rendering, guidance and navigation, the necessity of investigating the implications of ophthalmic procedures in virtual environments is apparent.

## 3.4 Contributions

As presented in the previous section, most of the proposals on the topic of VR for ophthalmic procedures are focused on preoperative training, planning and simulation. Part of this thesis aims at introducing advanced VR solutions to be employed in ophthalmic operating rooms. One could associate the absence of VR solutions in ophthalmic operating rooms to the fact that most surgical microscopes still provide direct view of the surgical scene via optics that redirect light to surgeons' eyes. This essentially means that ophthalmic surgeons are constrained to looking through ocular lenses that at most have data injection mechanisms for augmenting the view. The possibility of having a VR system in an ophthalmic operating room is only provided when three-dimensional monitors, head-mounted displays or other similar technologies are present. There is a growing interest in the concept of digital microscopes. That is, microscopes that act similar to endoscopes by having cameras that capture the surgical scene and a display or multiple displays that provide the view to the surgical staff. Such systems are theoretically less expensive to produce, easier to maintain and can have advanced computer-aided features. With more and more digital microscopes capable of using video see-though and 3D displays for visualization, there would be more interest in virtual reality solutions since VR techniques in other types of surgical interventions have proven to be advantageous. One other benefit of employing fully digital systems for surgical magnification is the improvement in ergonomics for operators. Since with a microscope equipped with 3D or head-mounted displays the need for the surgeon to look through ocular lenses is eliminated, the dynamics of the operating room can change towards the better. The empty space now available in front of the surgeon can also be utilized by devices that employ other means for communicating information such as spatial speakers and projectors. All in all, with all the new technologies in ophthalmic operating rooms around the corner, the case in this thesis for an ophthalmic VR solution to be used intraoperatively is more than ever relevant.

In the section below, the contribution of the author towards a virtual reality system that uses the full capacity of modern ophthalmic surgical microscopes equipped with intraoperative OCT acquisition engines is brought about. Even though the proposed approach is designed to provide support for a specific phase of a specific surgical intervention via VR, the principles, used techniques and proposed methodologies could be applied to other interventions. Similarly, the display technology used for rendering virtual assistive elements and cues could be substituted with any choice that fits the requirements of other task.

## 3.4.1 A Surgical Guidance System for Big-Bubble Deep Anterior Lamellar Keratoplasty (MICCAI 2016)

Virtual visual representation of a surgical scene is typically a depiction of the reality that has either elements simplified and removed, or elements enriched and added. An example is non-photorealistic rendering of segmented elements in a scene. Such rendering simplifies the scene by removing complexities such as shadows and small color variations and uses segmentation to put emphasis on certain parts of the scene. Often the simplified virtual representation of the scene is also augmented with navigation or other information. In contrast to proposals that follow the same pattern, in the work presented here, the focus of the proposed VR system is on optical coherence tomography. The attempted goal is to represent an ophthalmic procedure scene only using OCT. Interactions of the near-infrared light source used for OCT with the tissue do not produce a real representation of the scene but rather show boundaries and artifacts that are effects of the interaction of light with tissues. Continuous volumetric OCT if presented properly could in most cases be representative of the most important elements of the surgical scene for a surgeon to perform. In this work, a commercial ophthalmic surgical microscope capable of acquiring about 25000 A-scans per seconds that conventionally acquires a few high resolution B-scans continuously is modified to capture continuous low resolution OCT volumes. Commercial surgical microscopes equipped with intraoperative OCT engines use galvanometers to deflect the laser source beam to form B-scans out of multiple A-scans. The movement pattern of these galvanometers could be designed in such a way that continuous scan of a limited 3D region in the eye is possible. However, since the covered region is undersampled when such a technique is used, methods should be designed and incorporated for interpolation of data.
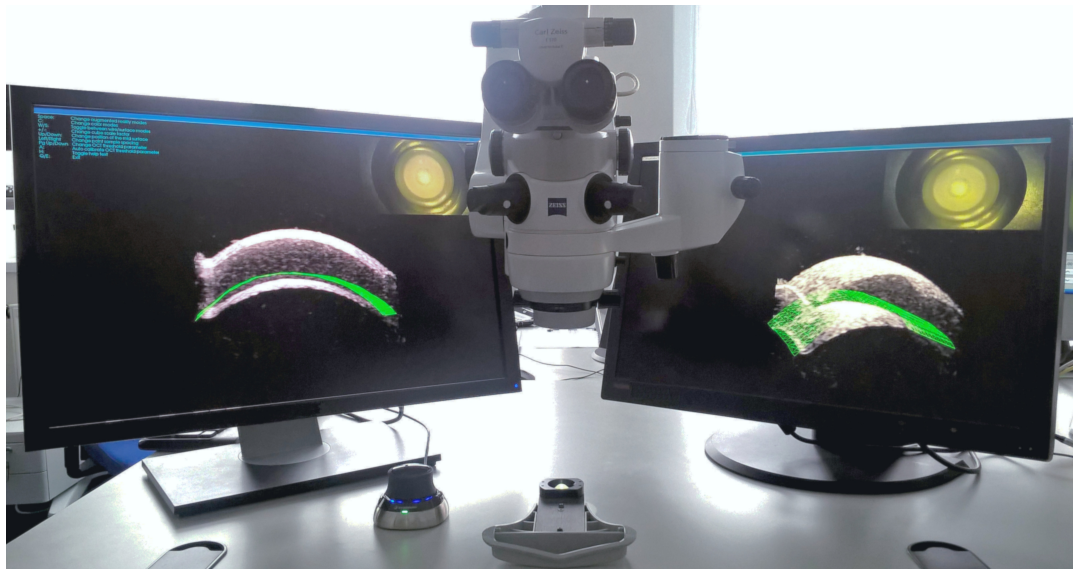


**Fig. 3.2.** The setup of the proposed virtual reality solution with two monitors, a 3D mouse and the microscope. The view angle of the visualization on the left monitor is automatically adjusted to show the side view and the view angle of the visualization on the right monitor is selected by the 3D mouse. The green wireframe mesh guides the surgeon to the target corneal depth.

In this work, the designed and demonstrated method is targeting the bubble formation step of the deep anterior lamellar keratoplasty using the big-bubble technique (BB-DALK). The reasoning behind this choice boils down to two facts. Firstly, the bubble formation step in the BB-DALK procedure is a delicate phase that often requires guidance provided by OCT to be performed. However conventional intraoperative OCT has the problem that not only the surgeon has to switch focus from the microscopic view to the OCT view often to see the needle penetration depth, but also keeping the location of a single B-scan capture exactly over the needle or at the tip of the needle is a constant problem. For every small movement of the needle, the surgeon has to use the foot pedal control to move the acquisition location to the new coordinates. Secondly, the bubble formation step is procedurally simple, straightforward to formulate and generalizable. This makes the proposed method extendable to cover similar procedures that require the same guidance. One example is subretinal injection or cannulation in which the approach is similarly guidance of the surgeon until a desired depth of penetration is achieved. In procedures like these, most of the elements visible in conventional microscopic view could be eliminated and absolute basic information kept in place for keeping the center of attention of the surgeon on the target. This is particularly true since in these procedures, the surgeon has only limited options in terms of freedom in the movement of the instrument or actions. Hence, providing a virtual representation of the scene that only contains the absolute basic information about the target and taking away the burden of controlling the acquisition device is extremely helpful in the opinion of the author. For a surgical phase that is concentrated around the target of reaching a certain depth, the idea of surgery using only continuous three-dimensional OCT with guidance is shown to work in this proposal.



**Fig. 3.3.** Surgeon performing DALK on an *ex vivo* porcine eye using the proposed VR solution. With the virtual representation of the procedure, the surgeon can perform the surgery without looking at the real scene.

When the proposed method is employed, the OCT engine with the modified scan pattern and data interpolation is used for data acquisition. The acquired volumes are rendered in real-time on regular monitors in front of the surgeon. Multiple monitors are used to

provide the surgeon with different view angles of the volume. Since the penetrated depth is measured with reference to corneal layers or retinal layers if the method is applied to posterior segment procedures, the rendered virtual representation is augmented with wireframe meshes emphasizing these layers. The target depth is also visualized with a hypothetical wireframe mesh. In the rendered volume, the color and opacity of voxels are determined based on adaptive transfer functions. These functions ensure that the used instrument is distinctly and continuously visualized in red, the background speckle noise is suppressed and the tissue opacity does not obscure the instrument. In the beginning of the procedure the surgeon can orient the instrument using the microscopic view rendered on the displays. Later on, the entire surgical phase is performed using a virtual augmented OCT view. For the corresponding full publication the reader is referred to appendix E of this manuscript.

Compared to ophthalmic surgery using the microscopic view with OCT B-scans, surgery with a virtual continuous three-dimensional OCT provides a better perception of the interactions between the instrument and tissues. As depicted in Fig. 3.3, with such a system in place, the surgeon is capable of performing surgery without the ocular view of the microscope. Fig. 3.4 better illustrates how the the scene, depths and interactions are perceived in an improved manner using the virtual depiction. The top row of the figure is showing a situation where an injection needle is kept above an *ex vivo* porcine cornea in a simulated anterior segment surgery. The needle is positioned in such a way that the beveled edge of the needle is pointing upwards. Image (a) shows an OCT cross-sectional view of the scene if the capture location is set exactly over the needle. Conventionally, this would be the best OCT illustration of the situation since the slightest movement of the needle moves it outside of the B-scan visible window and OCT is only captured and visualized as single high-resolution B-scans. However, although this view is the best possible, and even though OCT provides a better depth perception compared to the microscopic view, the total shadowing underneath the needle diminishes the perception of depth. Image (b) on the other hand shows a 3D rendered view of a similar scene formed by a customized version of the proposed VR method. In this render, no color transfer is used and the opacity transfer functions are chosen such that only the instrument, the epithelium and the endothelium are visible. The opacity of all other elements such as the speckle noise in the background or the stroma are close to zero. Although the region directly below the instrument harbors the same total shadowing effect, the surrounding regions provide a reference as to where the epithelium and endothelium layers should be below the instrument. Moreover, the possibility of changing the view point provides other cues regarding the distance of the instrument to the tissue or where the beveled edge and the injection needle opening is. The middle row of the figure is showing a situation where an injection needle is inserted into an *ex vivo* porcine cornea in a simulated anterior segment surgery. Image (c) is similarly showing a parallel OCT B-scan captured over the instrument and parallel to it. The sudden change in the slope of the instrument is due to the difference in the refractive index of the cornea and the air above it. Image (d) is the same scene rendered using the proposed VR method. Here not only the whole thickness of the cornea is clearly visible, the utilized color and opacity transfer functions differentiate the needle from the corneal stroma. The small region on the cornea that similar to the needle is illustrated in red, is the vertex of the cornea which has the most specular reflection, hence a profile similar to metallic instruments. By changing the viewpoint, it is possible to perceive how deep the needle is inside the cornea. The bottom row of the figure is showing a situation where a pair of forceps is simulating epiretinal membrane peeling on an *ex vivo* porcine retina. Here,
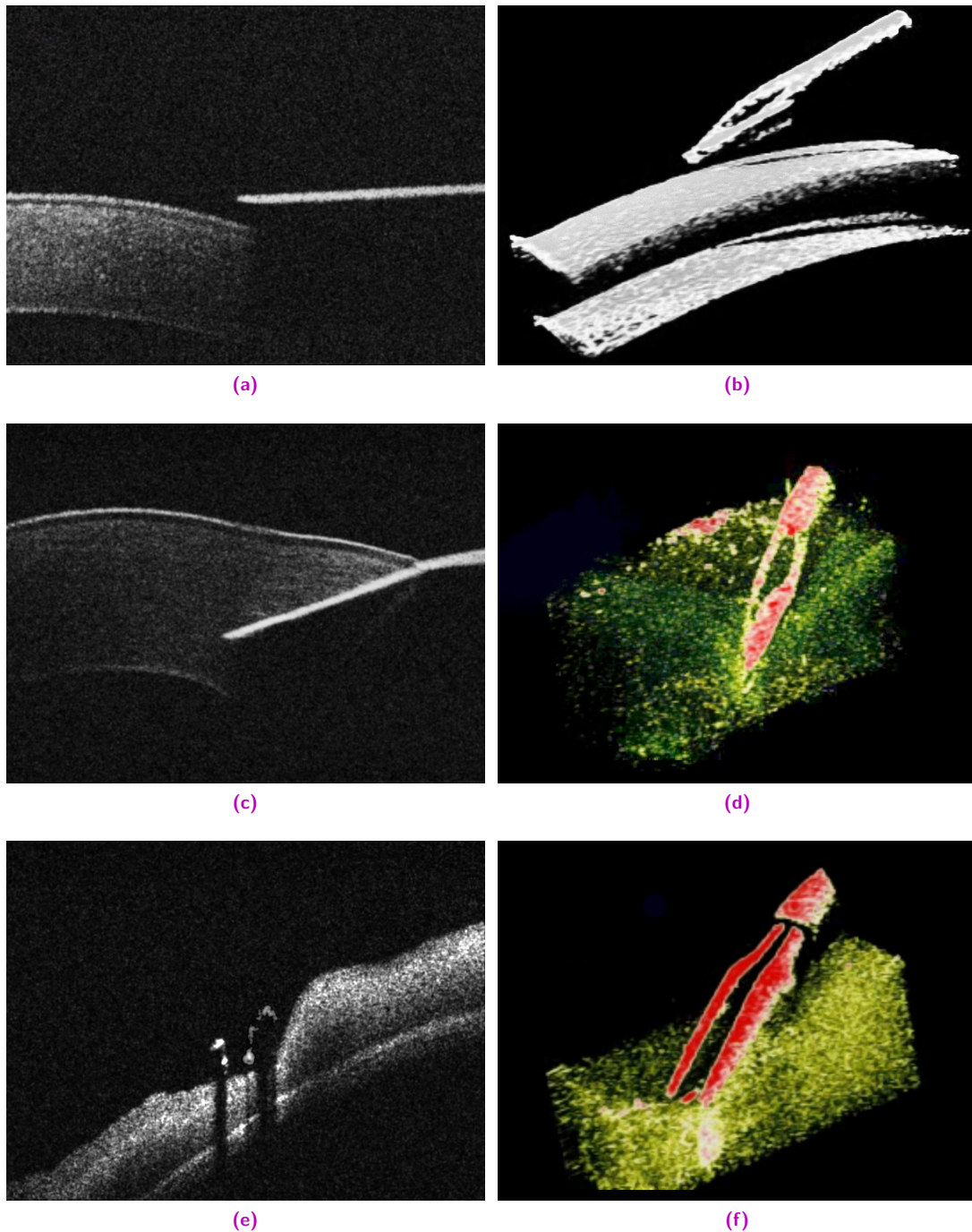
**Fig. 3.4.** Conventional OCT views and virtual views of simulated surgeries using the proposed method. (a), (c) and (e) are conventional OCT B-scans of surgical scenes as viewed on commercial ophthalmic surgical microscopes capable of intraoperative OCT acquisition. (b), (d) and (f) are virtual 3D views of the corresponding surgical scenes using continuous low-resolution volumetric OCT acquisitions.

image (e) shows an OCT B-scan that is captured at the tip of the instrument and perpendicular to the shaft of the instrument. For the case of forceps, this is a better view as both tips of the forceps and the interactions are visible. Image (f) is the virtual representation of the same scene with the proposed method. Similar to the previous case, the color and opacity transfer functions are in place to provide the maximum differentiation between the instrument and

the retina. Although the retinal layers are not as prominently visible in the 3D VR view, the interactions of the forceps with the retina are better visible in continuous mode. In general, since the VR system is capable of showing a better representation of the instrument and the regions below the instrument, it could substitute the microscopic and the conventional OCT view for certain types of surgery or certain phases of surgery.

# Conclusion

<div style="text-align: right; font-size: 3em; color: #2a6ebb;">4</div>

## 4.1 Discussion

This thesis encompasses multiple ideas in the form of method proposals, experiment reports and discussions on the topic of computer-aided ophthalmic procedures. Specifically, introducing extended reality to ophthalmic operating rooms via various techniques and for different use cases. The main contributions of this work could be summarized as follows:

- Introducing methods in form of augmented reality and virtual reality solutions that feature intraoperative optical coherence tomography as the primary imaging modality and the central extended view. The shift from extended reality systems that augment the conventional microscopic view with elements from OCT to solutions that extend the OCT view is shown to have a positive impact on surgical outcomes in simulated setups.

- Expanding extended reality solutions that are targeting ophthalmic procedures to include auditory information conveyance techniques. Using sonification to express information that is extracted from secondary imaging modalities is shown to positively impact surgical outcomes in simulated setups.

- Proposing augmented reality and virtual reality solutions that could be employed intraoperatively for ophthalmic procedures as opposed to XR systems that are solely suitable for preoperative planning, training and simulation. The proposed solutions are evaluated in simulated ophthalmic surgical setups.

To conclude the covered topics, here is a short discussion on the most notable findings and contributions made during the research leading to this thesis. On the topic of extended reality for ophthalmic procedures in general, one could refer to similar surgical interventions that have benefited throughout the years by using AR and VR solutions. One fitting example is microscopic neurosurgery where not only surgical microscopes similar to ophthalmic ones are used, but also general surgical conditions such as the availability of secondary imaging modalities such fluoroscopy or the limited surgical region are similar. Although there is not much literature around the topic of XR for ophthalmic procedures, the extensive work and the positive results of trials for neurosurgery and alike promise a transformational impact on ophthalmic procedures as well. Experiments conducted during the course of this work, although limited, show the positive impact such systems could have on surgical outcomes and surgeons' experience. As OCT in operating rooms becomes more and more commonplace and with the advent of even more imaging modalities such as Raman spectroscopy in operating rooms, the need for employing practical information conveyance methods and User Experience

(UX) designs become more apparent. Extended reality in the opinion of the author has the essential requirements to be used in ophthalmic operating rooms for the purpose.

Regarding the effect of using optical coherence tomography as the primary view of ophthalmic procedures instead of the microscopic view, both positive and negative points could be made. OCT inherently is more informative when it comes to describing anatomical layers or positions and movements in depth. Of course using OCT as a side imaging modality is beneficial on its own. However the constant change of focus from the primary imaging modality to the secondary is suboptimal. On the other hand, interpreting continuous OCT acquisitions are not trivial. artifacts such as total shadowing and mirroring could immensely affect one's perception of the view. Using OCT as the primary imaging modality and augmenting it with information extracted from within OCT acquisitions or gathered from other sources of information would eliminate such complications. One negative point against using OCT as the main imaging modality would be the loss of positional information because of the limited field of view of OCT devices and the similar loss of information regarding visual qualities such as color and reflectivity of tissues and instruments. Such information could be extracted and utilized for augmenting OCT. At the time of writing, OCT acquisition engines capable of acquiring millions of A-scans per second with much larger scan depths are already introduced. It is only a matter of time before such devices are employed intraoperatively. With the immense amount of data generated by such devices, turning OCT into the principal imaging modality for ophthalmic procedures could be more efficiently realized. The major finding in this work regarding the use of augmented OCT as the primary source of information has been the results of evaluations and opinions of surgeons emphasizing on the benefits of such approaches for certain surgical phases. Feasibility of such techniques for entire surgical procedures is however a subject that requires further analysis.

Regarding the adoption of techniques such as sonification and auditory augmented reality that use non-visual techniques for expressing information, the major finding of this work reveals two points. Firstly, the major obstacle to the success of XR systems for high-precision applications such as ophthalmic procedures is the fact that augmenting the visual field of the operator in a trivial manner negatively affects the performance. If image fusion techniques are employed without care, the obstruction of the scene negatively impacts the performance and if the augmented information is not in direct view, the required switch in focus is destructive. Designing XR solutions that circumvent problems like these by adhering to the requirements of each application exclusively is certainly a solution. On the other hand, employing auditory augmented reality is shown in this work to be as effective. Non-speech sound could provide one or multiple channels of communication between sources of information and operators such as surgeons. Indeed similar techniques with devices such as phacoemulsifiers are already used. Taking advantage of the benefits of sonification and auditory AR could provide intraoperative assistance to surgeons by providing information they typically have to refer to secondary modalities to acquire. The second major finding is that combination of visual and auditory information transfer methods is not as effective as only employing auditory mechanisms. This could be attributed to the fact that when many mechanisms are available to acquire information, humans tend to fall back to vision. The techniques proposed in this work could be employed for representation of intraoperative OCT using sound in order to avoid problems associated with visual data rendering. More research into implications, use cases and nuances

of employing sound for information conveyance in ophthalmic procedures could bring about improved surgical devices.

Regarding the contribution of introducing XR solutions that could be employed intraoperatively for ophthalmic procedures, it is worth mentioning that all the concepts in this work are designed with intraoperative support as the main purpose and investigated in resembling conditions. Quantitative evaluations show that as reported in similar disciplines, XR has the potential to benefit ophthalmic surgical interventions. Approaches presented here are among the first of their kind for use during eye surgeries. With the surge in ophthalmic procedures caused by population aging and the need for fast-paced efficient interventions guided by the state-of-the-art imaging techniques, XR could substantially contribute to the resolution of issues observed in conventional workflows.

