TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Robotik, Künstliche Intelligenz und Echtzeitsysteme

Towards Autonomy in Vitreoretinal Surgery: Robotic Subretinal Injection with Optical Tomography Coherence Images Navigation

Mingchuan Zhou

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Prof. Dr. Susanne Albers
Prüfer der Dissertation:	1.	Prof. Dr.-Ing. habil. Alois Knoll
	2.	Prof. Dr. Iulian I. Iordachita
	3.	Prof. Dr. Kai Huang

Die Dissertation wurde am 08.05.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 22.10.2019 angenommen.

Abstract

Vitreoretinal diseases are the main cause of the blindness in the worldwide. Some of these diseases, e.g. aged-related macular degeneration (AMD) and subretinal hemorrhage, require the drug to be injected directly into subretinal area. Subretinal injection is known to be a challenging operation, which requires very precise, dexterous manipulation, and profound experience. Due to this fact, the ophthalmologists are required not only to have enough training and clinical experiences but also need to be in a good physical condition to cope with hand tremors and a non-stop surgery session. However, for surgeons, clinical experience normally comes with age, but aging impairs physical skills. Image guided robot-assisted surgery (RAS) is a promising solution that brings significant improvements in outcomes and reduces the physical limitations of human surgeons. A certain of autonomy with image guided RAS has been shown to assist surgeon to obtain better surgical outcome. The main reasons behind are that the surgeon can obtain enhance information which is normally hard to observe by traditional modality and free the surgeon to focus more on the surgical plan and intraoperative decision.

In this thesis, we work on the safety evaluation of robot with remote center of motion (RCM) and how to perform the robotic subretinal injection with microscope intergraded optical coherence tomography (OCT) images navigation. Especially, we focus on the needle tip&pose estimation when the needle tip is above the retina, hand (robot) and eye (OCT device) calibration, and needle insertion depth detection when needle tip is under retina. All these technical problems and corresponding solutions are proposed towards giving the robotic subretinal injection a certain level of autonomy.

The results of this thesis try to feedback the possibility of autonomy in robotic assisted vitreoretinal surgery with OCT navigation. The main content developed for this thesis were published in international conferences and journals, indicating the credibility and novelty of the work.

Zusammenfassung

Vitreoretinale Erkrankungen sind weltweit die häufigste Ursache für Erblindung. Für die Therapie einiger dieser Krankheiten wie beispielsweise altersbedingte Makulardegeneration (AMD) und subretinale Blutungen wird dem Patienten ein Medikament direkt in den subretinalen Bereich verabreicht. Die subretinale Injektion ist dafür bekannt eine herausfordernder chirurgischer Eingriff zu sein, die sehr präzise und geschickte Handhabung der Instrumente und ein hohes Maß an Erfahrung benötigt. Aus diesem Grund benötigt der behandelnde Augenarzt nicht nur ein gutes Training und klinische Erfahrung, sondern muss ebenfalls in bester physischer Verfassung sein um diesen Eingriff trotz Tremor der Hand und einer langen ununterbrochenen Operation erfolgreich abzuschließen. Jedoch kommt die klinische Erfahrung bei Chirurgen gewöhnlich mit dem Alter, womit jedoch auch physische Einschränkungen einhergehen. Bild- und robotergestützte Chirurgie (RAS) ist eine vielversprechende Lösung dieses Problems, die sowohl bedeutende Verbesserungen der chirurgischen Ergebnisse liefert also auch die physische Limitierung des Chirurgen reduziert. Autonome Arbeitsschritte mit bild- und robotergestützter Chirurgie haben gezeigt, dass es mit ihrer Hilfe möglich ist dem Chirurgen zu bessern Ergebnissen des Eingriffs zu verhelfen. Der größte Vorteil hierbei liegt darin, dass der Chirurg zusätzlich Informationen bekommen kann, die mit den klassischen Methoden nur schwer zu beobachten wären. Dies ermöglicht dem Chirurgen sich stärker auf den Plan der Operation und intraoperative Entscheidungen zu konzentrieren.

In dieser Arbeit wird eine Beurteilung der Sicherheit eines Roboter, der sich um ein außerhalb des Roboters liegendes Zentrum (RCM) bewegen kann. Des Weiteren wird wird die Durchführung einer robotergestützten subretinalen Injektion mit bildgenützter Navigation durch ein in ein Mikroskop integrierten optischen Kohärenz-Tomographen (OCT) untersucht. Hier wird vor allem auf die Bestimmung der Nadel-Pose oberhalb der Retina, die sog. Hand-Auge-Kalibrierung des Roboters mit dem OCT-Gerät und die Detektion der Nadelspitze unterhalb der retinalen Oberfläche

eingegangen. All diese Technischen Probleme und deren Lösungen werden im Hinblick darauf entwickelt dem Roboter ein möglichst hohes Maß an Autonomie zu verleihen.

Das Ergebnis dieser Arbeit versucht die Möglichkeiten von Autonomen Arbeitsschritten bei robotergestützter vitreoretinaler Chirurgie unter OCT-Naviation zu beurteilen. Der inhaltliche Kern dieser Arbeit wurden auf internationalen Konferenzen und Journalen veröffentlich. Dies zeigt die Zuverlässigkeit der Ergebnisse und Neuartigkeit der Ansätze.