## 4.2  Future Work

As to how the work presented here could be continued, there are a few endeavors in the same direction that could be considered as possible future extensions. Defining quantitative metrics for measuring the performance of XR systems specifically designed as assistive tools for ophthalmic procedures would lead to more realistic comparisons between methodologies. There is still room for more methods designed for supporting specific types or phases of ophthalmic procedures.  Human cadaver experiments and clinical studies can shed light on aspects of the the application of XR in ophthalmic operating rooms that could not be investigated in the limited set of experiments in this work. The possibility of employing other information conveyance techniques or various combinations of techniques such as haptics and spatial sound can be explored for usage in ophthalmic operating rooms. And last but not least, application of artificial intelligence techniques including machine learning algorithms for information extraction and presentation could lead to more practical XR concepts.

# Appendix

# List of Authored and Co-authored Publications

<div style="text-align: right">

A

</div>

**2019**

[71] **Hessam Roodaki**, Wei-Jun Chen, Daniel Zapp, Abouzar Eslami. "Simulated Keratometry using Microscope-Integrated Optical Coherence Tomography". *Investigative Ophthalmology & Visual Science (IOVS) 60.9, 2019*.

[73] **Hessam Roodaki**, Matthias Grimm, Nassir Navab, Abouzar Eslami. "Real-time Scene Understanding in Ophthalmic Anterior Segment OCT Images". *Investigative Ophthalmology & Visual Science (IOVS) 60.11, 2019*.

**2018**

[102] Mingchuan Zhou, Kai Huang, Abouzar Eslami, **Hessam Roodaki**, Daniel Zapp, Mathias Maier, Chris P Lohmann, Alois Knoll, M Ali Nasseri. "Precision Needle Tip Localization using Optical Coherence Tomography Images for Subretinal Injection". *IEEE International Conference on Robotics and Automation (ICRA), 2018, Brisbane, Australia*.

[52] Sasan Matinfar, M Ali Nasseri, Ulrich Eck, Michael Kowalsky, **Hessam Roodaki**, Navid Navab, Chris P Lohmann, Mathias Maier, Nassir Navab. "Surgical Soundtracks: Automatic Acoustic Augmentation of Surgical Procedures". *International Journal of Computer Assisted Radiology and Surgery (IJCARS) 13.9, 2018*.

**2017**

[101] Mingchuan Zhou, Kai Huang, Abouzar Eslami, **Hessam Roodaki**, Haotian Lin, Chris P Lohmann, Alois Knoll, M Ali Nasseri. "Beveled Needle Position and Pose Estimation Based on Optical Coherence Tomography in Ophthalmic Microsurgery". *IEEE International Conference on Robotics and Biomimetics (ROBIO), 2017, Macao, China*.

[53] Sasan Matinfar, M Ali Nasseri, Ulrich Eck, **Hessam Roodaki**, Navid Navab, Chris P Lohmann, Mathias Maier, Nassir Navab. "Surgical Soundtracks: Towards Automatic Musical Augmentation of Surgical Procedures". *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2017, Quebec City, QC, Canada*.

[74] **Hessam Roodaki**, Navid Navab, Abouzar Eslami, Christopher Stapleton, Nassir Navab. "SonifEye: Sonification of Visual Information using Physical Modeling Sound Synthesis". *IEEE Transactions on Visualization and Computer Graphics (TVCG) 23.11, 2017*.

[103]  Mingchuan Zhou, **Hessam Roodaki**, Abouzar Eslami, Guang Chen, Kai Huang, Mathias Maier, Chris Lohmann, Alois Knoll, Mohammad Nasseri. "Needle Segmentation in Volumetric Optical Coherence Tomography Images for Ophthalmic Microsurgery". *Applied Sciences Special Issue Development and Application of Optical Coherence Tomography (OCT) 7.8, 2017*.

[31]  Abouzar Eslami, Stefan Duca, **Hessam Roodaki**, M Ali Nasseri, Sabrina Bohnacker, Daniel M Zapp, Mathias M Maier, Jochen Straub. "Incremental Enhancement of Live Intraoperative OCT Scans by Temporal Analysis". *Investigative Ophthalmology & Visual Science (IOVS) 58.8, 2017*.

[24]  Stefan Duca, Konstantinos Filippatos, Jochen Straub, M Ali Nasseri, Daniel M Zapp, Nassir Navab, Mathias M Maier, **Hessam Roodaki**, Abouzar Eslami. "Evaluation of Automatic Following of Anatomical Structures for Live Intraoperative OCT Repositioning". *Investigative Ophthalmology & Visual Science (IOVS) 58.8, 2017*.

## 2016

[2]  Mohamed Alsheakhali, Abouzar Eslami, **Hessam Roodaki**, Nassir Navab. "CRF-Based Model for Instrument Detection and Pose Estimation in Retinal Microsurgery". *Computational and Mathematical Methods in Medicine, 2016*.

[75]  **Hessam Roodaki**, Chiara Amat di San Filippo, Daniel Zapp, Nassir Navab, Abouzar Eslami. "A Surgical Guidance System for Big-Bubble Deep Anterior Lamellar Keratoplasty". *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2016, Athens, Greece*.

[43]  Megha Kalia, Christian Schulte zu Berge, **Hessam Roodaki**, Chandan Chakraborty, Nassir Navab. "Interactive Depth of Focus for Improved Depth Perception". *International Conference on Medical Imaging and Augmented Reality (MIAR), 2016*.

## 2015

[72]  **Hessam Roodaki**, Konstantinos Filippatos, Abouzar Eslami, Nassir Navab. "Introducing Augmented Reality to Optical Coherence Tomography in Ophthalmic Microsurgery". *IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2015*.

# Abstracts of Publications not Discussed in this Thesis

## CRF-Based Model for Instrument Detection and Pose Estimation in Retinal Microsurgery

Mohamed Alsheakhali, Abouzar Eslami, Hessam Roodaki, Nassir Navab

Detection of instrument tip in retinal microsurgery videos is extremely challenging due to rapid motion, illumination changes, the cluttered background, and the deformable shape of the instrument. For the same reason, frequent failures in tracking add the overhead of reinitialization of the tracking. In this work, a new method is proposed to localize not only the instrument center point but also its tips and orientation without the need of manual reinitialization. Our approach models the instrument as a Conditional Random Field (CRF) where each part of the instrument is detected separately. The relations between these parts are modeled to capture the translation, rotation, and the scale changes of the instrument. The tracking is done via separate detection of instrument parts and evaluation of confidence via the modeled dependence functions. In case of low confidence feedback an automatic recovery process is performed. The algorithm is evaluated on in vivo ophthalmic surgery datasets and its performance is comparable to the state-of-the-art methods with the advantage that no manual reinitialization is needed.

## Interactive Depth of Focus for Improved Depth Perception

Megha Kalia, Christian Schulte zu Berge, Hessam Roodaki,
Chandan Chakraborty, Nassir Navab

The need to look into human body for better diagnosis, improved surgical planning and minimally invasive surgery led to breakthroughs in medical imaging. But, intraoperatively a surgeon needs to look at multi-modal imaging data on multiple displays and to fuse the multi-modal data in the context of the patient. This adds extra mental effort for the surgeon in an already high cognitive load surgery. The obvious solution to augment medical object in the context of patient suffers from inaccurate depth perception. In the past, some visualizations have addressed the issue of wrong depth perception, but not without interfering with the natural intuitive view of the surgeon. Therefore, in the current work an interactive depth of focus (DoF) blur method for AR is proposed. It mimics the naturally present DoF blur effect in a microscope. DoF blur forces the cue of accommodation and convergence to come into

effect and holds potential to give near metric accuracy; its quality decreases with distance. This makes it suitable for microscopic neurosurgical applications with smaller working depth ranges.

# Evaluation of Automatic Following of Anatomical Structures for Live Intraoperative OCT Repositioning

Stefan Duca, Konstantinos Filippatos, Jochen Straub, M Ali Nasseri, Daniel M Zapp, Nassir Navab, Mathias M Maier, Hessam Roodaki, Abouzar Eslami

*Purpose*. The scan location of microscope integrated optical coherence tomography is primarily controlled through the foot panel joystick. We study the feasibility of automatically following an arbitrary structure by tracking the surgical view and updating the OCT scan location during porcine wet labs and on recorded operations.

*Methods*. Two Lumera 700 microscopes equipped with RESCAN 700 (Zeiss, Oberkochen) were modified to allow the surgeon to set the scan location of the live iOCT at any position such as macula to be automatically followed and acquire cross-sectional scans continuously. The system needed to be robust to illumination changes, instrument occlusion, out of focus views, and other common issues during actual surgeries. Two groups of surgeons including 5 posterior-segment surgeons and 1 anterior segment surgeon tried the modified system during wet labs. Following wet lab with porcine eyes, all surgeons answered yes or no to the questions: 1- Is the feature beneficial during surgery? 2- Is the system sufficiently fast? The second group of surgeons also answered yes or no to the questions: 3- Is the tracker robust to the illumination change? 4- Is the tracker robust to the occlusions by surgical instrument? In addition to the subjective tests, the performance of the tracker was evaluated by reviewing individual frames from several recorded operations. Landmarks were manually annotated every 5 frames and their position was compared with the estimated position from the algorithm.

*Results*. The answers of the surgeons to the questionnaire are reported in the following table. A total of 10250 fundus view frames from 38 operations were reviewed for accuracy and figure 1 shows the performance of the algorithm in following structures with different magnifications. The accuracy of the algorithm is measured as $11.19 \pm 9.48$ pixels and its Bland-Altman plot is shown in figure 2. Q1 (5 Y 1 N) Q2 (6 Y 0 N) Q3 (1 Y 1 N.A.) Q4 (2 Y 0 N)

*Conclusion*. Most surgeons participating in this study agree on the performance of the proposed feature for following anatomical structures. This is supported by the quantitative analyses.

# Incremental Enhancement of Live Intraoperative OCT Scans by Temporal Analysis

Abouzar Eslami, Stefan Duca, Hessam Roodaki, M. Ali Nasseri, Sabrina Bohnacker, Daniel M Zapp, Mathias M Maier, Jochen Straub

*Purpose*. Microscope-integrated optical coherence tomography provides live cross-sectional view of tissues during surgery. During porcine wet lab and recorded scans, we studied the feasibility of incremental enhancement of the live intraoperative OCT scans by analyzing their temporal consistency without interrupting the surgery.

*Methods*. The suggested concept is the detection of temporal consistency of the live OCT scans and incrementally enhancing it through spatiotemporal analysis. This concept is specialized to live intraoperative OCT scans in contrast to the conventional diagnostic OCT in which the captured signal is enhanced after acquisition. Two groups of surgeons (6 in total) tested the suggested concept during wet lab using two modified Lumera 700 Microscopes with RESCAN 700 (Zeiss, Oberkochen). The common question to all surgeons was: Q1- Are the live OCT scans enhanced? Additionally, the second group of surgeons answered to the following questions: Q2- Is there interference with surgical routine? Q3- Is the temporal consistency of the OCT signal successfully detected? Q4- Is the processing adequately fast for live OCT scan? In addition to the subjective tests, 27 recorded OCT scans were analyzed to quantify the achieved improvement based on their contrast to noise ratios (CNR). The CNR value is computed as the difference between the average signals over the foreground and background divided by the standard deviation of the background signal.

*Results*. Figure 1 shows examples of the performance of the proposed concept at anterior and posterior sides of the eye as well as the iris-cornea corner. The quantitative analysis showed 11.62 db increase in the contrast to noise ratio of the live OCT by the suggested concept. The YES(Y) or NO(N) answers of the participant surgeons are: Q1 (6Y 0N) Q2 (0Y 2N) Q3 (2Y 0N) Q4 (2Y 0N)

*Conclusion*. The surgeons participating in this study approved the proposed concept. The quantitative and qualitative assessments of the results proved the improvement in the live iOCT scans quality encouraging future research and development to provide better live iOCT images to the surgeons.

# Needle Segmentation in Volumetric Optical Coherence Tomography Images for Ophthalmic Microsurgery

Mingchuan Zhou, Hessam Roodaki, Abouzar Eslami, Guang Chen, Kai Huang, Mathias Maier, Chris P. Lohmann, Alois Knoll, Mohammad Ali Nasseri

Needle segmentation is a fundamental step for needle reconstruction and image-guided surgery. Although there has been success stories in needle segmentation for non-microsurgeries, the methods cannot be directly extended to ophthalmic surgery due to the challenges bounded to required spatial resolution. As the ophthalmic surgery is performed by finer and smaller surgical instruments in micro-structural anatomies, specifically in retinal domains, difficulties are raised for delicate operation and sensitive perception. To address these challenges, in this paper we investigate needle segmentation in ophthalmic operation on 60 Optical Coherence Tomography (OCT) cubes captured during needle injection surgeries on ex-vivo pig eyes. Furthermore, we developed two different approaches, a conventional method based on morphological features (MF) and a specifically designed full convolution neural networks (FCN) method, moreover, we evaluate them on the benchmark for needle segmentation in the volumetric OCT images. The experimental results show that FCN method has a better segmentation performance based on four evaluation metrics while MF method has a short inference time, which provides valuable reference for future works.

# Surgical Soundtracks: Towards Automatic Musical Augmentation of Surgical Procedures

Sasan Matinfar, M. Ali Nasseri, Ulrich Eck, Hessam Roodaki,
Navid Navab, Chris P. Lohmann, Mathias Maier, Nassir Navab

Advances in sensing and digitalization enable us to acquire and present various heterogeneous datasets to enhance clinical decisions. Visual feedback is the dominant way of conveying such information. However, environments rich with many sources of information all presented through the same channel pose the risk of over stimulation and missing crucial information. The augmentation of the cognitive field by additional perceptual modalities such as sound is a workaround to this problem. A major challenge in auditory augmentation is the automatic generation of pleasant and ergonomic audio in complex routines, as opposed to overly simplistic feedback, to avoid fatigue. In this work, without loss of generality to other procedures, we propose a method for aural augmentation of ophthalmic procedures via automatic modification of musical pieces. Evaluations of this first proof of concept regarding recognizability of the conveyed information along with qualitative aesthetics show the potential of our method.

# Beveled needle position and pose estimation based on optical coherence tomography in ophthalmic microsurgery

Mingchuan Zhou, Kai Huang, Abouzar Eslami, Hessam Roodaki,
Haotian Lin, Chris P. Lohmann, Alois Knoll, M. Ali Nasseri

The needle position and pose is essential information in ophthalmic microsurgery. Estimation of these parameters is still an open problem, which hinders the application of various robot-assisted surgeries, such as autonomous injection. This paper presents a method to estimate the position and pose of beveled needle using three-dimensional Optical Coherence Tomography (OCT) image cube. The incomplete 3D point cloud of beveled needle obtained from OCT image cube is used to reconstruct the needle by fitting the projection of needle points based on its geometric features. We demonstrate the applicability of our method in ophthalmic injection surgery using ex vivo pig eyes and evaluate the accuracy of position and pose information using a micromanipulator.

## Precision Needle Tip Localization Using Optical Coherence Tomography Images for Subretinal Injection

Mingchuan Zhou, Kai Huang, Abouzar Eslami, Hessam Roodaki, Daniel Zapp, Mathias Maier, Chris P. Lohmann, Alois Knoll, M. Ali Nasseri

Subretinal injection is a delicate and complex microsurgery, which requires surgeons to inject the therapeutic substance in a preoperatively defined and intraoperatively updated subretinal target area. Due to the lack of subretinal visual feedback, it is hard to sense the insertion depth during the procedure, thus affecting the results of surgical outcome and hindering the widespread use of this treatment. This paper presents a novel approach to estimate the 3D position of the needle under the retina using the information from microscope-integrated Intraoperative Optical Coherence Tomography (iOCT). We evaluated our approach on both tissue phantom and ex-vivo porcine eyes. Evaluation results show that the average error in distance measurement is 4.7 $\mu$m (maximum of 16.5 $\mu$m). We furthermore, verified the feasibility of the proposed method to track the insertion depth of needle in robot-assisted subretinal injection.

## Surgical soundtracks: automatic acoustic augmentation of surgical procedures

Sasan Matinfar, M. Ali Nasseri, Ulrich Eck, Michael Kowalsky, Hessam Roodaki, Navid Navab, Chris P. Lohmann, Mathias Maier, Nassir Navab

*Purpose*. Advances in sensing and digitalization enable us to acquire and present various heterogeneous datasets to enhance clinical decisions. Visual feedback is the dominant way of conveying such information. However, environments rich with many sources of information all presented through the same channel pose the risk of over stimulation and missing crucial information. The augmentation of the cognitive field by additional perceptual modalities such as sound is a workaround to this problem. A major challenge in auditory augmentation is the

automatic generation of pleasant and ergonomic audio in complex routines, as opposed to overly simplistic feedback, to avoid alarm fatigue.

*Methods*. In this work, without loss of generality to other procedures, we propose a method for aural augmentation of medical procedures via automatic modification of musical pieces.

*Results*. Evaluations of this concept regarding recognizability of the conveyed information along with qualitative aesthetics show the potential of our method.

*Conclusion*. In this paper, we proposed a novel sonification method for automatic musical augmentation of tasks within surgical procedures. Our experimental results suggest that these augmentations are aesthetically pleasing and have the potential to successfully convey useful information. This work opens a path for advanced sonification techniques in the operating room, in order to complement traditional visual displays and convey information more efficiently.

# Real-time Scene Understanding in Ophthalmic Anterior Segment OCT Images

Hessam Roodaki, Matthias Grimm, Nassir Navab, Abouzar Eslami

*Purpose*. Machine learning algorithms are useful and efficient at interpreting medical images and segmenting anatomies. Here we present an approach that goes one step further by gaining scene understanding using cutting-edge machine learning techniques. Our method reliably detects anatomies of the anterior segment of the eye in OCT B-scans and implicitly understands the location of acquisition.

*Methods*. The utilized neural network architecture in our work is U-net, a convolutional classifier without any fully connected layers, with multiple modifications. Batch renormalization is introduced to increase training efficiency. Squeeze and excitation layers are added to improve interdependencies between channels. Dilated convolutions are also used to increase the receptive field of the network. Our design emphasizes on scene understanding to accurately learn the correct position of anatomies relative to each other. The ADAM algorithm is used for training with cross entropy loss as cost function. The neural network classifies input image pixels as one of cornea, sclera, iris or background classes. A spectral domain OCT system is used for data acquisition. An automated method is employed to capture random OCT B-scans with varying parameters such as size, scale, location and gain from ex-vivo porcine eyes as training dataset. Multiple images are captured from each location as a form of data augmentation. Annotation of the anatomies in the acquired dataset is initialized by multiple automated algorithms and then manually refined.

*Results*. In total 7503 training images and 1136 validation images are used. The network achieves an accuracy of 95.62% pixel classification over all classes and the entire validation dataset. Inference on a $1024 \times 1024$ OCT B-scan takes about 50 milliseconds.

*Conclusion*. We have presented a reliable method using machine learning for real-time OCT anatomy classification with acceptable accuracy. The algorithm succeeds in segmentation independent of the input image size, scale or location. However, to train a neural network effectively for clinical use, gathering large datasets of human subjects or domain adaptation is crucial.

# Simulated keratometry using microscope-integrated optical coherence tomography

Hessam Roodaki, Wei-Jun Chen, Daniel Zapp, Abouzar Eslami

*Purpose*. Intraoperative simulated keratometry has the potential to have a positive impact on anterior segment surgeries. Microscope-integrated Optical Coherence Tomography (OCT) is one of the modalities that can be employed for that purpose. This study compares astigmatism measurements by a gold standard method and measurements by an ophthalmic surgical microscope with a modified OCT engine in a porcine wet lab.

*Methods*. A Lumera 700 microscope with RESCAN 700 (Zeiss, Oberkochen, Germany) is modified and equipped with an automated algorithm to carry out simulated keratometry. The process starts with centering the OCT scan region to the corneal vertex. Then, an OCT volume covering a region of 6mm (width) × 6mm (height) × 2mm (depth) is captured from ex-vivo porcine eyes. Additional B-scans are captured and used for the correction of motions induced by environmental effects. The algorithm then segments the cornea in the motion-corrected volume. To measure the astigmatism of the corneal surface, anterior axial power map is calculated and the simulated keratometry values of the flattest and the steepest meridians are extracted. The difference between the two k-readings is considered as the measured astigmatism. The corneal topography and astigmatism of each porcine eye is also measured using an ATLAS 9000 corneal topography system (Zeiss, Dublin, USA).

*Results*. 20 porcine eyes are tested for the purpose of this study. To eliminate measurement inaccuracies, each eye is tested 5 times with each method and the median reading is used. 5 eyes are rejected due to inconsistent measurements. The calculated central astigmatism in each eye is compared to the reference measurements by ATLAS 9000. The readings have a correlation coefficient of 0.67. The observed deviation between the values by the proposed method and the reference method has a mean of 0.62, a standard deviation of 0.41, and a median of 0.57 (all in diopters).

*Conclusion*. Our results show that astigmatism measurements using microscope-integrated OCT is comparable to the measurements using the gold standard method with a correlation coefficient of 0.67. Further research is needed to examine the repeatability of measurements using the proposed method.

# Introducing Augmented Reality to Optical Coherence Tomography in Ophthalmic Microsurgery

C

Hessam Roodaki[1], Konstantinos Filippatos[2], Abouzar Eslami[2], Nassir Navab[1,3]

[1]Technische Universität München
[2]Carl Zeiss Meditec AG
[3]Johns Hopkins University

**Summary.** This publication proposes an augmented reality technique that is targeting posterior segment ophthalmic procedures such as vitrectomy and epiretinal membrane peeling. In these procedures, surgeons maneuver miniature surgical instruments around the retina observed under a surgical microscope. Since unnecessary touch of the retina is not desirable and the microscopic view of the surgical scene provides limited depth perception, such interventions tend to be challenging. Microscope-integrated optical coherence tomography has the potential to provide valuable depth information to complement surgeons' microscopic view. However, since surgical instruments are metallic, they cause total shadowing artifact in OCT images and only the top segment of the instrument is visible. This work proposes a novel method that can be employed to extract intrinsic cross-sectional information about the instrument tip. Using this information, the view of the surgeon can be augmented to eliminate the effect of the shadowing artifact. Furthermore, the distance between the tip of the instrument to the retina is reported using color-coding. The second novel contribution of this work is in introducing augmented reality to OCT. By augmenting the OCT view of the surgeon as opposed to the microscopic view, not only OCT is elevated from a secondary imaging modality to a primary source of information, but also distraction from the main focus area is avoided. The technique starts with tracking the tip of the instrument in the microscopic view. This allows two OCT scans to be continuously captured in a cross configuration at the tip of the instrument. In each OCT scan instrument top segments and shadows are then extracted. Instrument cross-sections are geometrically estimated using the extracted information from OCT. This information is overlaid into the OCT view of the surgeon using color codes. The accuracy in distance estimation of the proposed method is evaluated using a micromanipulator for different tools and positions in many trials.

**Contribution.** The contribution of the author of this thesis to this publication consists of the main idea, implementation, experiment design and evaluation. Writing of the manuscript was done and coordinated by the author of this thesis. Refinement of the details in implementation as well as the revision of the publication was done jointly with the co-authors.

# Introducing Augmented Reality to Optical Coherence Tomography in Ophthalmic Microsurgery

Hessam Roodaki*
Technische Universität München

Konstantinos Filippatos†
Carl Zeiss Meditec AG

Abouzar Eslami‡
Carl Zeiss Meditec AG

Nassir Navab§
Technische Universität München
and Johns Hopkins University

## ABSTRACT

Augmented Reality (AR) in microscopic surgery has been subject of several studies in the past two decades. Nevertheless, AR has not found its way into everyday microsurgical workflows. The introduction of new surgical microscopes equipped with Optical Coherence Tomography (OCT) enables the surgeons to perform multimodal (optical and OCT) imaging in the operating room. Taking full advantage of such elaborate source of information requires sophisticated intraoperative image fusion, information extraction, guidance and visualization methods. Medical AR is a unique approach to facilitate utilization of multimodal medical imaging devices. Here we propose a novel medical AR solution to the long-known problem of determining the distance between the surgical instrument tip and the underlying tissue in ophthalmic surgery to further pave the way of AR into the surgical theater. Our method brings augmented reality to OCT for the first time by augmenting the surgeon's view of the OCT images with an estimated instrument cross-section shape and distance to the retinal surface using only information from the shadow of the instrument in intraoperative OCT images. We demonstrate the applicability of our method in retinal surgery using a phantom eye and evaluate the accuracy of the augmented information using a micromanipulator.

**Keywords:** Augmented reality, optical coherence tomography, ophthalmic surgery, instrument cross-section.

**Index Terms:** H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

## 1 INTRODUCTION

Ophthalmic posterior segment surgery is a complex manual task performed under a surgical stereo microscope that involves manipulation of delicate retinal tissue. Surgical accuracy required in a microsurgery task is in the scale of tens of micrometers which is at or close to the limit of human positional ability. The interactions between the surgical instrument and the retinal membranes are also below the limit of human tactile sensation. Limited positional and tactile sensation in a vitreoretinal surgery leads to the reliance of the surgeon primarily on the microscopic visual feedback which is known to dramatically reduce the manipulation quality, speed and accuracy [14]. Perceiving depth and instrument-tissue proximity to avoid retinal damage is a challenging burden for the surgeon. For this reason, development of assistive algorithms such as retinal surgery instrument detection or proximity estimation in microscopic images has become a research interest in the computer vision community [13, 18]. Such a task requires real-time

---

*e-mail: he.roodaki@tum.de
†e-mail: konstantinos.filippatos@zeiss.com
‡e-mail: abouzar.eslami@zeiss.com
§e-mail: navab@in.tum.de

tracking under challenging conditions such as lens distortions, rapid instrument motions, partial occlusions and illumination variations imposed by handheld endo-illumination probe.

Compared to conventional microscopic imaging, Optical Coherence Tomography (OCT) is a fundamentally different imaging modality that has become a major non-invasive tool in ophthalmology. OCT B-scans provide real-time cross-sectional images of retina with resolutions close to that of histopathology and higher than any other non-invasive imaging modality. Diagnosis of chorioretinal diseases and monitoring morphological retinal changes over time using OCT are well-established approaches [5, 6]. Until recently, OCT scanners were available only in tabletop form factors and restricted to the diagnostic setting. Faster image acquisition by Spectral Domain OCT (SD-OCT) led to introduction of handheld OCT scanners. However, their inability to visualize ocular structures in real-time during surgical maneuvers has limited their intraoperative applicability [15]. Since optical microscopy and OCT can share the same optical path in a device, integration of OCT and surgical microscope was a natural next step to address limitations in maneuver, accuracy and speed associated with handheld OCT scanners [11, 20]. Microscope-mounted intraoperative OCT (iOCT) solution by Carl Zeiss Meditec (*RESCAN 700*) with foot pedal control of the OCT and real-time overlay on microscopic view using Heads-Up Display (HUD) has been the first in this category to be described in human use [9, 10].

Integration of OCT into the operating room has a significant impact on surgical decision-making by providing information regarding surgical maneuvers and their effect on tissues [8]. Examples include membrane identification and removal confirmation in Epiretinal Membrane (ERM) peeling procedure [12, 4] and intraoperative measurement of the anterior lens capsule position [16]. Intraoperative application of OCT has also limitations including latency and positional drifts. Also, optical properties of routinely used metallic instruments include high reflectivity resulting in total shadowing and obstructing the view of the underlying tissues in OCT images. By turning a limitation into an opportunity, we present a novel method to use information provided by reflections and shadows of surgical instruments to estimate the instrument cross-section shape and its distance to retina. Providing this extra information helps in perceiving instrument-tissue proximity to avoid unnecessary touch.

The method of choice for the visualization of extra cross-sectional and depth information is a key element of this work. Trials of Augmented Reality (AR) in the operating room show its ability to seamlessly integrate into the surgical workflow [17]. Since vitreoretinal surgery is performed under a surgical microscope, augmenting the surgeon's view with additional information is a reasonable approach. Augmented stereoscopic microscope has already been used both for neuro- and ophthalmo-surgical interventions [3, 7, 19]. Also preliminary use of an overlay system to display preoperative data during a retinal microsurgical intervention has shown to improve the gesture accuracy and reduce the operating time [13]. In this paper, we introduce AR to intraoperative OCT in conventional vitreoretinal microsurgery which to our knowledge is the first of its kind.

## 2 METHODS

Our setup is shown in fig.1. It consists of an *OPMI LUMERA 700* microscope equipped with the *RESIGHT* fundus imaging system and an integrated *RESCAN 700* OCT engine connected to a *CALLISTO eye* assistance computer system (all from Carl Zeiss Meditec, Jena, Germany). A phantom eye was fixated under the optical pathway of the microscope. An artificial lens was attached to the phantom eye which optically mimicked a human eye lens. To simulate a realistic surgical setting, two 23-gauge sclera trocars were inserted into the phantom eye. A wide-angle 23-gauge endo-illumination probe was inserted into the phantom eye through the left trocar and the right trocar was used for insertion of the surgical instruments. The *OPMI LUMERA 700* microscope was focused on the retinal surface of the phantom eye while the *RESCAN 700* was configured to provide 2 real-time orthogonal OCT B-scans in form of a cross. The OCT B-scans were visualized both on the *CALLISTO eye* display and directly into the eyepiece of the microscope via HUD.
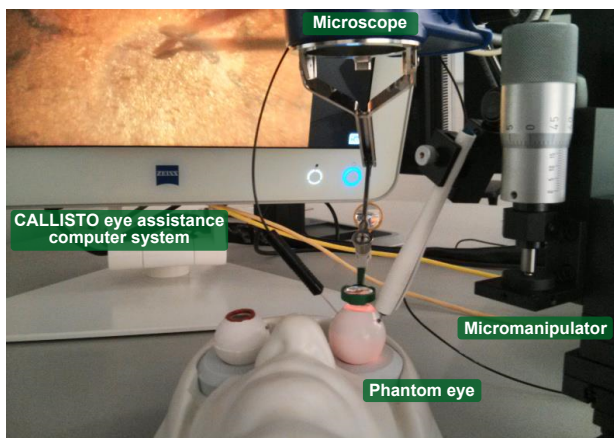


Figure 1: Our retinal surgery simulation setup. The surgical microscope is equipped with a fundus imaging system and a live OCT engine. The microscope is connected to an assistance computer system. A phantom eye is put under the optical pathway of the microscope. For evaluation purposes, the instruments are positioned using a micromanipulator.

Vitreoretinal instruments have tips with either one or two parts and are generally close to cylindrical in folded state for the ease of insertion. Instruments used in this preliminary work are considered of low geometrical complexity and include diamonized end-gripping vitreoretinal forceps, ILM end-gripping vitreoretinal forceps from various producers and posterior vitrectomy probe (Carl Zeiss Meditec, Jena, Germany). The common feature of these instruments is their cross-section being very close to either an ellipse or half-ellipses.

The basic principle behind our method of augmenting the surgeon's view of the OCT images with the instrument cross-section is illustrated in fig.3. Visualization of the instrument cross-section requires estimation of several instrument parameters such as its radius and orientation. These parameters are estimated based on the reflection of the instrument in the two perpendicular OCT images. For this purpose, the OCT scan position must be continuously kept at the instrument tip. Also the orientation of the scan must be such that one OCT B-scan is acquired perpendicular to the instrument and the other in the direction and over the instrument. Thus, the algorithm begins with the tracking of the instrument in the microscope image stream. For tracking the instrument in retinal images, we use the method described in detail in [2] which provides the instrument tip and its orientation.

The second step is the detection of the retinal surface, the instrument reflections and their corresponding shadows in both OCT B-scans. Differentiation of the vitreous humor from the anatomical or artificial structures such as the retinal surface, the Internal Limiting Membrane (ILM) or the instrument reflections is done by thresholding the OCT images. The threshold value is adaptively defined based on statistical measures of each B-scan. This simple yet effective method has been used and evaluated in automatic segmentation of structures in OCT images in [1]. As a later step, morphological filtering is applied to eliminate thin vitreous membranes detected in the OCT images. Then the topmost surface is segmented and considered as the retinal surface. Scanning from left to right, any vertical jump and drop in this surface layer is detected and considered as the beginning and the end of an instrument reflection respectively. To avoid misdetection of anatomical features of retina as a reflection caused by a surgical instrument, only reflections with no visible structures below are confirmed as instrument reflections (fig.2b). For estimation and visualization of the cross-section, the instrument tip point in the OCT along the instrument and the topmost point or the mean of the topmost points of the instrument reflection in the OCT perpendicular to the instrument is extracted.

A cylindrical surgical instrument has a rectangular and an elliptical cross-section along its shaft and perpendicular to its shaft respectively. Augmenting both OCT B-scans with the cross-section of the instrument requires 5 parameters to estimate. The ellipse parameters are the center point $(C_x, C_y)$, the major and minor axes $(a, b)$ and the orientation angle $(\phi)$. The rectangle parameters are the top-left point $(T_x, T_y)$, the sides $(b, l)$ and the orientation angle $(\theta)$ as seen in fig.4.
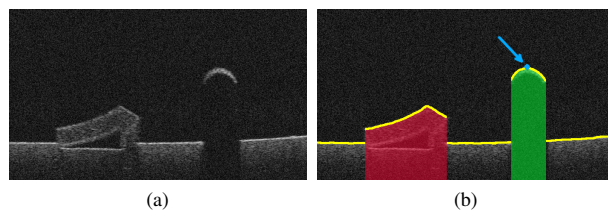


Figure 2: (a) A posterior segment OCT B-scan of a phantom eye featuring instrument shadowing and an abnormality on the retinal surface. (b) Detected features in the sample OCT image. Cuts in the topmost surface (yellow) are considered as instrument reflection candidates (red and green regions), but only reflections with no visible structures underneath get confirmed (green). The peak point of the instrument reflection (blue) is also detected.

The tracking of the instrument in the microscopic view guarantees one OCT image to be along the instrument shaft and the other perpendicular to that. Hence, the ellipse orientation angle $(\phi)$ is always zero and the orientation angle of the rectangle $(\theta)$ is the instrument orientation angle which can be estimated based on the slope of the instrument reflection in the OCT image along the instrument. Moreover the width of the instrument shadow in the perpendicular OCT equals the diameter of the instrument $(2r)$. Which quickly gives the shorter side of the rectangular cross-section and the elliptical cross-section minor axis $(b)$. Based on the cylindric section formulation, the following holds:

$$b = 2r, \qquad a = 2r \sec\theta \qquad (1)$$

The top-left point of the rectangular cross-section $(T)$ could be set to the instrument tip point detected in the OCT over the instrument. To estimate the center point of the ellipse, the detected peak point
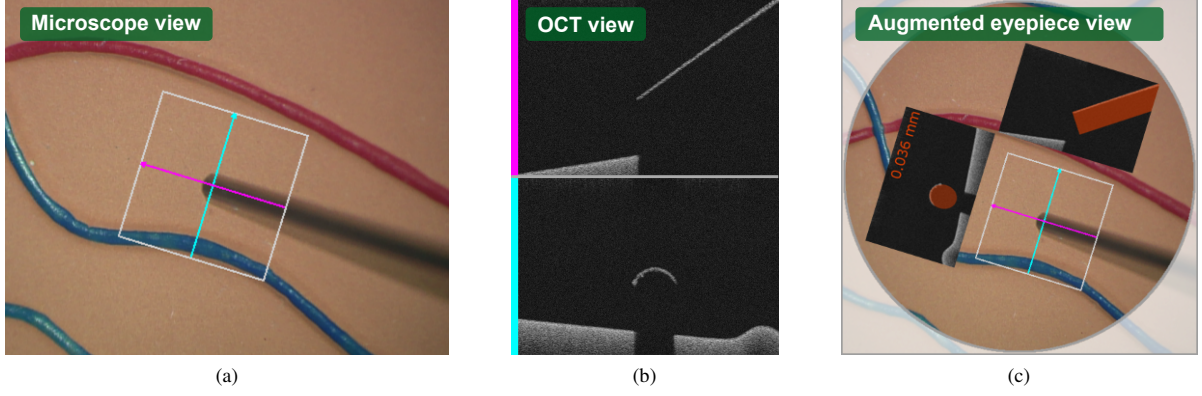
Figure 3: General instrument cross-section estimation scheme. (a) Tip of the instrument and its orientation is continuously detected in the microscopic view. (b) Two real-time OCT scans are constantly acquired over the instrument and perpendicular to the instrument at the instrument tip. Estimation of the instrument cross-sectional information is performed based on the instrument reflections in the two OCT scans. (c) The cross-sectional information is augmented into the OCT images and are directly put into the eyepiece of the microscope.

($P$) of the instrument reflection in the OCT perpendicular to the instrument can be used as follows:

$$C_x = P_x$$
$$C_y = P_y + \frac{a}{2} \qquad (2)$$

Finally, since the OCT scans are centered on the instrument tip, the instrument is generally considered to be extended towards outside of the longer side of the rectangle. Thus the longer side of the rectangular cross-section ($l$) is set such that the rectangle is extended to the edge of the OCT image in both top and right edges.
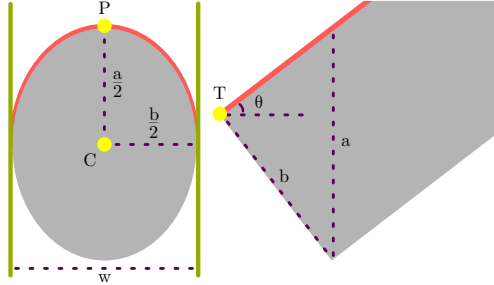


Figure 4: Cross-sections of a cylindrical instrument held at $\theta$ degrees. The upper red segments are the instrument reflections visible in OCT images, $P$ and $w$ are the instrument reflection peak point and shadow width in the perpendicular OCT image. The parameters to estimate are the instrument angle ($\theta$), the ellipse axes ($a$, $b$), the ellipse center ($C$) and the rectangle top-left point ($T$).

Estimation of the cross-section in instruments with two-part tips (e.g. forceps) in the perpendicular OCT needs further calculations. Such instruments in their most basic form have cylindrical shafts, but cylindrical tips cut into halves along the axis of the shaft. Hence, the cross-sections to estimate in the OCT perpendicular to the instrument are half-ellipses. The instrument tracking provides the projection of the two parts of the tip on the instrument centerline as the instrument tip. Thus the instrument casts one shadow in one OCT scan which is treated the same as before to estimate the instrument slope and two shadows in the other OCT image. Although all the parameters of both half-elliptical cross-sections

are the same, to avoid propagation of the estimation error, each shadow is considered independent from the other. By doing so, one pair of independent parameters is estimated for the half-elliptical cross-sections (fig.5). This consideration guarantees less estimation error in case of misalignment of the OCT acquisition direction and the instrument caused by errors in instrument tracking.

In contrast to instruments with single-part tips (e.g. vitrectomy probe), a pair of forceps can also be tilted along the shaft axis. This instrument tilt angle ($\phi$) must be taken into account for accurate cross-section estimation. The instrument tilt angle can be determined using the difference in the position of the two instrument reflection topmost points in the OCT image. This estimation of the instrument tilt is accurate since the OCT image is acquired perpendicular to the instrument shaft. The estimation of the axes of the ellipses is done with further considerations. The ellipses are now rotated by $\phi$ degrees. Moreover, since only half of the elliptical cross-section actually casts a shadow in the OCT image, the shadow width of each tip is extended to a value which would be covered if the tip was fully cylindrical. The concept is illustrated in fig.5. The width of each of the elliptical cross-sections axis-aligned bounding-box is relative to the ellipse axes. If both tips of the instrument were full cylinders, the radius of each one could be defined as:

$$w_i + e_i = \sqrt{a_i^2 \sin^2 \phi + b_i^2 \cos^2 \phi}$$
$$b_i = 2r_i \qquad (3)$$
$$a_i = 2r_i \sec \theta, \qquad i = 1, 2$$

where $i$ is the tip index. As illustrated in fig.5, the extended shadow length ($w_i + e_i$) can be estimated based on each ellipse semi-major axis and the ellipse rotation angle:

$$w_i + e_i = 2(e_i + x_i)$$
$$x_i = \frac{a_i}{2} \sin \phi, \qquad i = 1, 2 \qquad (4)$$

By substituting $w_i + e_i$ from equation 4 into equation 3, the radius of each hypothetical cylindrical tip can be derived as follows:

$$r_i = \frac{w_i}{\sqrt{\frac{\sin^2 \phi}{\cos^2 \theta} + \cos^2 \phi + \frac{\sin \phi}{\cos \theta}}}, \qquad i = 1, 2 \qquad (5)$$

which immediately gives the axes of the ellipses. The ellipses are then cut into halves according to the tilt of the instrument. Finally,

each ellipse center point is defined as:

$$C_{i_x} = S_{i_x} + \frac{w_i + e_i}{2}$$
$$C_{i_y} = P_{i_y} + \frac{1}{2}\sqrt{a_i^2 \cos^2 \phi + b_i^2 \sin^2 \phi}, \quad i = 1,\ 2$$

(6)

in which $P_i$ is the detected peak point. $S_{i_x}$ is the start of the occurrence of the $i^{th}$ shadow in the x-axis in the perpendicular OCT. In other words, $S_{i_x}$ is the $x$ element of the leftmost points of instrument reflections in the perpendicular OCT.
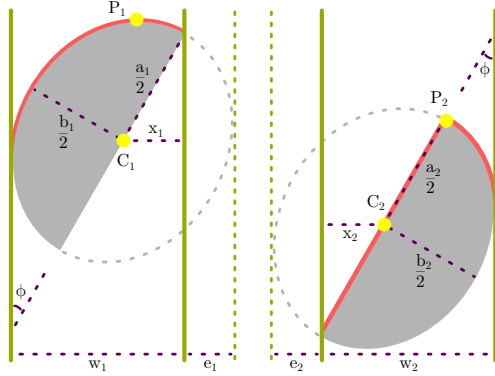


Figure 5: Cross-sections of a cylindrical instrument with a two-part tip held at $\theta$ degrees in the perpendicular OCT. The upper red segments are the instrument reflections visible in the OCT image. Estimation of the half-elliptical cross-sections, in excess of the parameters in the case of instruments with single-part tips, requires the extended shadow ($w_i + e_i$) and the instrument tilt angle ($\phi$).

In either case of instruments with single-part tips or instruments with two-part tips, an estimation of the distance between the instrument tip and the underlying tissue is presented to the surgeon in real-time. For the estimation of the distance between the instrument tip and the retinal surface, the bottommost point of the estimated cross-section is needed. This point is straightforwardly determined based on the instrument reflection topmost point and the height of the axis-aligned bounding box of the estimated cross-section ellipse or half-ellipse. The difference between the mean of the retinal surface height detected immediately before and after the shadow region and the cross-section bottommost point is the desired distance to report. Since the pixel size of the OCT device is known, the distance in pixels is converted to the actual distance and is reported with micrometer resolution.

The augmentation of the surgeon's view is done via overlaying the estimated cross-sections and the estimated distance between the instrument tip and retina over the OCT images acquired at the instrument tip. Since concentration of the surgeon on the distance value is not feasible during surgery, a color coding mechanism is employed to facilitate the presentation of the estimated distance. The color of the instrument cross-sections are chosen from green to red spectrum based on the distance between the instrument and the retinal surface. Green represents $1mm$ or more and red is for $0.2mm$ or less. Also, the cross-sections are shown with less opacity if they are closer than $0.3mm$ to retina to avoid deviating the focus of the surgeon from the structure over retina which is under manipulation. These values are determined based on ex vivo experiments on pig eyes and geometrical footprints of the most typical vitreoretinal instruments used in the most common posterior manipulations such as ERM peeling and Vitrectomy.

The final synthesized image consists of a cross indicating the position of the two OCT scans and the augmented OCT images

(fig.3c). The OCT images are put at the top and left side of the cross if the instrument enters the scene from bottom right and at the top and right side if the instrument enters from bottom left. The synthesized image is then displayed in the HUD system which is optically coupled into the field-of-view (FOV) of one microscope ocular lens. Since the HUD is a half-mirror, the image is overlaid into the microscopic view. The HUD system of the microscope is internally registered to the scene under view. Thus, by providing the synthesized images in the same pixel size of the microscope, the overlay is correctly registered to the microscopic scene. An intraoperative OCT device requires thorough calibration between the OCT acquisition engine and the microscope. The HUD is also calibrated to the microscope FOV. The proposed method requires no further calibration. The instrument cross-section estimation and the augmentation is performed in real-time. Hence, the surgeon constantly gets feedback of the performed maneuvers.

## 3 RESULTS

The method is successfully employed to augment the OCT image stream of the described setup with the extra cross-sectional information (fig.6). The OCT B-scans used in our experiments are 8-bit, $1024 \times 1024$ pixels representing $3mm$ in width and $2mm$ in depth. The microscopic images used for instrument tracking are 24-bit, $720 \times 540$ pixels representing $8mm$ in width and $6mm$ in height. The microscope images along with the OCT B-scans are acquired at the rate of 25 frames per second. The computation and memory requirements of the developed algorithm are far above real-time limits. Although the method is intended to be an assistive tool in posterior ophthalmic procedures, tests in the anterior setup show the same overall results in terms of user experience and performance. A video demonstration of the proposed method is provided as supplementary material.
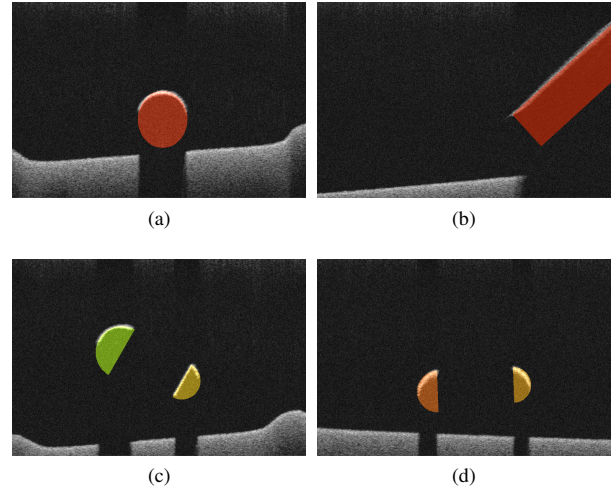


(a)

(b)

(c)

(d)

Figure 6: Examples of augmented OCT images in a simulated setup. (a) is augmented perpendicular OCT image of a vitreoretinal probe, (b) is augmented parallel OCT image of a diamonized forceps and (c) and (d) are augmented perpendicular OCT images of an ILM end-gripping forceps. The color of the cross-sections shifts towards red as the instruments get closer to the retinal surface.

To evaluate the accuracy of the proposed method, a relative measurement approach was applied in a simulated posterior setup. The detection accuracy of the instrument reflection topmost point in the perpendicular OCT as well as the estimation accuracy of the distance between the instrument tip and the

Table 1: Instrument Reflection Topmost Point Detection and Instrument-Retina Distance Estimation Accuracy

| Instrument | Real movement in X and Z $[\mu m]$ | Trial frames | Estimated topmost point movement in X $[\mu m]$ | Estimated topmost point movement in Z $[\mu m]$ | Estimated distance change $[\mu m]$ |
|---|---|---|---|---|---|
| Posterior vitrectomy probe | 5 | 82 | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) |
| ILM end-gripping forceps* | 5 | 394 | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) |
| Diamonized forceps* | 5 | 181 | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) |
| Posterior vitrectomy probe | 10 | 122 | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) |
| ILM end-gripping forceps* | 10 | 70 | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) |
| Diamonized forceps* | 10 | 173 | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) | (box-and-whisker plot, axis 0–18) |

*For the forceps the mean of the estimated values for the two parts of the tip are considered.

retinal surface was evaluated using a 3-axes $HS$-6 manual micromanipulator (Märzhäuser Wetzlar, Wetzlar, Germany). The employed micromanipulator has a resolution of 5 micrometers in each axis. In different trials, a 25-gauge posterior vitrectomy probe and two different types of 23-gauge vitreoretinal forceps were attached to the manipulator. The manipulator was positioned such that one axis of the manipulator was in the OCT depth direction and the other two axes in the height and width orientations of the microscope image plane. To avoid any interference in the movement of the instruments, instead of a trocar, a larger hole was created in the sclera region of the phantom eye. By accurately changing the position of the instrument in the orientation of the OCT width and depth, the relative change in the topmost point position detection and the distance to retina estimation was observed. For each instrument, a multitude of different trials at different locations with different distances between the instrument tip and retina were performed to suppress any effect caused by varying reflectivity of the phantom eye or the instruments. At each trial, several OCT frames were grabbed for evaluation. Also, instead of the instrument tracking algorithm, the position and orientation of the OCT scans were manually set according to the instrument to avoid the effect of the tracking error.

The evaluation results are presented in table 1. Each row of the table is representing the combination of a particular instrument and a fixed movement of $0.005mm$ or $0.01mm$ of the manipulator in the orientation of the OCT width and depth. The estimated topmost point movement and the estimated change in the distance between the instrument and the retinal surface is presented using box-and-whisker plots. In all of the plots, the whiskers are showing the minimum and maximum estimated change while the start and the end of the box are the first and third quartiles. The band and the dot inside the box are the median and mean of the estimated change respectively. The actual value against which the estimation should be compared is highlighted in blue on the horizontal axis of the plots. For the case of forceps, the mean of the estimated values for the two parts of the tip are plotted. The quantitative evaluation of the proposed AR solution shows that the maximum error in the estimation of the distance to retina is 9.76 micrometers which is the equivalent of 5 pixels in the OCT images of the test dataset.

Errors in detection of the topmost point induce error in estimation of the distance to retina. This is noticeable particularly in case of the posterior vitrectomy probe where a wide interquartile range in estimation of the topmost point movement leads to a wide interquartile range in the estimated distance change. In general, the more non-uniform the reflection from the instrument surface is, the wider the range between the minimum and maximum whiskers get. Hence, the observed wide minimum to maximum range is caused by variant reflections of the instrument in different OCT scans. In case of instruments with two-part tips, when the instrument is in the folded state it is considered to be an instrument with a single-part tip. Although this modeling causes very little error on its own, the irregular reflection due to the complex shape causes errors in detection of the topmost point.

## 4  CONCLUSION

In this paper, we introduced a novel method for augmenting the surgeon's view of intraoperative OCT images in ophthalmic surgery with instrument cross-section. We demonstrated that the information provided via the reflectivity characteristic of metallic vitreoretinal instruments in intraoperative OCT imaging can be employed to provide the surgeon with additional information regarding his maneuvers. Although the proposed method is an initial work in this direction, it shows the vast range of its applications and has the potential to be used in real clinical settings. Injecting extra information to the OCT images extends the reach of augmented reality to a new limit. The experiments demonstrate the robustness and the potential of the current approach for usage both in anterior and posterior ophthalmic setups. In [14], Gupta et al. have demonstrated that the majority of retinal surgery is performed without the surgeon being able to feel instrument-tissue interactions since most of the manipulations are below the human tactile sensation threshold. Hence, an estimation of the distance of the instrument tip to retina is valuable to avoid unnecessary touch of the delicate retinal tissue and the proposed AR paradigm is a suitable platform for providing this additional information. Also, the same study indicates that movements below tens of micrometers are believed to be below the limit of human positional ability. Thus, the proposed medical augmented reality method can be considered

as an assistive intraoperative tool considering its maximum error in estimation of the instrument to retina distance.

It is evident from the results that a robust augmentation of the parallel OCT is challenging in some cases due to the irregular shape of the instrument reflection. Hence, future work will direct towards extending the current approach for challenging instrument reflections and instruments with more complex geometries such as scissors. Having the 3D model of the instrument as prior knowledge would be one way to deal with instruments with complex shapes or augmenting the parallel OCT more smoothly. Also removing the reliability of the procedure on instrument tracking could be an applicable extension to the method. This would be particularly beneficial when the OCT is needed to be fixed on a specific location. Finally, design and application of a more sophisticated evaluation process to assess the accuracy of the estimated cross-section parameters such as the estimated radius of the instrument is of great importance. Instruments with known geometrical parameters at every location of the tip are required for such a task. A realistic clinical test to get feedback from surgeons and evaluation of the method in more realistic scenarios is also of great benefit.

## REFERENCES

[1] C. Ahlers, C. Simader, W. Geitzenauer, G. Stock, P. Stetson, S. Dastmalchi, and U. Schmidt-Erfurth. Automatic segmentation in three-dimensional analysis of fibrovascular pigmentepithelial detachment using high-definition optical coherence tomography. *British Journal of Ophthalmology*, 92(2):197–203, 2008.

[2] M. Alsheakhali, M. yigitsoy, A. Eslami, and N. Navab. Surgical tool detection and tracking in retinal microsurgery. In *Proceedings of SPIE Medical Imaging*. SPIE, 2015.

[3] J. W. Berger and D. S. Shin. Computer-vision-enabled augmented reality fundus biomicroscopy. *Ophthalmology*, 106(10), 1999.

[4] S. Binder, C. I. Falkner-Radler, C. Hauger, H. Matz, and C. Glittenberg. Feasibility of intrasurgical spectral-domain optical coherence tomography. *Retina*, 31(7):1332–1336, 2011.

[5] R. A. Costa, M. Skaf, L. A. M. Jr., D. Calucci, J. A. Cardillo, J. C. Castro, D. Huang, and M. Wojtkowski. Retinal assessment using optical coherence tomography. *Progress in Retinal and Eye Research*, 25(3):325 – 353, 2006.

[6] W. Drexler and J. G. Fujimoto. State-of-the-art retinal optical coherence tomography. *Progress in Retinal and Eye Research*, 27(1):45 – 88, 2008.

[7] P. Edwards, A. King, J. Maurer, C.R., D. De Cunha, D. Hawkes, D. Hill, R. Gaston, M. Fenlon, A. Jusczyzck, A. Strong, C. Chandler, and M. Gleeson. Design and evaluation of a system for microscope-assisted guided interventions (magi). *Medical Imaging, IEEE Transactions on*, 19(11):1082–1093, Nov 2000.

[8] J. P. Ehlers, W. J. Dupps, P. K. Kaiser, J. Goshe, R. P. Singh, D. Petkovsek, and S. K. Srivastava. The prospective intraoperative and perioperative ophthalmic imaging with optical coherence tomography (pioneer) study: 2-year results. *American Journal of Ophthalmology*, July 2014.

[9] J. P. Ehlers, P. K. Kaiser, and S. K. Srivastava. Intraoperative optical coherence tomography using the rescan 700: preliminary results from the discover study. *British journal of Ophthalmology*, 98(10):1329–1332, 2014.

[10] J. P. Ehlers, S. K. Srivastava, D. Feiler, A. I. Noonan, A. M. Rollins, and Y. K. Tao. Integrative advances for oct-guided ophthalmic surgery and intraoperative oct: Microscope integration, surgical instrumentation, and heads-up display surgeon feedback. *PLoS ONE*, 9(8):e105224, 08 2014.

[11] J. P. Ehlers, Y. K. Tao, S. Farsiu, R. Maldonado, J. A. Izatt, and C. A. Toth. Integration of a spectral domain optical coherence tomography system into a surgical microscope for intraoperative imaging. *Investigative Ophthalmology & Visual Science*, 52(6):3153–3159, 2011.

[12] J. P. Ehlers, Y. K. Tao, and S. K. Srivastava. The value of intraoperative optical coherence tomography imaging in vitreoretinal surgery. *Current Opinion in Ophthalmology*, 25(3):221–227, 2014.

[13] I. Fleming, S. Voros, B. Vagvolgyi, Z. Pezzementi, J. Handa, R. Taylor, and G. Hager. Intraoperative visualization of anatomical targets in retinal surgery. In *Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on*, pages 1–6, Jan 2008.

[14] P. K. Gupta, P. S. Jensen, and J. de Juan, Eugene. Surgical forces and tactile perception during retinal microsurgery. In C. Taylor and A. Colchester, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI99*, volume 1679 of *Lecture Notes in Computer Science*, pages 1218–1225. Springer Berlin Heidelberg, 1999.

[15] P. Hahn, J. Migacz, R. O'Connell, R. S. Maldonado, J. A. Izatt, and C. A. Toth. The use of optical coherence tomography in intraoperative ophthalmic imaging. *Ophthalmic surgery, lasers & imaging : the official journal of the International Society for Imaging in the Eye*, 42 Suppl:S8594, July 2011.

[16] N. Hirnschall, S. Norrby, M. Weber, S. Maedel, S. Amir-Asgari, and O. Findl. Using continuous intraoperative optical coherence tomography measurements of the aphakic eye for intraocular lens power calculation. *British journal of Ophthalmology*, 2014.

[17] N. Navab, T. Blum, L. Wang, A. Okur, and T. Wendler. First deployments of augmented reality in operating rooms. *Computer*, 45(7):48–55, July 2012.

[18] R. Richa, M. Balicki, R. Sznitman, E. Meisner, R. Taylor, and G. Hager. Vision-based proximity detection in retinal surgery. *Biomedical Engineering, IEEE Transactions on*, 59(8):2291–2301, Aug 2012.

[19] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays: A literature review of augmented reality. *Display Technology, Journal of*, 4(4):451–467, Dec 2008.

[20] Y. K. Tao, S. K. Srivastava, and J. P. Ehlers. Microscope-integrated intraoperative oct with electrically tunable focus and heads-up display for imaging of ophthalmic surgical maneuvers. *Biomedical optics express*, 5(6):18771885, June 2014.

# SonifEye: Sonification of Visual Information using Physical Modeling Sound Synthesis

Hessam Roodaki[1,3], Navid Navab[2], Abouzar Eslami[3], Christopher Stapleton[4], Nassir Navab[1,5]

[1] Computer Aided Medical Procedures, Technische Universität München, Munich, Germany
[2] Topological Media Lab, Concordia University, Montreal, QC, Canada
[3] Carl Zeiss Meditec AG, Munich, Germany
[4] Simiosys Real World Laboratory, Oviedo, FL, USA
[5] Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD, USA

**Summary.** This publication proposes an auditory augmented reality technique that is targeting high-precision focus demanding tasks. The motivation behind using sound for information conveyance is that visual augmented reality could be distracting. Sonification is an alternative method with advantages such as bringing about faster reaction times and omnidirectionality. The primary contribution of this work is in employing auditory augmented reality for conveying information that is extracted from visual sensors for the purpose of guidance in micron-level tasks. By doing so, the visual field of subjects performing such tasks are not cluttered with excessive visual information. The secondary novelty in this work, is the use of physical modeling for sound synthesis. Employing physical modeling, provides the ability to generate sounds that resemble natural phenomena flexibly. This leads into less mental effort for interpretation which is a necessity in high-precision tasks. The technique proposed here starts with receiving information from visual sensors. This information is passed as parameters into mathematical functions that generate sounds that for instance resemble tapping on a plate made from a specific material in intervals. When visual parameters change, parameters defining sound such as the material of the plate change accordingly providing intrinsic cues. In this paper three examples are given describing effective methods to map visual information to non-speech sound guidance. The effect of using auditory AR is compared against using visual AR in two controlled experiments. Results from both experiments indicate that using auditory augmented reality can decrease some errors and have no effects on others. When both visual and auditory AR techniques are used however, errors tend to increase as users fallback to visual cues for guidance and lose focus.

**Contribution.** The contribution of the author of this thesis to this publication consists of the main idea, the majority of the implementation, experiment design and evaluation. Writing of the manuscript was done and coordinated by the author of this thesis. Sound designs, discussions on the findings and the revision of the publication was done jointly with the co-authors.

# SonifEye: Sonification of Visual Information using Physical Modeling Sound Synthesis

Hessam Roodaki, Navid Navab, Abouzar Eslami, Christopher Stapleton, and Nassir Navab

**Abstract**— Sonic interaction as a technique for conveying information has advantages over conventional visual augmented reality methods specially when augmenting the visual field with extra information brings distraction. Sonification of knowledge extracted by applying computational methods to sensory data is a well-established concept. However, some aspects of sonic interaction design such as aesthetics, the cognitive effort required for perceiving information, and avoiding alarm fatigue are not well studied in literature. In this work, we present a sonification scheme based on employment of physical modeling sound synthesis which targets focus demanding tasks requiring extreme precision. Proposed mapping techniques are designed to require minimum training for users to adapt to and minimum mental effort to interpret the conveyed information. Two experiments are conducted to assess the feasibility of the proposed method and compare it against visual augmented reality in high precision tasks. The observed quantitative results suggest that utilizing sound patches generated by physical modeling achieve the desired goal of improving the user experience and general task performance with minimal training.

**Index Terms**—Aural augmented reality, sonification, sonic interaction, auditory feedback

◆

## 1 INTRODUCTION

Augmented Reality (AR) as a means of human computer interaction and information conveyance has been explored in studies with various foci. As techniques and technologies involved in AR progress, more applications for methods revolving around the idea of interaction through augmented environments emerge. AR has found its way even in the more complex set of professions including medical procedures where extreme levels of precision and reliance are necessary [7]. However, a quick review of the publications in the AR community reveals the large gap between visual augmented reality and augmentation via other human sensory modalities including haptics and audio. There has been attempts to examine the feasibility of employing multisensory perception in complex human performance procedures. These efforts include increasing situational awareness of human subjects via tactile feedback [1]. However, the full potential of multisensory augmented reality and information conveyance by means of multisensory perception has not been exploited to the benefit of assistance in complex procedures. Considering the achievements of visual AR, multisensory AR is a proper candidate for scenarios with multivariate sources of information. There are unique qualities of hearing separate from all other senses that make it the communication channel of choice for many applications. Audition being omnidirectional, provides the ability to gather information beyond spatial restrictions of other senses. Hearing gives a subject the capability to articulate the nuances of complex sonic environments.

---

- *Hessam Roodaki is with the Chair for Computer Aided Medical Procedures, Technische Universität München, 85748 Munich, Germany and also Carl Zeiss Meditec AG, 81379 Munich, Germany (e-mail: he.roodaki@tum.de).*
- *Navid Navab is with the Topological Media Lab, Concordia University, Montreal, QC, H3G 2W1 Canada (e-mail: navid.nav@gmail.com).*
- *Abouzar Eslami is with Carl Zeiss Meditec AG, 81379 Munich, Germany (e-mail: abouzar.eslami@zeiss.com).*
- *Christopher Stapleton is with Simiosys Real World Laboratory, Oviedo, FL 32765 USA (e-mail: cstapleton@simiosys.com).*
- *Nassir Navab is with the laboratory for Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA and also the Chair for Computer Aided Medical Procedures, Technische Universität München, 85748 Munich, Germany (email: navab@cs.tum.edu).*

### 1.1 Background

Humans have a complex and sophisticated auditory system. From an evolutionary point of view, auditory perception skills and having the ability to detect and decode information in environmental sounds has had a significant survival advantage. Exploiting human auditory perception skills and employing non-speech audio to convey information, i.e., sonification is the idea that has been around for a long time. Sonification is often defined as the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation [5]. From simple beeps and bleeps generated by home appliances to navigation using audio feedback, the concept of sonification of quantitative parameters is everywhere. As visualization of data streams is an essential tool for both analyzing and exchanging scientific data, sonification is slowly finding its way in science. Since mid 1990s, there have been several studies on the subject of sonification in data representation and guidance [4]. Since many applications of AR (e.g., medical AR) already require high visual concentration on focus demanding tasks, aural augmentation is a suitable candidate for avoiding obstruction of the visual field of focus with added information. Experiments in the field of psychophysics strongly suggest that warning signals in the form of auditory stimuli capture spatial attention better than visual signals. Using auditory cues to indicate the occurrence of an event or conveying spatial or mechanical information has proven to be effective and reliable [11].

Mapping parameters extracted form a data stream to sounds has similarities and dissimilarities to visualization of the same parameters. On one hand, pitch, volume, timbre, and other qualities of synthesized sounds could be used in the same way that color, intensity, opacity, and other qualities of synthesized images are used to convey multichannel information. On the other hand, many aspects including the rate of sensation decay, the learning time required for users, and the conscious effort required for perceiving the presented information is not directly comparable in the two mechanisms. Hence, thorough usability testings are required for individual methods incorporated for aural augmentation. One of the aspects involved in sonification of data streams to investigate is the phenomenon known as alarm fatigue. Usage of non-speech audio to express information could lead to an overload of sensory stimuli which in turn could lead to misinterpretation or loss of significant information. This phenomenon is critical in the medical field where a multitude of devices generate sounds that carry different information at various levels of importance often simultaneously. Moving away from devices that sonify streams of data with tones that are unintuitive, incompatible with human perception, and negatively perceived is a necessity. Employing sound synthesis methods that adhere to the basics

of aesthetics has the capability of improving work environment of users of such systems (e.g., physicians) as well as human subjects who are exposed to generated sounds but do not necessarily need to understand the underlying information (e.g., patients). Another aspect involved in designing a feasible auditory display system is how intuitively understandable the presented audio feedback is. Similar to the way a siren sound indicates urgency or the sound of tapping on a hollow object expresses qualitative information regarding its construction, a proper auditory display design can convey information in such a way that requires minimum cognitive effort to interpret. Intuitive embodiment of information in auditory feedback can potentially accelerate the learning process to the level that subconscious association of auditory cues to events is achieved.

Auditory display systems employ a multitude of methods for mapping data to sound. In [2] Dubus and Bresin have conducted a review of mapping strategies in 179 scientific publications involving sonification of physical quantities. They conclude that pitch is by far the most used auditory dimension in sonification mappings. More than half of the most popular associations between physical and auditory dimensions involve Pitch. The design process for a sonification system typically starts by mapping the most important data dimension to the frequency of a pure tone. Moreover, the authors show that natural perceptual associations between sounds and physical quantities are limited to straightforward concepts such as associating distance to loudness or location to spatialization of sound. In [6] the authors illustrate that sonification strategies which give the users the opportunity to hear a sound that can be compared with the guidance tone, generally achieve higher accuracy but at the cost of longer identification time. This shows that most common sonification schemes that use tones without any reference sound, lack the proper human cognition requirements needed in an interactive system.

Physical modeling of sound is the set of mathematical methods by which the generation of sound in resonating matter is simulated. Sounds generated using physical modeling closely resemble sounds made by natural phenomena that are known to human subjects. Utilizing sounds made by physical modeling in an auditory display system has the potential of providing instinctive reference to human subjects beside minimizing the mental effort required for correlating auditory stimuli to events. In this work, we present methods for designing a sonic interaction system that uses physical modeling sound synthesis to represent quantitative information in an intuitive way. We leverage the resemblance of generated sounds using physical modeling to natural phenomena to address both the problem of alarm fatigue as well as occupying the least amount of cognitive load for users.

### 1.2 Previous Work

Mathematical methods behind physical modeling sound synthesis simulate vibrations in physical objects and sample the amplitude and frequencies of these vibrations at certain time intervals. This could be achieved via assembling linear or non-linear mathematical functions describing displacement in a resonating matter when force is applied to it. A leading example of such mathematical simulations is the Modalys [3] software environment. Modalys' synthesis model is based on four types of elements. Objects, accesses, connections, and controllers which can be assembled in order to build virtual instruments and synthesize sounds. Modalys objects describe vibrating structures defined by their geometrical characteristics (e.g. strings, plates, or membranes). Modalys connections describe the mode of interaction between objects (e.g. strike, pluck, or bow). Modalys accesses specify the locations on objects at which they interact with each other. And Modalys controllers are used to specify the trajectories of all time-varying synthesis parameters (e.g. coordinates of accesses, connection parameters, forces, and controller inputs) [3]. Internally, analytical and finite element methods are used for sampling vibrations of the constructed virtual instruments. A similar approach to sound synthesis is using numerical methods to solve non-linear models representing a resonating object. In [8] Roze and Bensoam present an example of sound synthesis with one interaction based on a nonlinear physical model.

In a sonic interaction design, the mentioned sound synthesis methods are utilized to map quantitative parameters to resonance. A fine example of such approach in action is the work by Shelly *et al.* [9] in which sonification using physical modeling is used to communicate information about the curve shape and curvature of a virtual three-dimensional object. Results achieved by their experiments suggest that humans, previously unexposed to the system, are capable of establishing a clear connection between the sound and the underlying data. This is consistent with the assumption that usage of physical modeling leads to less cognitive load necessary for understanding the underlying data as in their experiments the sound is representing a nontrivial mathematical term expressing curvature of a three-dimensional object. More recently, physical modeling of sound has been successfully used to interactively sonify the foot-floor contacts resulting from jumping on an elastic trampoline in real-time [12]. Their results indicate that the proposed auditory feedback can modulate the perception of foot-haptic sensations of the surface utilized when jumping. These results and similar sonification approaches signify the feasibility of physical modeling sound synthesis in various applications. In continuation of the aforementioned works, here we propose a general aural augmented reality system to convey precise gestural information via sound.

## 2 METHODS

Under certain circumstances the sensation of reaching an object and the moment of contacting its surface could be intangible. Examples of such conditions are touching an object with high elasticity using a miniature instrument or reaching an object under a microscope where depth perception is not immediate. Similarly, the amount of pressure applied to an object could be imperceivable. When the act of manually applying and maintaining precise force to an object is required for high precision tasks, tactile overstimulation negatively affects the outcome of the job at hand. Moving an object around while preserving its physical pose needs constant conscious regard. Current sensory technology allows accurate measurement of such physical quantities. However, presentation of the data stream generated by sensors using visual augmented reality brings the burden of requiring repeated change of visual focus. Here, three sonification mechanisms are introduced that employ sound synthesized by physical modeling to embody touch, applied pressure and angle of approach in high precision tasks regardless of the sensing method.

To represent the moment of gently touching a surface, the sound of resonating a bounded elastic membrane when struck by hand is used. The recorded acceleration on impact is mapped directly into the sound volume and is inversely proportional to the elasticity of the virtual membrane. To understand the reasoning behind the choice of sound in this scenario one could imagine the similarity between the described synthesized sound and the actual sound generated when an egg yolk is touched (and amplified by several orders of magnitude). Mapping the acceleration on impact to sound volume comes natural. Inversely proportioning the acceleration to the elasticity of the virtual membrane is meant to eliminate the possibility of missing the low volume sound signal when the real object is approached slowly and under utmost care.

Representation of pressure applied to a real object is done by synthesizing the sound of tapping on a wooden square plate in intervals. To convey information regarding the amount of pressure, two thresholds are considered. A soft threshold that can indicate either the desired pressure or the high but tolerable pressure and a hard threshold marking the maximum tolerated pressure. The amount of pressure is embodied using two mechanisms. Firstly, as the pressure rises the intervals between the taps become shorter. When the pressure is at the hard threshold, the intervals between the taps tend to zero turning the taps into a solid tone. Secondly, as the pressure rises, the material which the virtual plate is composed of gradually turns from wood to metal. The change of plate material acts as an implicit reference. As the user applies more pressure, they can sense the effect in the timbre and quality of the feedback auditory signal. Once the tapping sounds tend to turn metallic, they can stop and maintain the current applied pressure. The difference between the soft threshold and the current applied pressure is inversely mapped to the sound volume as a guide to target. The sound

volume is determined using the following function:

$$v(c) = \frac{c^2 - c}{t^2 - t}, \qquad v, c, t \in [0, 1] \tag{1}$$

in which $v$ is the sound volume level of the playback device, $c$ is the current pressure, and $t$ is the soft threshold. All the parameters affecting the material of the virtual plate could be described using a vector of values generated by a function with its input being a value calculated by the following piecewise function:

$$\begin{cases} p(c) = \frac{m-w}{T}c + w & c \in [0, T] \\ p(c) = m & c \in (T, 1] \end{cases}, \tag{2}$$

in which $T$ is the hard threshold, $c$ is the current pressure, and $P(c)$ holds a value between $w$ and $m$ for wood and metal respectively which is then mapped to a vector of values responsible for the material the virtual plate is constructed from. The idea behind this scheme of sonification is to provide an implicit cognitive reference in the generated audio. This method is in contrast to conventional sonification schemes that use explicit auditory icons or earcons to represent thresholds.

Embodiment of angle or pose in sound is done with synthesizing resonance generated by plucking and dampening a bounded string in intervals. The goal of the scenario in this case is to find and maintain a desired angle or pose. When the recorded angle is less than the desired one, the pitch of the plucked string is low and when the recorded angle is higher than the one desired, the pitch is high. There is no continuous mapping of angle to pitch but rather two fixed pitches are chosen to lower the mental effort for the user. The recorded angle is continuously mapped to the tapping intervals. Contrary to the common sonification schemes, in the proposed method, tapping intervals are inversely proportioned to the distance to the desired angle. A recorded angle too far away from the desired angle turns the tapping to a solid tone and perfectly matching with the desired angle is expressed using total silence. This choice came after initial experiments with the designed system. Feedback both from untrained users and experts familiar with sonification systems indicated that continuous feedback regarding the correctness of the maintained pose is rather destructive. A typical task requiring a user to employ an instrument in a specific pose or angle has two phases. First, searching for the proper pose or angle by sweeping through the range of possible options. Second, maintaining the correct pose or angle while manipulating an object by the instrument. The second phase is in essence both longer and more focus demanding compared to the first phase. Continuously informing a user about correctness of the maintained pose during a long and focus demanding period is both unnecessary and distracting. Deviation from the correct pose would cause the user to hear tapping sounds again requiring him to adjust the pose accordingly which achieves the guidance goal without imposing additional burden.

Multiple mathematical methods are available for synthesizing the sound of stroking a membrane ranging from solving differential equations of acoustically resonating thin plates (e.g. Timoshenko formulation) to deterministic waveguide finite element method. For synthesizing the sound of plucking a string, a famous mathematical method is the following Karplus-Strong recursive formulation:

$$y_n = x_n + \frac{y_{n-N} + y_{n-N-1}}{2}, \qquad N = \frac{f_s}{f_0}, \tag{3}$$

in which $y_n$ is the $n^{\text{th}}$ frequency sample, $f_s$ is the sampling rate (44100 Hz is a typical and sufficient choice since it can adequately cover human hearing range), $f_0$ is the fundamental frequency of the synthesized plucked string and could be set to the two selected pitch frequencies in the case of presenting pose via sound, and $x_n$ is an excitation noise signal of length $N$. In this work, the Modalys software framework is used for sound synthesis with fine tuning parameters for aesthetic purposes.
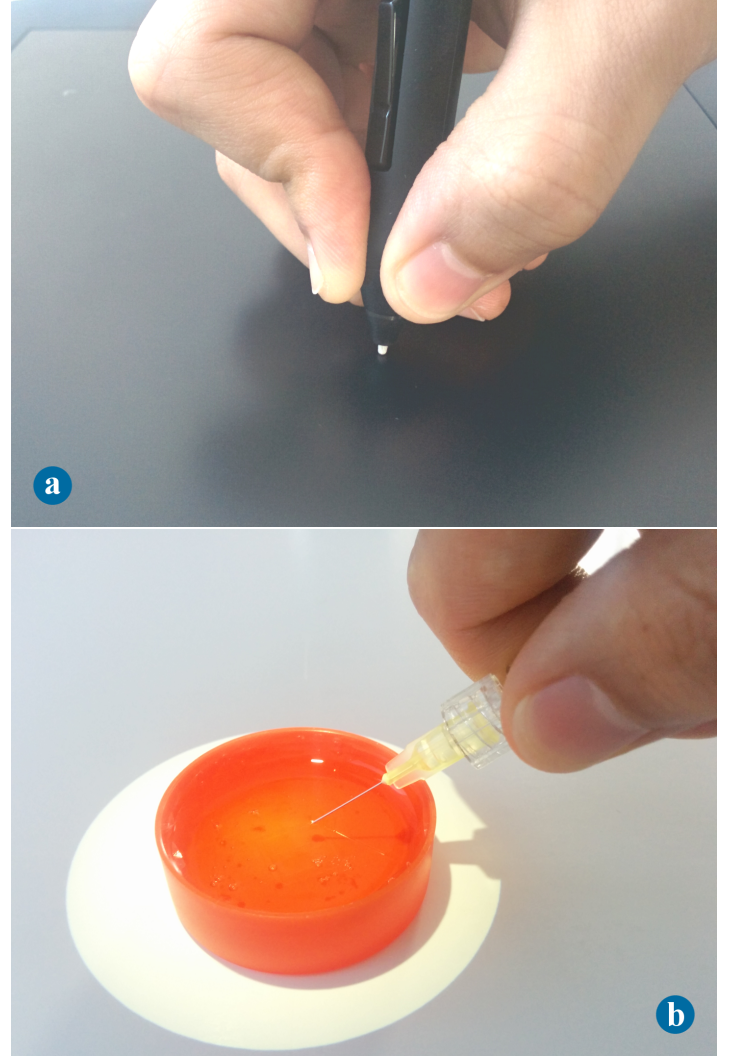
Fig. 1. Setups of experiments involving high precision manual Tasks. (a) Moving a stylus on a digitizer with the aim of tracking an object while applying a predetermined pressure on the board. (b) Moving an injection needle under a microscope with the aim of gently touching a highly viscous liquid with a predefined angle of approach.

## 3 EXPERIMENTS

To verify the effectiveness of the proposed method, two experiments are designed and conducted. These experiments are designed with the aim of comparing environments that are augmented only visually, solely aurally or with both techniques simultaneously. The basic principle kept consistent in the design of both experiments is concentrating on tasks that require extreme precision. A pair of over-the-ear stereo headphones is employed for sound playback. All the software required for sound synthesis, visual augmented reality and recording experiment measures are executed on a single workstation computer. To compare various combinations of augmented reality methods, error measures are recorded in each experiment for every subject separately.

### 3.1 Task 1

This assignment is designed to simulate a high precision manual task using a rod-like instrument. The aim in this task is to use a digital stylus on a digitizer board (Wacom Intuos Pro, Wacom Co. Ltd., Japan) to track a moving object while keeping the pressure applied by the stylus at a target value that changes over time (Fig.1a). The working area on the digitizer ($224mm \times 148mm$) is mapped to a monitor in front of the subject. A white filled circle as the target object to track is rendered on a black background on the screen. Every 9 seconds, a two-dimensional speed vector with a random size of maximum $10mm$ per second and a
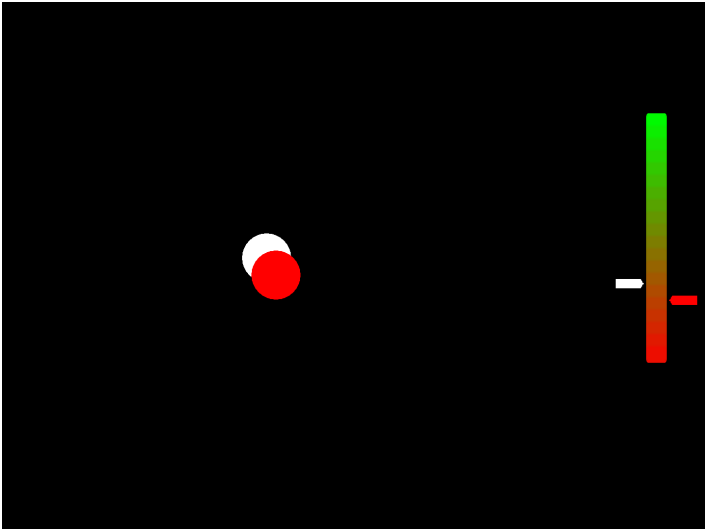
Fig. 2. The view on the screen for task 1. The white circle moves with random direction and speed. The red circle indicates where the users stylus location is mapped to. The gauge shows the current versus the target pressure with red and white indicators respectively. The gauge is not visible when only sonification is employed to give feedback regarding the applied and target pressure values.
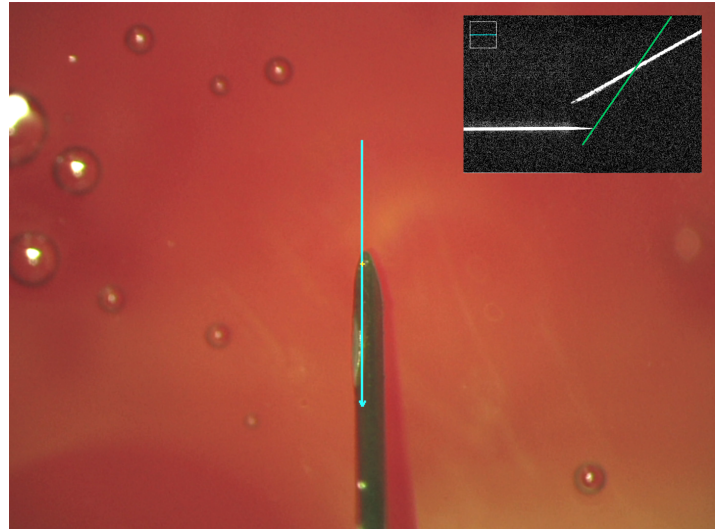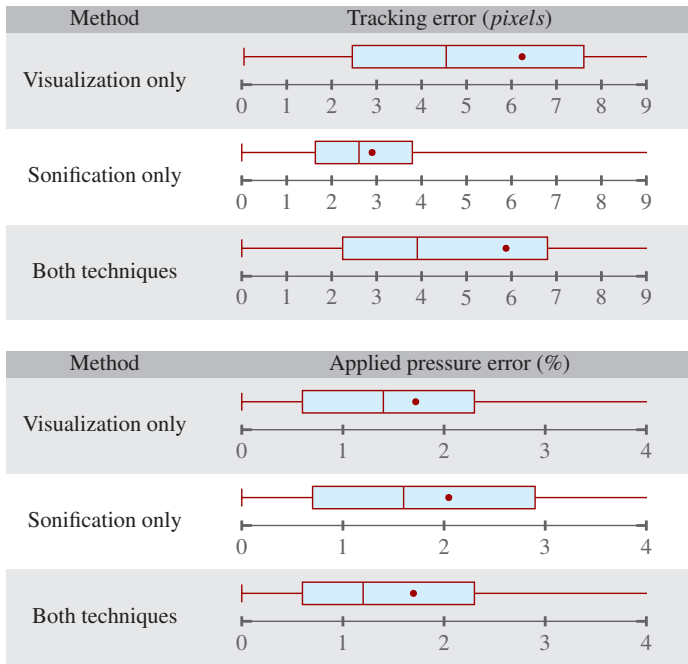


Fig. 3. The augmented microscopic view on the HUD for task 2. The cyan line indicates the OCT scan location. The top right overlay image is the OCT acquisition showing the cross-sectional view of the liquid surface and the top segment of the needle. The green line on the OCT image shows the target angle of the needle that the participant is expected to maintain when the liquid surface is touched.

random direction is generated and assigned to the circle. At the same time intervals, a random target pressure value ranging from 0 to 1 in 0.001 increments is selected. The location on the screen where the stylus tip is mapped to is indicated with a red filled circle (Fig.2). The software designed for this task records the Euclidean distance between the two circles centers and the absolute difference between the selected target pressure and the actual pressure applied to the digitizer by the test subject in 3 intermediate seconds of each interval. To inform the participants of the target and the applied pressure, three mechanisms are tested. In one run, the pressure is rendered on one side of the screen using a linear color marked gauge. The target pressure and the current pressure applied to the digitizer are indicated on the gauge by a white and a red line respectively. In the second run of the task, the proposed sonification method is utilized with no visual cue. Reaching the target pressure is emphasized with a louder hit on the virtual plate. A Schmitt trigger mechanism avoids the system from reconfirmation of reaching the target pressure when the applied pressure oscillates around it. The third run of the test engages both visual and aural techniques as assistance. Participants perform several mock runs to learn the procedure. Then, each participant performs 6 trials with each of the runs in a randomly determined order.

### 3.2 Task 2

This assignment is designed to test the capabilities of the proposed system under more realistic conditions. For the purpose of this task, an ophthalmic surgical microscope (OPMI Lumera 700, Carl Zeiss Meditec AG, Germany) is used. Test subjects are asked to move a needle with a diameter of 0.3$mm$ towards a transparent and highly viscous liquid and gently touch it while maintaining a certain angle (Fig.1b). The procedure is performed under the surgical microscope while looking through the ocular lenses. For each trial a random target angle between 10 and 80 degrees is selected. To assist the participants, a cross-sectional Optical Coherence Tomography (OCT) image of the needle and the surface is continuously acquired. Since OCT uses infrared to capture images and is acquired from top, no signal can be recorded beyond the needle and only the top segment of the instrument is effectively visible. The actual angle of the needle, the location of the tip of the needle, and whether the needle tip is touching the surface of the liquid or not are extracted from the OCT images by computer vision techniques. To pass on this information to the participants, three AR mechanisms are tested. In one run, the OCT image and the target

angle line are rendered onto a Heads-Up Display (HUD) integrated in the ocular lens of the microscope (Fig.3). Hence, the user is able to view the extra information on top of the procedure scene. When the system detects a touch, a red bar appears at the top left corner of the scene. In the second run, the HUD is only used to show the raw OCT image and only the proposed sonification method to convey information regarding angle and touch is employed. In the third run, both AR methods are employed. Test subjects are asked to keep the needle under a line visible in the HUD that indicates the location of the OCT scan. Participants press a button on a foot pedal when they can confirm the tip of the needle touches the surface. Each participant performs the task several times in all runs. The software designed for this task records the difference between the selected target angle and the actual maintained angle at the moment the employed computer vision algorithm detects an actual touch event. The system also records every time a touch occurs and whether the test subject has correctly identified it using the foot pedal within a period of 1.5 seconds before and after the moment of touch. Participants perform 2 trial runs for each method of AR to learn the procedure. The order of the runs in recorded sessions is determined randomly.

## 4 RESULTS

A total of 18 participants performed Task 1. The screen resolution for the experiment was set to $1720 \times 1140$ pixels resulting to a pixel size of roughly 13 $\mu m$. The error in tracking the visualized circle on the screen using the stylus by all participants was on average 6.235 pixels (with standard deviation of 5.964 pixels) corresponding to 810 $\mu m$ when only visualization of the applied pressure was used. When Sonification was the only method for conveying information regarding pressure, the average error in tracking dropped to 2.899 pixels (with standard deviation of 1.763 pixels) corresponding to 370 $\mu m$ for all users. Using both visualization and sonification techniques resulted in an average tracking error of 5.878 pixels (with standard deviation of 6.195 pixels) corresponding to 764 $\mu m$. The digitizer used in the first experiment reports a certain range of pressure applied on it with a resolution of 1000 units. The measured errors in pressure are reported in percentage of the minimum to maximum range of pressure the digitizer can measure. Employing only visualization on average resulted to an error of 1.72% in the applied pressure while employing only sonification raised the average error to 2.04%. Using both techniques resulted in an average error of 1.69% similar to the case when visualization was the only

Table 1. Measurements for Task 1

| Method | Tracking error (*pixels*) |
|---|---|
| Visualization only | (box plot, scale 0–9) |
| Sonification only | (box plot, scale 0–9) |
| Both techniques | (box plot, scale 0–9) |

| Method | Applied pressure error (%) |
|---|---|
| Visualization only | (box plot, scale 0–4) |
| Sonification only | (box plot, scale 0–4) |
| Both techniques | (box plot, scale 0–4) |

Table 2. Measurements for Task 2

| Method | Angle error (*degrees*) |
|---|---|
| Visualization only | (box plot, scale 0–8) |
| Sonification only | (box plot, scale 0–8) |
| Both techniques | (box plot, scale 0–8) |

| Method | Touch incidences (%) |
|---|---|
| Visualization only | True positives: 66.7% <br> False negatives: 27.8% <br> False positives: 5.5% |
| Sonification only | True positives: 77.8% <br> False negatives: 22.2% <br> False positives: 0.0% |
| Both techniques | True positives: 77.8% <br> False negatives: 22.2% <br> False positives: 0.0% |

utilized method. The standard deviation in pressure error was 1.6% for the case of sonification and when both techniques were used, and 1.76% when sonification was the only used method. Details of measurements for task 1 are presented in Table 1.

Task 2 was tested on 6 participants for a total of 54 runs. The error in the maintained angle when visual AR using HUD is employed to show the target angle was on average 5.06 degrees with a standard deviation of 4.21 degrees. This figure dropped to 4.31 degrees with a standard deviation of 6.15 degrees when only aural AR was used. Using both visual and aural AR resulted in an average angle error of 4.64 degrees with a standard deviation of 5.13. Regarding touch events, when a touch is registered by both the participant and the corresponding computer vision algorithm, a true positive incidence is recorded. When the user misses a touch, a false negative incidence is recorded and when the test subject registers a touch but the system does not recognize a touch event, a false positive incidence is recorded. Using visual AR to signify touching the surface, 66.7% true positive was observed. This figure improved to 77.8% when aural AR was employed. Introducing both AR techniques kept the figure at the rate of 77.8%. Details of measurements for task 2 are presented in Table 2.

In both tables 1 and 2, box plots are used to present the recorded data in experiments. The start and the end of each box indicate the first and the third quartile while the inner band and the dot represent the median and the average of the recorded data respectively.

## 5 DISCUSSION

Results from the first experiment indicate that when a system contains multiple sources of visual information, introducing a sonification method with minimum cognitive load to convey part of the information using channels other that vision is helpful. Converting part of the visual information to sounds results in more focus on the tasks that require the remaining visual information to proceed. The clear connection could be observed in the improvement of tracking error from 810 $\mu m$ to less than half (370 $\mu m$). Substituting visualization with sonification lowered the accuracy of maintained pressure only by 0.32% which suggests that the particular employed method of sonification is effective in replacing visualization and conveying information. Results from the case where both sonification and visualization are used show that users have the tendency to focus more on visualization when available though measures still show slight improvement in errors. One could argue that moving the visual cue along with the target
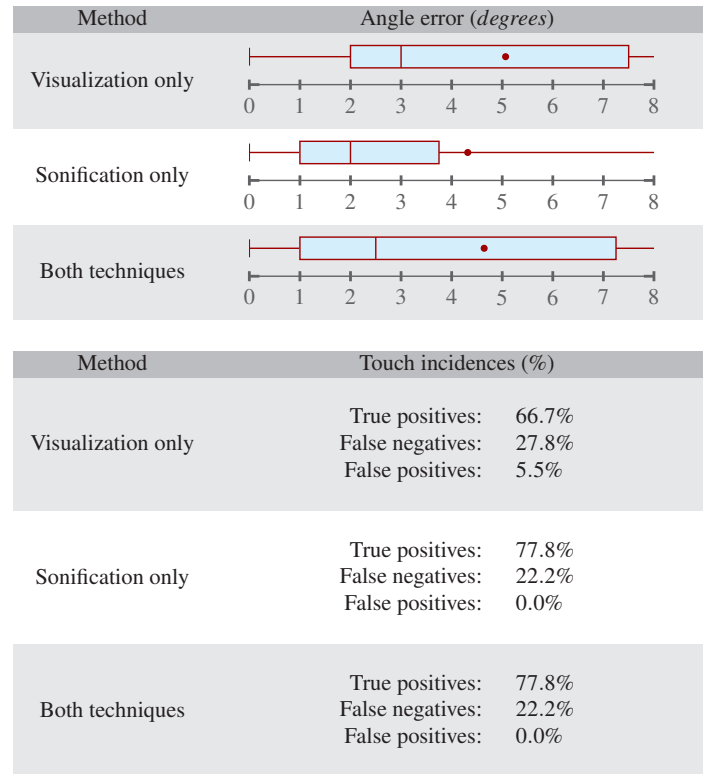
representing the tip of the instrument would improve the user tracking error when only visualization is used. However, in scenarios requiring high focus and accuracy, which the proposed sonification scheme is targeted for, obstructing the focus of the visual field with added information is commonly considered bad practice. As an example, in the case of surgical AR, a major obstacle in the way of smooth integration of medical AR devices into the medical workflow is the fact that they restrict the field of view of the physicians by introducing augmented information [10]. Moreover, a common concern associated with visual AR systems is the accuracy of the engaged tracking algorithms responsible for accurately overlaying augmented information. Rendering a visual cue such as a gauge next to the instrument and moving it as the instrument moves in a realistic scenario is prone to errors in automatic tracking. Omnidirectionality of sound perception means that misdetection of objects and loss of tracking do not influence an aural augmented reality system in the same way they affect visual augmented reality systems.

Task 2 is considered to be extremely difficult to perform by human subjects who do not have experience with manipulation of objects under a microscope. Skills such as hand-eye coordination, tremor and breathing control, and accurate depth perception play significant roles in the general outcome of a microscopic task. Test subjects for task 2 were not trained for microscopic manipulation experiments. This explains the large interquartile ranges in the recorded data for this experiment. One of the goals of this study is to experiment with representing natural phenomena such as events or number readings from sensors using sounds that are generated by modeling natural resonating interactions with objects rather than beeps and tones that require training to be understood. At first look, the sound of plucking a string to represent the angle of an instrument seems to contradict this goal. However, one should consider the emphasis of the employed sonification mechanism at hand. Two distinct pitches in combination with intervals and silence are representing angle. Detecting pitch differences for users with no musical training is easier when contextual cues including tones and timbre qualities of a musical instrument such as a damped bounded string is added to the audio signal compared to when simple solid tones or sine waves are used [13]. The increase in the accuracy of

the maintained angles seen in task 2 when audio is the sole source of information confirms the effectiveness of the proposed natural sound synthesis scheme. Using both sonification and visualization also shows slight improvement in the recorded errors, though the effect is not as strongly observed as in the case of only sonification. Regarding the information on touch incidences, while the data does not suggest strong improvement when sound is alone used to inform the participants, the fact that measures are similar in the case of using only aural augmented reality and the case of using both visual and aural AR, concludes a tested scenario where sonification and visualization could be combined. Using the combination, has the advantage of benefiting the user with the additional information presented in multiple ways and yet not changing the workflow of the procedure by removing visual augmented information. Another point to consider is the fact that counting true positive incidences relies on the accuracy of the automatic touch detection algorithm in use. The assumption that all positive touch events are correctly detected by the employed computer vision techniques and no false touch event is recorded is unrealistic. Overall, improvements in the performance of the test subjects leading to lower errors recorded in both experiments show the feasibility of the utilized methods to present sensory data.

## 6 CONCLUSION

The main contribution of this work is presenting a new model to follow for applications that desire to use auditory channels for information conveyance regardless of sensory means of input data. The aim of the presented work is to employ physical sound synthesis to generate feedback sounds derived from visual information to improve the accuracy achieved in performing high precision tasks. In designing the proposed sonification scheme, adherence to the basics of aesthetics, use of implicit references and employing minimalistic sound patterns to avoid alarm fatigue were among top aims. We specifically demonstrated the feasibility of employing aural augmented reality as a feedback tool in high precision tasks where visual augmented reality could bring distraction from the field of focus. The presented results from initial experiments indicate the effectiveness of using such sonification systems for extremely focus demanding tasks.

Future work will be in the direction of designing experiments that evaluate the temporal aspects of aural augmented reality systems specially for tasks requiring rapid reactions. Comparing alternative parameter mappings for each particular type of information could result in the design of systems that are more robust to noise, missing signal, and minimum user training. Finally, evaluating the proposed methods in real scenarios e.g. augmented medical workflows is of great benefit for future designs.

## REFERENCES

[1] J. C. Brill, R. D. Gilson, M. Mouloua, P. A. Hancock, and P. I. Terrence. Increasing situation awareness of dismounted soldiers via directional cueing. *Human Performance, Situation Awareness, and Automation: Current Research and Trends*, 1:130–132, 2005.

[2] G. Dubus and R. Bresin. A systematic review of mapping strategies for the sonification of physical quantities. *PloS one*, 8(12), 2013.

[3] G. Eckel, F. Iovino, and R. Caussé. Sound synthesis by physical modelling with modalys. In *Proc. International Symposium on Musical Acoustics*, pages 479–482, 1995.

[4] T. Hermann, A. Hunt, and J. G. Neuhoff. *The sonification handbook*. Logos Verlag Berlin, 2011.

[5] G. Kramer. *An introduction to auditory display, in Auditory Display: Sonification, Audification and Auditory Interfaces*. Reading, MA: Addison-Wesley, 1994.

[6] G. Parseihian, C. Gondre, M. Aramaki, S. Ystad, and R. Kronland-Martinet. Comparison and evaluation of sonification strategies for guidance tasks. *IEEE Transactions on Multimedia*, 18(4):674–686, 2016.

[7] H. Roodaki, C. A. di San Filippo, D. Zapp, N. Navab, and A. Eslami. A surgical guidance system for big-bubble deep anterior lamellar keratoplasty. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 378–385. Springer, 2016.

[8] D. Roze and J. Bensoam. Nonlinear physical models of vibration and sound synthesis. In *Unfold Mechanics for Sounds and Music*, 2014.

[9] S. Shelley, M. Alonso, J. Hollowood, M. Pettitt, S. Sharples, D. Hermes, and A. Kohlrausch. Interactive sonification of curve shape and curvature data. In *International Conference on Haptic and Audio Interaction Design*, pages 51–60. Springer, 2009.

[10] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays: A literature review of augmented reality. *Journal of Display Technology*, 4(4):451–467, 2008.

[11] C. Spence and J. Driver. Audiovisual links in exogenous covert spatial orienting. *Attention, Perception, & Psychophysics*, 59(1):1–22, 1997.

[12] L. Turchet, R. Pugliese, and T. Takala. Physically based sound synthesis and control of jumping sounds on an elastic trampoline. In *Interactive Sonification Workshop (ISon 2013)*, pages 87–94, 2013.

[13] C. M. Warrier and R. J. Zatorre. Influence of tonal context and timbral variation on perception of pitch. *Attention, Perception, & Psychophysics*, 64(2):198–207, 2002.

# A Surgical Guidance System for Big-Bubble Deep Anterior Lamellar Keratoplasty

Hessam Roodaki[1], Chiara Amat di San Filippo[1], Daniel Zapp[3], Nassir Navab[1,2], Abouzar Eslami[4]

[1] Computer Aided Medical Procedures, Technische Universität München, Germany
[2] Computer Aided Medical Procedures, Johns Hopkins University, USA
[3] Augenklinik rechts der Isar, Technische Universität München, Germany
[4] Carl Zeiss Meditec AG, Germany

**Summary.** This publication proposes a virtual reality technique that provides guidance for an anterior segment ophthalmic procedure. Deep anterior lamellar keratoplasty using the big-bubble technique is a corneal transplantation intervention that has a critical phase in which surgeons insert a needle to a depth of around 90% of the cornea to inject air. Not reaching the target depth would lead into the failure of the procedure and needle getting too close to the endothelial layer of the cornea could also lead into perforation and failure. Hence, OCT plays a crucial role in BB-DALK. However, total shadowing caused by metallic instruments, artifacts and noise in OCT scans, limited field of view, and difficulty in control of OCT scan parameters are among shortcomings that make OCT inferior. The proposed method employs 3D OCT scans over time that are visualized using adaptive color transfer functions to minimize artifacts and maximize the contrast between the instrument and tissue. By doing so, the artifacts in the 3D+t view of the surgical scene are suppressed and the field of view of OCT is enhanced from a few square millimeters to several cubic millimeters. Corneal layers are tracked and visualized using virtual elements. This tracking allows the system to automatically adjust the view of the surgeon to provide maximal depth information. The target depth of the cornea which could be selected intraoperatively is visualized using a virtual surface to guide the surgeons throughout the procedure. The proposed virtual reality technique provides a complete representation of the surgical field without relying on the microscopic view. Hence, surgeons are able to perform the entire air injection phase of the intervention by looking at the virtual representation of the surgical scene. The accuracy of corneal layer tracking which plays a critical role in the proposed technique is evaluated using a micromanipulator.

**Contribution.** The contribution of the author of this thesis to this publication consists of the main idea, the implementation, experiment design and evaluation. Writing of the manuscript was done and coordinated by the author of this thesis. Discussions on the findings and the revision of the publication was done jointly with the co-authors.

# A Surgical Guidance System for Big-Bubble Deep Anterior Lamellar Keratoplasty

Hessam Roodaki[1(✉)], Chiara Amat di San Filippo[1], Daniel Zapp[3],
Nassir Navab[1,2], and Abouzar Eslami[4]

[1] Computer Aided Medical Procedures,
Technische Universität München, Munich, Germany
`he.roodaki@tum.de`
[2] Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, USA
[3] Augenklinik rechts der Isar, Technische Universität München, Munich, Germany
[4] Carl Zeiss Meditec AG, Munich, Germany

**Abstract.** Deep Anterior Lamellar Keratoplasty using Big-Bubble technique (BB-DALK) is a delicate and complex surgical procedure with a multitude of benefits over Penetrating Keratoplasty (PKP). Yet the steep learning curve and challenges associated with BB-DALK prevents it from becoming the standard procedure for keratoplasty. Optical Coherence Tomography (OCT) aids surgeons to carry out BB-DALK in a shorter time with a higher success rate but also brings complications of its own such as image occlusion by the instrument, the constant need to reposition and added distraction. This work presents a novel real-time guidance system for BB-DALK which is practically a complete tool for smooth execution of the procedure. The guidance system comprises of modified 3D+t OCT acquisitions, advanced visualization, tracking of corneal layers and providing depth information using Augmented Reality. The system is tested by an ophthalmic surgeon performing BB-DALK on several *ex vivo* pig eyes. Results from multiple evaluations show a maximum tracking error of 8.8 micrometers.

## 1 Introduction

Ophthalmic anterior segment surgery is among the most technically challenging manual procedures. Penetrating Keratoplasty (PKP) is a well-established transplant procedure for the treatment of multiple diseases of the cornea. In PKP, the full thickness of the diseased cornea is removed and replaced with a donor cornea that is positioned into place and sutured with stitches. Deep Anterior Lamellar Keratoplasty (DALK) is proposed as an alternative method for corneal disor-ders not affecting the endothelium. The main difference of DALK compared to PKP is the preservation of the patient's own endothelium. This advantage reduces the risk of immunologic reactions and graft failure while showing simi-lar overall visual outcomes. However, DALK is generally more complicated and time-consuming with a steep learning curve particularly when the host stroma is manually removed layer by layer [4]. In addition, high rate of intraoperative

perforation keeps DALK from becoming surgeons' method of choice [7]. To overcome the long surgical time and high perforation rate of DALK, in [1] Anwar *et al.* have proposed the big-bubble DALK technique (BB-DALK). The fundamental step of the big-bubble technique is the insertion of a needle into the deep stroma where air is injected with the goal of separating the posterior stroma and the Descemet's Membrane (DM). The needle is intended to penetrate to a depth of more than 60 % of the cornea, where the injection of air in most cases forms a bubble. However, in fear of perforating the DM, surgeons often stop the insertion before the target depth, where air injection results only in diffuse emphysema of the anterior stroma [7]. When bubble formation is not achieved, effort on exposing a deep layer nearest possible to the DM carries the risk of accidental perforation which brings further complications to the surgical procedure.

Optical Coherence Tomography (OCT) has been shown to increase the success rate of the procedure by determining the depth of the cannula before attempting the air injection [2]. Furthermore, recent integration of Spectral Domain OCT (SD-OCT) into surgical microscopes gives the possibility of continuous monitoring of the needle insertion. However, current OCT acquisition configurations and available tools to visualize the acquired scans are insufficient for the purpose. Metallic instruments interfere with the OCT signal leading to obstruction of deep structures. The accurate depth of the needle can only be perceived by removing the needle and imaging the created tunnel since the image captured when the needle is in position only shows the reflection of the top segment of the metallic instrument [2]. Also, limited field of view makes it hard to keep the OCT position over the needle when pressure is applied for insertion.

Here we propose a complete system as a guidance tool for BB-DALK. The system consists of modified 3D+t OCT acquisition using a microscope-mounted scanner, sophisticated visualization, tracking of the epithelium (top) and endothelium (bottom) layers and providing depth information using Augmented Reality (AR). The method addresses all aspects of the indicated complex procedure, hence is a practical solution to improve surgeons' and patients' experience.

## 2 Method

As depicted in Fig. 1, the system is based on an OPMI LUMERA 700 microscope equipped with a modified integrated RESCAN 700 OCT device (Carl Zeiss Meditec, Germany). A desktop computer with a quad-core Intel Core i7 CPU, a single NVIDIA GeForce GTX TITAN X GPU and two display screens are connected to the OCT device. Interaction with the guidance system is done by the surgeon's assistant via a 3D mouse (3Dconnexion, Germany). The surgeon performs the procedure under the microscope while looking at the screens for both microscopic and OCT feedback. The experiments are performed on *ex vivo* pig eyes as shown in Fig. 3a using 27 and 30 gauge needles. For evaluations, a micromanipulator and a plastic anterior segment phantom eye are used.
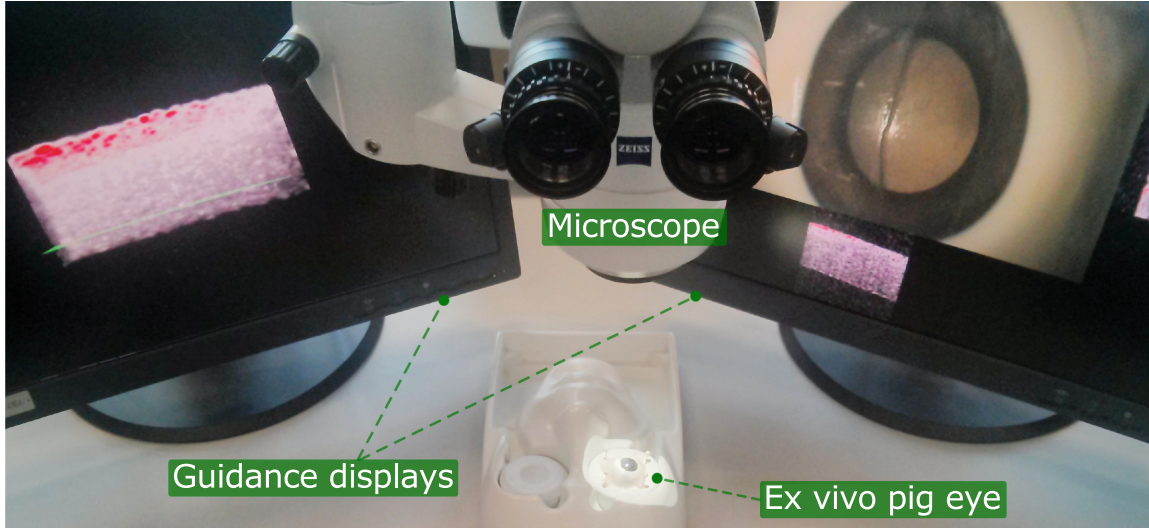
**Fig. 1.** Experimental setup of the guidance system.

## 2.1 OCT Acquisition

The original configuration of the intraoperative OCT device is set to acquire B-scans consisting of either 512 or 1024 A-scans. It can be set to acquire a single B-scan, 2 orthogonal B-scans or 5 parallel B-scans. For the proposed guidance system, the OCT device is set to provide 30 B-scans each with 90 A-scan samples by reprogramming the movement of its internal mirror galvanometers. B-scans are captured in a reciprocating manner for shorter scanning time. The scan region covered by the system is 2 mm by 6 mm. The depth of each A-scan is 1024 pixels corresponding to 2 mm in tissue. The concept is illustrated in Fig. 2a.

The cuboid of $30 \times 90 \times 1024$ voxels is scanned at the rate of 10 volumes per second. Since the cuboid is a 3D grid of samples from a continuous scene, it is interpolated using tricubic interpolants to the target resolution of $180 \times 540 \times 180$



(a)                         (b)
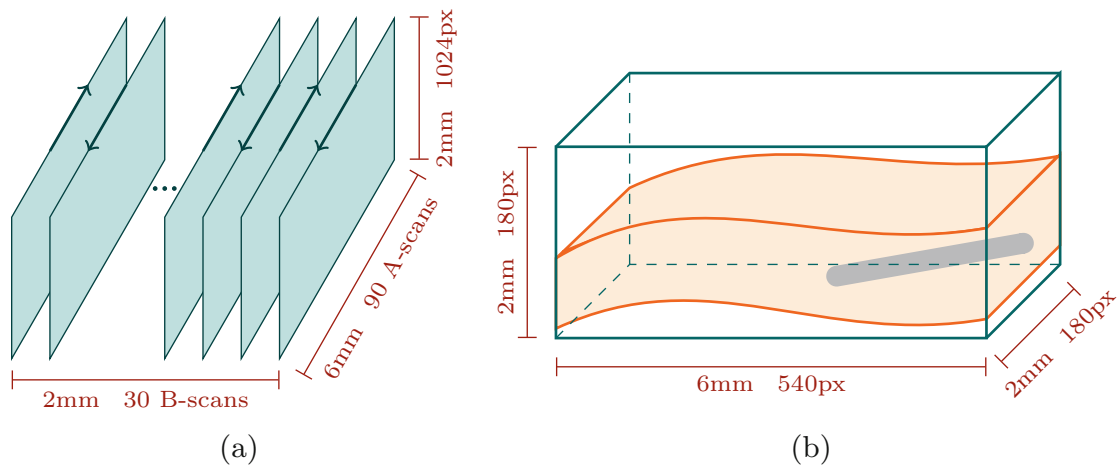
**Fig. 2.** (a): The modified pattern of OCT acquisition. (b): The lateral visualization of the cornea (orange) and the surgical needle (gray) in an OCT cuboid.

voxels (Fig. 2b). For that, frames are first averaged along the depth to obtain 30 frames of $90 \times 30$ pixels. Then in each cell of the grid, a tricubic interpolant which maps coordinates to intensity values is defined as follows:

$$f(x, y, z) = \sum_{i,j,k=0}^{3} c_{ijk} x^i y^j z^k, \qquad x, y, z \in [0, 1], \tag{1}$$

in which $c_{ijk}$ are the 64 interpolant coefficients calculated locally from the grid sample points and their derivatives. The coefficients are calculated by multiplication of a readily available $64 \times 64$ matrix and the vector of 64 elements consisting of 8 sample points and their derivatives [6]. The interpolation is implemented on the CPU in a parallel fashion.

## 2.2 Visualization

The achieved 3D OCT volume is visualized on both 2D monitors using GPU ray casting with 100 rays per pixel. Maximum information in OCT images is gained from high-intensity values representing boundaries between tissue layers. Hence, the Maximum Intensity Projection (MIP) technique is employed for rendering to put an emphasis on corneal layers. Many segmentation algorithms in OCT imaging are based on adaptive intensity thresholding [5]. Metallic surgical instruments including typical needles used for the BB-DALK procedure have infrared reflectivity profiles that are distinct from cellular tissues. The 3D OCT volume is segmented into the background, the cornea and the instrument by taking advantage of various reflectivity profiles and employing K-means clustering. The initial cluster mean values are set for the background to zero, the cornea to the volume mean intensity ($\mu$) and the instrument to the volume mean intensity plus two standard deviations ($\mu + 2\sigma$). The segmentation is used to dynamically alter the color and opacity transfer functions to ensure the instrument is distinctly and continuously visualized in red, the background speckle noise is suppressed and the corneal tissue opacity does not obscure the instrument (Fig. 3b, c).



(a)      (b)      (c)
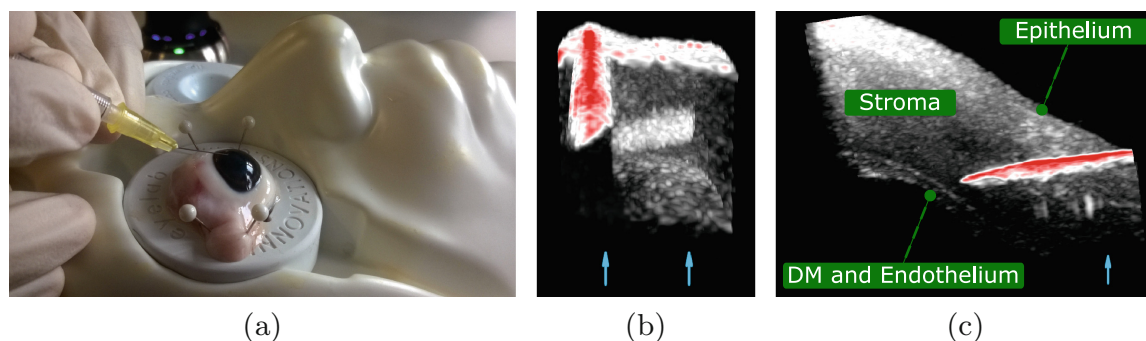
**Fig. 3.** (a): Needle insertion performed by the surgeon on the *ex vivo* pig eye. (b), (c): 3D visualization of the OCT cuboid with frontal and lateral viewpoints. The needle is distinctly visualized in red while endothelium (arrow) is not apparent.

The OCT cuboid could be examined from different viewpoints according to the exact need of the surgeon. For this purpose, one of the two displays could be controlled by the surgeon's assistant using a 3D mouse with zooming, panning and 3D rotating functionalities. The proposed guidance system maintains an automatic viewpoint of the OCT volume next to the microscopic view in the second display using the tracking procedure described below.

## 2.3 Tracking

The corneal DM and endothelial layer are the main targets of the BB-DALK procedure. The DM must not be perforated while the needle must be guided as close as possible to it. However, the two layers combined do not have a footprint larger than a few pixels in OCT images. As an essential part of the guidance system, DM and endothelium 3D surfaces are tracked for continuous feedback by solid visualization. The advancement of the needle in a BB-DALK procedure is examined and reported by percentage of the stroma that is above the needle tip. Hence, the epithelium surface of the cornea is also tracked to assist the surgeon by the quantitative guidance of the insertion.

Tracking in each volume is initiated by detection of the topmost and bottommost 3D points in the segmented cornea of the OCT volume. Based on the spherical shape of the cornea, two half spheres are considered as models of the endothelium and epithelium surfaces. The models are then fitted to the detected point clouds using iterative closest point (ICP) algorithm. Since the insertion of the needle deforms the cornea, ICP is utilized with 3D affine transformation at its core [3]. If the detected and the model half sphere point clouds are respectively denoted as $P = \{p_i\}_{i=1}^{N_P} \in \mathbb{R}^3$ and $M = \{m_i\}_{i=1}^{N_M} \in \mathbb{R}^3$, each iteration of the tracker algorithm is consecutively minimizing the following functions:

$$C(i) = \underset{j \in \{1,...,N_P\}}{\arg\min} \|(A_{k-1}m_i + t_{k-1}) - p_i\|_2^2, \qquad \text{for all } i \in \{1,..,N_M\}. \quad (2)$$

$$(A_k, t_k) = \underset{A,t}{\arg\min} \frac{1}{N} \sum_{i=1}^{N} \|(Am_i + t) - p_{C(i)}\|_2^2. \quad (3)$$

Equation 2 finds the correspondence $C(i)$ between $N \leq min(N_P, N_M)$ detected and model points. Equation 3 minimizes the Euclidean distance between the detected points and the transformed points of the model. $A_k$ and $t_k$ are the desired affine and translation matrices at iteration $k$. For each incoming volume, ICP is initialized by the transformation that brings the centroid of the model points to the centroid of the detected points. The algorithm stops after 30 iterations.

The lateral view of the OCT volume gives a better understanding of the needle dimensions and advancement. Also, the perception of the distance between the instrument and the endothelium layer is best achieved from viewing the scene parallel to the surface. Therefore, the viewpoint of the second display is constantly kept parallel to a small plane at the center of the tracked endothelium
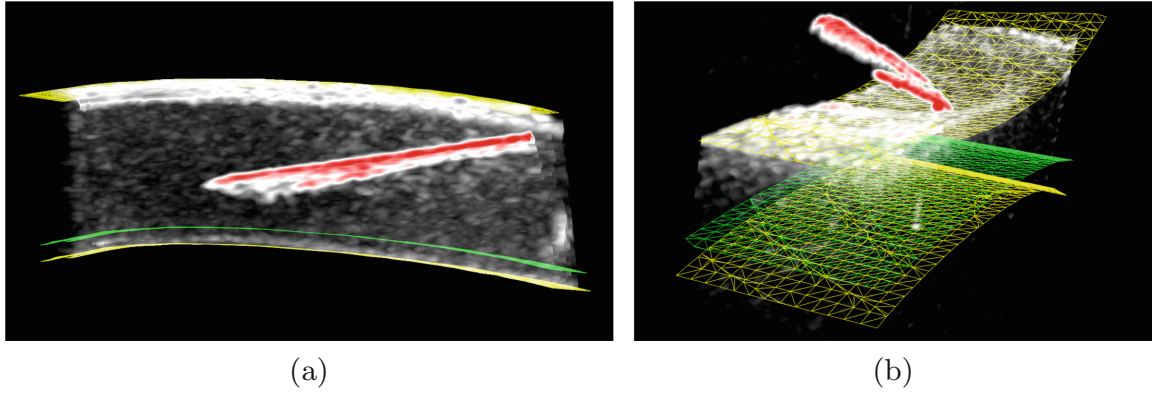
**Fig. 4.** Augmented Reality is used to solidly visualize the endothelium and epithelium surfaces (yellow) using wireframes. A hypothetical surface (green) is rendered to indicate the insertion target depth.

surface (Fig. 4a). The pressure applied for insertion of the needle leads to deformation of the cornea. To keep the OCT field of view centered on the focus of the procedure despite the induced shifts, the OCT depth range is continuously centered to halfway between top and bottom surfaces. This is done automatically to take the burden of manual repositioning away from the surgeon.

## 2.4 Augmented Reality

To further assist the surgeon, a hypothetical third surface is composed between the top and bottom surfaces indicating the insertion target depth (Fig. 4). The surgeon can choose a preferred percentage of penetration at which the imaginary surface would be rendered. Each point of the third surface is a linear combination of the corresponding points on the tracked epithelium and endothelium layers according to the chosen percentage. To visualize the detected surfaces, a wireframe mesh is formed on each of the three point sets. The two detected surfaces are rendered in yellow at their tracked position and the third surface is rendered
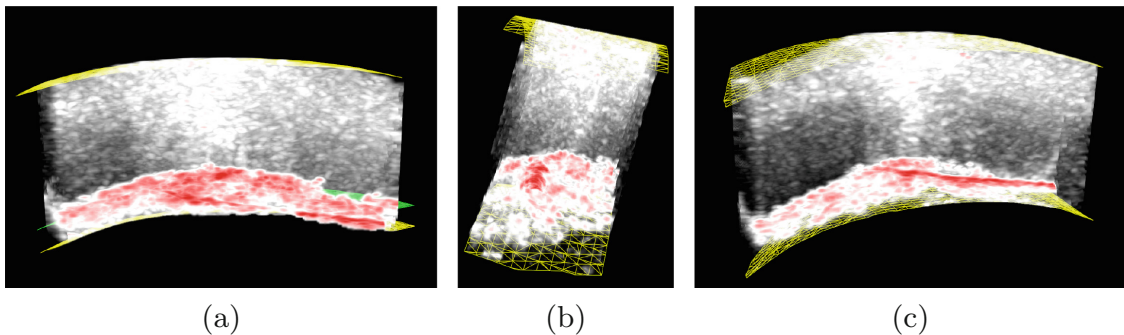


**Fig. 5.** Results of air injection in multiple pig eyes visualized from various viewpoints. The concentration of air in the bottommost region of the cornea indicates the high insertion accuracy. Deep stroma is reached with no sign of perforation.

in green at its hypothetical location. Visualization of each surface could be turned off if necessary. After injection, the presence of air leads to high-intensity voxels in the OCT volume. Therefore, the separation region is visualized effectively in red and could be used for validation of separation (Fig. 5).
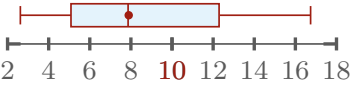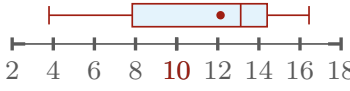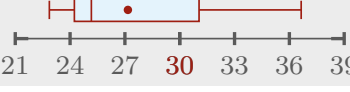
## 3   Experiments and Results

The proposed guidance system is tested by an ophthalmic surgeon experienced in corneal transplantation procedure on several *ex vivo* pig eyes. The visualization gives a new dimension never seen before in conventional systems in his comment. His experience with the system signifies the ability of the real-time guidance solution to help in deep needle insertions with fewer perforation incidents.

For the purpose of real-time OCT acquisition, the surgical scene is sparsely sampled via a grid of A-scans and interpolated. To evaluate the accuracy of interpolation against dense sampling, four fixed regions of a phantom eye ($2\,mm \times 6\,mm \times 2\,mm$) are scanned once with real-time sparse sampling ($30\,px \times 90\,px \times 1024\,px$) and two times with slow dense sampling ($85\,px \times 512\,px \times 1024\,px$). The sparse volumes are then interpolated to the size of the dense volumes. Volume pixels have intensities in the range of $[0, 1]$. For each of the four regions, Mean Absolute Error (MAE) of pixel intensities is once calculated for the two dense captures and once for one of the dense volumes and the interpolated volume. A maximum pixel intensity error of 0.073 is observed for the dense-sparse comparison while a minimum pixel intensity error of 0.043 is observed for the dense-dense comparison. The reason for the observed error in dense-dense comparison lies in the presence of OCT speckle noise which is a known phenomenon. The error observed for the dense-sparse comparison is comparable with the error induced by speckle noise hence the loss in sparse sampling is insignificant.

Human corneal thickness is reported to be around $500\,\mu m$. To ensure a minimum chance of perforation when insertion is done to the depth of $90\,\%$, the tracking accuracy required is around $50\,\mu m$. To evaluate tracking accuracy of the proposed solution, a micromanipulator with a resolution of $5\,\mu m$ is used. A phantom eye and a pig eye are fixed to a micromanipulator and precisely moved upwards and downwards while the epithelium and endothelium surfaces are tracked. At each position, the change in the depth of the tracked surfaces corresponding points are studied. Results are presented in Table 1 using box-and-whisker plots. The whiskers are showing the minimum and maximum recorded change of all tracked points while the start and the end of the box are the first and third quartiles. Bands and dots represent medians and means of the recorded changes respectively. The actual value against which the tracking accuracy should be compared is highlighted in red on the horizontal axis of the plots. Overall, the maximum tracking error is $8.8\,\mu m$.

**Table 1.** Evaluation of Tracking



| Experiment | Actual move ($\mu m$) | Detected epithelium displacement ($\mu m$) | Detected endothelium displacement ($\mu m$) |
|---|---|---|---|
| Phantom eye | 10 | | |
| Phantom eye | 30 | | |
| Pig eye | 10 | | |
| Pig eye | 30 | | |

## 4 Conclusion

This work presents a novel real-time guidance system for one of the most challenging procedures in ophthalmic microsurgery. The use of medical AR aims at facilitation of the BB-DALK learning process. Experiments on *ex vivo* pig eyes suggest the usability and reliability of the system leading to more effective yet shorter surgery sessions. Quantitative evaluations of the system indicate its high accuracy in depicting the surgical scene and tracking its changes leading to precise and deep insertions. Future work will be in the direction of adding needle tracking and navigation, further evaluations and clinical *in vivo* tests.

## References

1. Anwar, M., Teichmann, K.D.: Big-bubble technique to bare Descemet's membrane in anterior lamellar keratoplasty. J. Cataract Refract. Surg. **28**(3), 398–403 (2002)
2. De Benito-Llopis, L., Mehta, J.S., Angunawela, R.I., Ang, M., Tan, D.T.: Intra-operative anterior segment optical coherence tomography: a novel assessment tool during deep anterior lamellar keratoplasty. Am. J. Ophthalmol. **157**(2), 334–341 (2014)
3. Du, S., Zheng, N., Ying, S., Liu, J.: Affine iterative closest point algorithm for point set registration. Pattern Recogn. Lett. **31**(9), 791–799 (2010)
4. Fontana, L., Parente, G., Tassinari, G.: Clinical outcomes after deep anterior lamellar keratoplasty using the big-bubble technique in patients with keratoconus. Am. J. Ophthalmol. **143**(1), 117–124 (2007)
5. Ishikawa, H., Stein, D.M., Wollstein, G., Beaton, S., Fujimoto, J.G., Schuman, J.S.: Macular segmentation with optical coherence tomography. Invest. Ophthalmol. Vis. Sci. **46**(6), 2012–2017 (2005)
6. Lekien, F., Marsden, J.: Tricubic interpolation in three dimensions. Int. J. Numer. Meth. Eng. **63**(3), 455–471 (2005)
7. Scorcia, V., Busin, M., Lucisano, A., Beltz, J., Carta, A., Scorcia, G.: Anterior segment optical coherence tomography-guided big-bubble technique. Ophthalmology **120**(3), 471–476 (2013)

# Bibliography

[1] M. K. Adam, S. Thornton, C. D. Regillo, C. Park, A. C. Ho, and J. Hsu. "Minimal endoillumination levels and display luminous emittance during three-dimensional heads-up vitreoretinal surgery". In: *Retina* 37.9 (2017), pp. 1746–1749 (cit. on p. 12).

[2] M. Alsheakhali, A. Eslami, H. Roodaki, and N. Navab. "CRF-based model for instrument detection and pose estimation in retinal microsurgery". In: *Computational and mathematical methods in medicine* 2016 (2016) (cit. on p. 60).

[3] M. Anwar and K. D. Teichmann. "Big-bubble technique to bare Descemet's membrane in anterior lamellar keratoplasty". In: *Journal of Cataract & Refractive Surgery* 28.3 (2002), pp. 398–403 (cit. on p. 43).

[4] M. Balicki, R. Richa, B. Vagvolgyi, et al. "Interactive OCT annotation and visualization for vitreoretinal surgery". In: *Workshop on Augmented Environments for Computer-Assisted Interventions*. Springer. 2012, pp. 142–152 (cit. on p. 24).

[5] K. U. Bartz-Schmidt, B. Kirchhof, and K. Heimann. "Primary vitrectomy for pseudophakic retinal detachment." In: *British journal of ophthalmology* 80.4 (1996), pp. 346–349 (cit. on p. 10).

[6] D. A. Belyea, S. E. Brown, and L. Z. Rajjoub. "Influence of surgery simulator training on ophthalmology resident phacoemulsification performance". In: *Journal of Cataract & Refractive Surgery* 37.10 (2011), pp. 1756–1761 (cit. on p. 45).

[7] J. W. Berger, M. E. Leventon, N. Hata, W. Wells, and R. Kikinis. "Design considerations for a computer-vision-enabled ophthalmic augmented reality environment". In: *CVRMed-MRCAS'97*. Springer. 1997, pp. 399–408 (cit. on p. 22).

[8] J. W. Berger and B. Madjarov. "Augmented reality fundus biomicroscopy: A working clinical prototype". In: *Archives of Ophthalmology* 119.12 (2001), pp. 1815–1818 (cit. on p. 23).

[9] J. Bergmeier, D. Kundrat, A. Schoob, L. A. Kahrs, and T. Ortmaier. "Methods for a fusion of optical coherence tomography and stereo camera image data". In: *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*. Vol. 9415. International Society for Optics and Photonics. 2015, p. 94151C (cit. on p. 25).

[10] P. R. Bhadri, A. P. Rowley, R. N. Khurana, et al. "Evaluation of a stereoscopic camera-based three-dimensional viewing workstation for ophthalmic surgery". In: *American journal of ophthalmology* 143.5 (2007), pp. 891–892 (cit. on p. 24).

[11] N. Bitterman. "Technologies and solutions for data display in the operating room". In: *Journal of clinical monitoring and computing* 20.3 (2006), pp. 165–173 (cit. on p. 26).

[12] D Black, B Kocev, H Meine, A Nabavi, and R Kikinis. "Towards uncertainty-aware auditory display for surgical navigation". In: *Computer Assisted Radiology and Surgery (in print)* (2016) (cit. on p. 28).

[13] D. Black, C. Hansen, A. Nabavi, R. Kikinis, and H. Hahn. "A survey of auditory display in image-guided interventions". In: *International journal of computer assisted radiology and surgery* 12.10 (2017), pp. 1665–1676 (cit. on p. 26).

[14] D. Black, J. Hettig, M. Luz, C. Hansen, R. Kikinis, and H. Hahn. "Auditory feedback to support image-guided medical needle placement". In: *International journal of computer assisted radiology and surgery* 12.9 (2017), pp. 1655–1663 (cit. on p. 28).

[15] G. E. Boynton and M. A. Woodward. "Evolving techniques in corneal transplantation". In: *Current surgery reports* 3.2 (2015), p. 2 (cit. on p. 8).

[16] G. J. Chader and A. Taylor. "Preface: the aging eye: normal changes, age-related diseases, and sight-saving approaches". In: *Investigative ophthalmology & visual science* 54.14 (2013), ORSF1–ORSF4 (cit. on p. 2).

[17] K.-S. Choi, S. Soo, and F.-L. Chung. "A virtual training simulator for learning cataract surgery with phacoemulsification". In: *Computers in biology and medicine* 39.11 (2009), pp. 1020–1031 (cit. on p. 44).

[18] L. N. Darlow, S. S. Akhoury, and J. Connan. "A review of state-of-the-art speckle reduction techniques for optical coherence tomography fingertip scans". In: *Seventh International Conference on Machine Vision (ICMV 2014)*. Vol. 9445. International Society for Optics and Photonics. 2015, p. 944523 (cit. on p. 15).

[19] J. F. De Boer, R. Leitgeb, and M. Wojtkowski. "Twenty-five years of optical coherence tomography: the paradigm shift in sensitivity and speed provided by Fourier domain OCT". In: *Biomedical optics express* 8.7 (2017), pp. 3248–3280 (cit. on p. 16).

[20] M. Dewan, P. Marayong, A. M. Okamura, and G. D. Hager. "Vision-based assistance for ophthalmic micro-surgery". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2004, pp. 49–57 (cit. on p. 45).

[21] B. J. Dixon, M. J. Daly, H. Chan, A. Vescan, I. J. Witterick, and J. C. Irish. "Augmented real-time navigation with critical structure proximity alerts for endoscopic skull base surgery". In: *The Laryngoscope* 124.4 (2014), pp. 853–859 (cit. on p. 27).

[22] B. J. Dixon, M. J. Daly, H. Chan, A. D. Vescan, I. J. Witterick, and J. C. Irish. "Surgeons blinded by enhanced navigation: the effect of augmented reality on attention". In: *Surgical endoscopy* 27.2 (2013), pp. 454–461 (cit. on p. 36).

[23] W. Drexler and J. G. Fujimoto. *Optical coherence tomography: technology and applications*. Springer Science & Business Media, 2008 (cit. on p. 16).

[24] S. Duca, K. Filippatos, J. Straub, et al. "Evaluation of Automatic Following of Anatomical Structures for Live Intraoperative OCT Repositioning". In: *Investigative Ophthalmology & Visual Science (IOVS)* 58.8 (2017), pp. 4832–4832 (cit. on p. 60).

[25] C. Eckardt and E. B. Paulo. "Heads-up surgery for vitreoretinal procedures: an experimental and clinical study". In: *Retina* 36.1 (2016), pp. 137–147 (cit. on p. 13).

[26] P. Edwards, A. King, D. Hawkes, et al. "Stereo augmented reality in the surgical microscope." In: *Studies in health technology and informatics* 62 (1999), pp. 102–108 (cit. on p. 22).

[27] T. Edwards, K Xue, H. Meenink, et al. "First-in-human study of the safety and viability of intraocular robotic surgery". In: *Nature biomedical engineering* 2.9 (2018), p. 649 (cit. on p. 11).

[28] J. P. Ehlers, W. J. Dupps, P. K. Kaiser, et al. "The prospective intraoperative and perioperative ophthalmic imaging with optical coherence tomography (PIONEER) study: 2-year results". In: *American journal of ophthalmology* 158.5 (2014), pp. 999–1007 (cit. on p. 17).

[29] J. P. Ehlers, J. Goshe, W. J. Dupps, et al. "Determination of feasibility and utility of microscope-integrated optical coherence tomography during ophthalmic surgery: the DISCOVER Study RESCAN Results". In: *JAMA ophthalmology* 133.10 (2015), pp. 1124–1132 (cit. on p. 17).

[30] J. P. Ehlers, P. K. Kaiser, and S. K. Srivastava. "Intraoperative optical coherence tomography using the RESCAN 700: preliminary results from the DISCOVER study". In: *British journal of Ophthalmology* 98.10 (2014), pp. 1329–1332 (cit. on p. 17).

[31] A. Eslami, S. Duca, H. Roodaki, et al. "Incremental Enhancement of Live Intraoperative OCT Scans by Temporal Analysis". In: *Investigative Ophthalmology & Visual Science (IOVS)* 58.8 (2017), pp. 3124–3124 (cit. on p. 60).

[32] A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser. "Optical coherence tomography-principles and applications". In: *Reports on progress in physics* 66.2 (2003), p. 239 (cit. on pp. 13, 14).

[33] I. N. Fleming, S. Voros, B. Vagvolgyi, et al. "Intraoperative visualization of anatomical targets in retinal surgery". In: *2008 IEEE Workshop on Applications of Computer Vision*. IEEE. 2008, pp. 1–6 (cit. on p. 23).

[34] J. C. Folk, R. A. Adelman, C. J. Flaxel, L. Hyman, J. S. Pulido, and T. W. Olsen. "Idiopathic epiretinal membrane and vitreomacular traction Preferred Practice Pattern® Guidelines". In: *Ophthalmology* 123.1 (2016), P152–P181 (cit. on p. 10).

[35] G. Y. Fujii, E. de Juan Jr, M. S. Humayun, et al. "A new 25-gauge instrument system for transconjunctival sutureless vitrectomy surgery". In: *Ophthalmology* 109.10 (2002), pp. 1807–1812 (cit. on p. 9).

[36] S. N. Gillan and G. M. Saleh. "Ophthalmic surgical simulation: a new era". In: *JAMA ophthalmology* 131.12 (2013), pp. 1623–1624 (cit. on p. 44).

[37] C. A. Giller, A. M. Murro, Y. Park, S. Strickland, and J. R. Smith. "EEG sonification for epilepsy surgery: A clinical work-in progress". In: Georgia Institute of Technology. 2012 (cit. on p. 27).

[38] P. K. Gupta, P. S. Jensen, and E. de Juan. "Surgical forces and tactile perception during retinal microsurgery". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 1999, pp. 1218–1225 (cit. on pp. 2, 21).

[39] R. S. Haluck and T. M. Krummel. "Computers and virtual reality for surgical education in the 21st century". In: *Archives of surgery* 135.7 (2000), pp. 786–792 (cit. on p. 44).

[40] C. Hansen, D. Black, C. Lange, et al. "Auditory support for resection guidance in navigated liver surgery". In: *The International Journal of Medical Robotics and Computer Assisted Surgery* 9.1 (2013), pp. 36–43 (cit. on pp. 27, 28).

[41] P.-A. Heng, C.-Y. Cheng, T.-T. Wong, et al. "A virtual-reality training system for knee arthroscopic surgery". In: *IEEE Transactions on Information Technology in Biomedicine* 8.2 (2004), pp. 217–227 (cit. on p. 44).

[42] A. Jain, R. Bansal, A. Kumar, and K. Singh. "A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students". In: *International Journal of Applied and Basic Medical Research* 5.2 (2015), p. 124 (cit. on p. 37).

[43] M. Kalia, C. S. zu Berge, H. Roodaki, C. Chakraborty, and N. Navab. "Interactive depth of focus for improved depth perception". In: *International Conference on Medical Imaging and Augmented Reality (MIAR)*. Springer,Cham. 2016, pp. 221–232 (cit. on p. 60).

[44] S. Khwarg, F. Linstone, S. Daniels, et al. "Incidence, risk factors, and morphology in operating microscope light retinopathy." In: *American journal of ophthalmology* 103.3 Pt 1 (1987), pp. 255–263 (cit. on p. 12).

[45] S. Kishi. "Impact of swept source optical coherence tomography on ophthalmology". In: *Taiwan journal of ophthalmology* 6.2 (2016), pp. 58–68 (cit. on p. 16).

[46] T. Kohnen, M. Baumeister, D. Kook, O. K. Klaproth, and C. Ohrloff. "Cataract surgery with implantation of an artificial lens". In: *Deutsches Ärzteblatt International* 106.43 (2009), p. 695 (cit. on p. 2).

[47] I. Kozak, P. Banerjee, J. Luo, and C. Luciano. "Virtual reality simulator for vitreoretinal surgery using integrated OCT data". In: *Clinical ophthalmology (Auckland, NZ)* 8 (2014), p. 669 (cit. on pp. 44, 45).

[48] T. M. Krummel. "Surgical simulation and virtual reality: the coming revolution." In: *Annals of surgery* 228.5 (1998), p. 635 (cit. on p. 44).

[49] G Lafortune, G Balestat, and A Durand. "Comparing activities and performance of the hospital sector in Europe: how many surgical procedures performed as inpatient and day cases". In: *Labour and Social Affairs* (2012) (cit. on p. 2).

[50] M. A. Leitritz, F. Ziemssen, D. Suesskind, et al. "Critical evaluation of the usability of augmented reality ophthalmoscopy for the training of inexperienced examiners". In: *Retina* 34.4 (2014), pp. 785–791 (cit. on p. 24).

[51] J. Marescaux, J.-M. Clément, V. Tassetti, et al. "Virtual reality applied to hepatic surgery simulation: the next revolution." In: *Annals of surgery* 228.5 (1998), p. 627 (cit. on p. 44).

[52] S. Matinfar, M. A. Nasseri, U. Eck, et al. "Surgical soundtracks: automatic acoustic augmentation of surgical procedures". In: *International journal of computer assisted radiology and surgery (IJCARS)* 13.9 (2018), pp. 1345–1355 (cit. on p. 59).

[53] S. Matinfar, M. A. Nasseri, U. Eck, et al. "Surgical soundtracks: Towards automatic musical augmentation of surgical procedures". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer,Cham. 2017, pp. 673–681 (cit. on p. 59).

[54] C. McAlinden. "Corneal refractive surgery: past to present". In: *Clinical and Experimental Optometry* 95.4 (2012), pp. 386–398 (cit. on p. 7).

[55] C. A. McCannel. "Simulation surgical teaching in ophthalmology". In: *Ophthalmology* 122.12 (2015), pp. 2371–2372 (cit. on p. 45).

[56] C. A. McCannel, D. C. Reed, and D. R. Goldman. "Ophthalmic surgery simulator training improves resident performance of capsulorhexis in the operating room". In: *Ophthalmology* 120.12 (2013), pp. 2456–2461 (cit. on p. 45).

[57] R. McCloy and R. Stone. "Virtual reality in surgery". In: *Bmj* 323.7318 (2001), pp. 912–915 (cit. on p. 44).

[58] J. R. Merril, N. F. Notaroberto, D. M. Laby, A. M. Rabinowitz, and T. E. Piemme. "The Ophthalmic Retrobulbar Injection Simulator (ORIS): an application of virtual reality to medical education." In: *Proceedings of the Annual Symposium on Computer Application in Medical Care*. American Medical Informatics Association. 1992, p. 702 (cit. on p. 44).

[59] N. Navab, T. Blum, L. Wang, A. Okur, and T. Wendler. "First deployments of augmented reality in operating rooms". In: *Computer* 45.7 (2012), pp. 48–55 (cit. on p. 3).

[60] P. F. Neumann, L. L. Sadler, and J. Gieser. "Virtual reality vitrectomy simulator". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 1998, pp. 910–917 (cit. on p. 44).

[61] S. R. Palma, B. C. Becker, L. A. Lobes, and C. N. Riviere. "Comparative evaluation of monocular augmented-reality display for surgical microscopes". In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2012, pp. 1409–1412 (cit. on p. 24).

[62] Y. Peng, L. Tang, and Y. Zhou. "Subretinal injection: a review on the novel route of therapeutic delivery for vitreoretinal diseases". In: *Ophthalmic research* 58.4 (2017), pp. 217–226 (cit. on p. 11).

[63] F Peugnet, P Dubois, and J. Rouland. "Virtual reality versus conventional training in retinal photocoagulation: a first clinical assessment". In: *Computer Aided Surgery: Official Journal of the International Society for Computer Aided Surgery (ISCAS)* 3.1 (1998), pp. 20–26 (cit. on p. 45).

[64] J. Plazak, S. Drouin, L. Collins, and M. Kersten-Oertel. "Distance sonification in image-guided neurosurgery". In: *Healthcare technology letters* 4.5 (2017), pp. 199–203 (cit. on p. 27).

[65] D. Z. Reinstein, T. J. Archer, and M. Gobbe. "Small incision lenticule extraction (SMILE) history, fundamentals of a new refractive surgery technique and clinical outcomes". In: *Eye and Vision* 1.1 (2014), p. 3 (cit. on p. 7).

[66] D. Z. Reinstein, M. Gobbe, T. J. Archer, R. H. Silverman, and D. J. Coleman. "Epithelial thickness in the normal cornea: three-dimensional display with Artemis very high-frequency digital ultrasound". In: *Journal of refractive surgery* 24.6 (2008), pp. 571–581 (cit. on p. 7).

[67] B. Reitinger, A. Bornik, R. Beichel, and D. Schmalstieg. "Liver surgery planning using virtual reality". In: *IEEE Computer Graphics and Applications* 26.6 (2006), pp. 36–47 (cit. on p. 44).

[68] S. Resnikoff, W. Felch, T.-M. Gauthier, and B. Spivey. "The number of ophthalmologists in practice and training worldwide: a growing gap despite more than 200 000 practitioners". In: *British Journal of Ophthalmology* 96.6 (2012), pp. 783–787 (cit. on p. 2).

[69] S. Resnikoff, D. Pascolini, D. Etya'Ale, et al. "Global data on visual impairment in the year 2002". In: *Bulletin of the world health organization* 82 (2004), pp. 844–851 (cit. on p. 6).

[70] M. Röhlig, P. Rosenthal, C. Schmidt, H. Schumann, and O. Stachs. "Visual Analysis of Optical Coherence Tomography Data in Ophthalmology." In: *EuroVA@ EuroVis*. 2017, pp. 37–41 (cit. on p. 46).

[71] H. Roodaki, W.-J. Chen, D. Zapp, and A. Eslami. "Simulated keratometry using microscope-integrated optical coherence tomography". In: *Investigative Ophthalmology & Visual Science* 60.9 (2019), pp. 2128–2128 (cit. on p. 59).

[72] H. Roodaki, K. Filippatos, A. Eslami, and N. Navab. "Introducing augmented reality to optical coherence tomography in ophthalmic microsurgery". In: *2015 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2015, pp. 1–6 (cit. on p. 60).

[73] H. Roodaki, M. Grimm, N. Navab, and A. Eslami. "Real-time Scene Understanding in Ophthalmic Anterior Segment OCT Images". In: *Investigative Ophthalmology & Visual Science* 60.11 (2019), PB095–PB095 (cit. on p. 59).

[74] H. Roodaki, N. Navab, A. Eslami, C. Stapleton, and N. Navab. "Sonifeye: Sonification of visual information using physical modeling sound synthesis". In: *IEEE transactions on visualization and computer graphics (TVCG)* 23.11 (2017), pp. 2366–2371 (cit. on p. 59).

[75] H. Roodaki, C. A. di San Filippo, D. Zapp, N. Navab, and A. Eslami. "A surgical guidance system for big-bubble deep anterior lamellar keratoplasty". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer,Cham. 2016, pp. 378–385 (cit. on p. 60).

[76] M. A. Sagar, D. Bullivant, G. D. Mallinson, and P. J. Hunter. "A virtual environment and model of the eye for surgical simulation". In: *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. ACM. 1994, pp. 205–212 (cit. on p. 44).

[77] G. M. Saleh, J. Lamparter, P. M. Sullivan, et al. "The international forum of ophthalmic simulation: developing a virtual reality training curriculum for ophthalmology". In: *British Journal of Ophthalmology* 97.6 (2013), pp. 789–792 (cit. on p. 45).

[78] J. M. Schmitt. "Optical coherence tomography (OCT): a review". In: *IEEE Journal of selected topics in quantum electronics* 5.4 (1999), pp. 1205–1215 (cit. on pp. 14, 15).

[79] J. M. Schmitt, S. Xiang, and K. M. Yung. "Speckle in optical coherence tomography". In: *Journal of biomedical optics* 4.1 (1999), pp. 95–106 (cit. on p. 15).

[80] J. P. Schulze, C. Schulze-Döbold, A. Erginay, and R. Tadayoni. "Visualization of three-dimensional ultra-high resolution OCT in virtual reality." In: *MMVR*. 2013, pp. 387–391 (cit. on p. 46).

[81] M. Selvander and P. Åsman. "Virtual reality cataract surgery training: learning curves and concurrent validity". In: *Acta ophthalmologica* 90.5 (2012), pp. 412–417 (cit. on p. 45).

[82] J. H. Shuhaiber. "Augmented reality in surgery". In: *Archives of surgery* 139.2 (2004), pp. 170–174 (cit. on p. 22).

[83] T. Sielhorst, M. Feuerstein, and N. Navab. "Advanced medical displays: A literature review of augmented reality". In: *Journal of Display Technology* 4.4 (2008), pp. 451–467 (cit. on p. 19).

[84] S. Singhy and C. Riviere. "Physiological tremor amplitude during retinal microsurgery". In: *Proceedings of the IEEE 28th Annual Northeast Bioengineering Conference (IEEE Cat. No. 02CH37342)*. IEEE. 2002, pp. 171–172 (cit. on p. 2).

[85] C. V. Stewart, C.-L. Tsai, and B. Roysam. "The dual-bootstrap iterative closest point algorithm with application to retinal image registration". In: *IEEE transactions on medical imaging* 22.11 (2003), pp. 1379–1394 (cit. on p. 23).

[86] J. Teichmann, M. Valtink, M. Nitschke, et al. "Tissue engineering of the corneal endothelium: a review of carrier materials". In: *Journal of functional biomaterials* 4.4 (2013), pp. 178–208 (cit. on p. 7).

[87] A. S. S. Thomsen, D. Bach-Holm, H. Kjærbo, et al. "Operating room performance improves after proficiency-based virtual reality cataract surgery training". In: *Ophthalmology* 124.4 (2017), pp. 524–531 (cit. on p. 45).

[88] A. S. S. Thomsen, P. Smith, Y. Subhi, et al. "High correlation between performance on a virtual-reality simulator and real-life cataract surgery". In: *Acta ophthalmologica* 95.3 (2017), pp. 307–311 (cit. on p. 45).

[89] S. Tuft and D. Coster. "The corneal endothelium". In: *Eye* 4.3 (1990), p. 389 (cit. on p. 7).

[90] D. Verma, D. Wills, and M Verma. "Virtual reality simulator for vitreoretinal surgery". In: *Eye* 17.1 (2003), p. 71 (cit. on p. 44).

[91] P. Vickers and A. Imam. "The use of audio in minimal access surgery". In: (1999) (cit. on p. 26).

[92] J. Vroomen and B. d. Gelder. "Sound enhances visual perception: cross-modal effects of auditory organization on vision." In: *Journal of experimental psychology: Human perception and performance* 26.5 (2000), p. 1583 (cit. on p. 37).

[93] F. Wang, T. Poston, C. L. Teo, K. M. Lim, and E. Burdet. "Multisensory learning cues using analytical collision detection between a needle and a tube". In: *12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2004. HAPTICS'04. Proceedings.* IEEE. 2004, pp. 339–346 (cit. on p. 26).

[94] G. O. Waring III, W. M. Bourne, H. F. Edelhauser, and K. R. Kenyon. "The corneal endothelium: normal and pathologic structure and function". In: *Ophthalmology* 89.6 (1982), pp. 531–590 (cit. on p. 8).

[95] J. R. Wilkins, C. A. Puliafito, M. R. Hee, et al. "Characterization of epiretinal membranes using optical coherence tomography". In: *Ophthalmology* 103.12 (1996), pp. 2142–2151 (cit. on p. 10).

[96] G. Williams. "25-, 23-, or 20-gauge instrumentation for vitreous surgery?" In: *Eye* 22.10 (2008), p. 1263 (cit. on p. 10).

[97] K Xue, M Groppe, A. Salvetti, and R. MacLaren. "Technique of retinal gene therapy: delivery of viral vector into the subretinal space". In: *Eye* 31.9 (2017), p. 1308 (cit. on p. 11).

[98] Z. Yaqoob, J. Wu, and C. Yang. "Spectral domain optical coherence tomography: a better OCT imaging strategy". In: *Biotechniques* 39.6 (2005), S6–S13 (cit. on p. 16).

[99] E. Yiannakopoulou, N. Nikiteas, D. Perrea, and C. Tsigris. "Virtual reality simulators and training in laparoscopic surgery". In: *International Journal of Surgery* 13 (2015), pp. 60–64 (cit. on p. 44).

[100] M. Zhou, X. Hao, A. Eslami, et al. "6DOF Needle Pose Estimation for Robot-Assisted Vitreoretinal Surgery". In: *IEEE Access* 7 (2019), pp. 63113–63122 (cit. on p. 11).

[101] M. Zhou, K. Huang, A. Eslami, et al. "Beveled needle position and pose estimation based on optical coherence tomography in ophthalmic microsurgery". In: *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE. 2017, pp. 308–313 (cit. on p. 59).

[102] M. Zhou, K. Huang, A. Eslami, et al. "Precision Needle Tip Localization Using Optical Coherence Tomography Images for Subretinal Injection". In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 1–8 (cit. on p. 59).

[103] M. Zhou, H. Roodaki, A. Eslami, et al. "Needle segmentation in volumetric optical coherence tomography images for ophthalmic microsurgery". In: *Applied Sciences,Special Issue Development and Application of Optical Coherence Tomography (OCT)* 7.8 (2017), p. 748 (cit. on p. 60).

# List of Figures