Acknowledgements

First of all, I would like to express my sincere gratitude to Prof. Dr. Alois Knoll for giving me the opportunity of a PhD study. Without his support, this thesis would have not been possible. His academic guidance, together with his willingness to share his experience, have benefited myself a lot during the period.

I would like to thank Prof. Susanne Albers being president at my defense. I would also like to thank the member of my defense committee, Prof. Iulian I. Iordachita. His reviews, comments and fruitful suggestions helped me a lot to improve the quality of this thesis. I would like to thank Prof. Kai Huang for his guidance in research, paper writing, and being the member of my defense committee.

I would like to thank Dr. M. Ali Nasseri for his guidance, inspirations on the research direction, sustained support for this project, and bringing me to surgical robot field. I would like to thank Prof. Dr. Nassir Navab for a lot of guidance in his group. I would like to thank Dr. Eslami Abouzar for giving me the chance to participant his research and use imaging device in Zeiss Meditec AG. I would like to thank my senior colleague Dr. Feihu Zhang, Dr. Caicai Xia, Dr. Long Cheng, Dr. Biaohu, and Dr. Guang Chen for their valuable suggestions during the study. I would also like to thank my colleague Mr. Zhenshan Bing and Mr. Xiebing Wang for the research cooperation; Mr. Hessam Roodaki and Mr. Jakob Weiss for the research cooperation and future advice; Ms. Silvia Magdici for being nice of three years of office mate. Furthermore, I would like to thank all my former and current colleagues of the whole Robotics and Embedded System group for their company and support, especially Ms. Amy Bücherl and Ms. Ute Lomp for their nice supporting. I also want to thank all the students whom I ever worked with. Finally, my dearest thanks go to my wife Dr. Di Cui for her continuous support, trust, passion, and encouragement. I am thankful to my five-month daughter Elena. You are so great in the past months which helps me great focusing on the thesis. I am thankful to my parents for their trust and encouragement in my study.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Eye is one of the most delicate and fragile organs in the human body. The first eye surgery can be dating back as early as 1800 BC, and the cataract treatment starting in the fifth century BC [2]. Some eye problems are minor and do not last long, but some can lead to a permanent loss of vision even blindness. In 2002, the World Health Organization analyzed available data about low vision and blindness from 55 countries which shows that 37 million people are blind and 124 million people have low vision [3]. The majority causes of blindness are cataract (48%), the glaucomas (12%), corneal scarring including trachoma (9%), age-related macular degeneration (AMD) (9%), and diabetic retinopathy (DR) (5%) [4]. In 2010, 39 million people were blind and 246 million people were vision-impaired [5]. The vitreoretinal surgery is an important method to help the patients to regain lost sight for diseases like AMD, DR and Retinal Vein Occlusion (RVO). AMD is the leading cause of blindness in developed countries [6]. The predicted population with AMD in 2020 is 196 million and will be increased up to 288 million in 2040 according to to [7], due to demographic changes and aging. It most commonly occurs in people over the age of fifty and in the United States is the most common cause of vision loss in this age group [2]. About 0.4% of people between 50 and 60 have the disease, while it occurs in 0.7% of people 60 to 70, 2.3% of those 70 to 80, and nearly 12% of people over 80 years old [8].

AMD is a medical condition that may result in blurred or no vision in the center of the visual field [2]. Early on there are often no symptoms. Over time, however, some people experience a gradual worsening of vision that may affect one or both eyes. While it does not result in complete blindness, loss of central vision can make it hard to recognize faces, drive, read, or perform other activities of daily life. Visual hallucinations may also occur but these do not represent a mental illness. The diagnosis is by a complete eye exam. The severity is divided into early, intermediate, and late types. The late type is additionally divided into ”dry” and ”wet” forms. The ”dry” form tends to progress more slowly than the ”wet” type. The ”wet” form is dangerous and accounts for approximately 90% of all cases of severe vision loss from the disease. During the disease, abnormal blood vessels under the retina begin to grow toward the macula. Because these new blood vessels are abnormal, they tend to break, bleed, and leak fluid, damaging the macula and causing it to lift up and pull away from its base. This can result in a rapid and severe loss of central vision. One of the promising surgery is the inject drug directly to the subretinal area which is named subretinal injection. The subretinal injection is a typical vitreoretinal surgery that has been successfully used in clinical trials to deliver therapeutic cargos of proteins, viral agents, and cells to the interphotoreceptor or subretinal compartment that has direct exposure to photoreceptors and the Retinal Pigment Epithelium (RPE) [9, 10].

The surgery can be divided into different categories based on the location of the diseases. Vitreoretinal surgery refers to an operation to treat eye problems involving the retina, macula, and vitreous fluid which mainly happens in the posterior segment of the eye, as shown in Fig. 1.1 (a). The RPE layer can be seen by zooming up the retina, as shown in Fig. 1.1 (b). Ideally, the drug is delivered into the potential space between the retinal pigment epithelium layer and photoreceptors in the outer nuclear layer [11]. The retina thickness is typically around 250 µm which proposes a big challenge for accurate injection depth control.

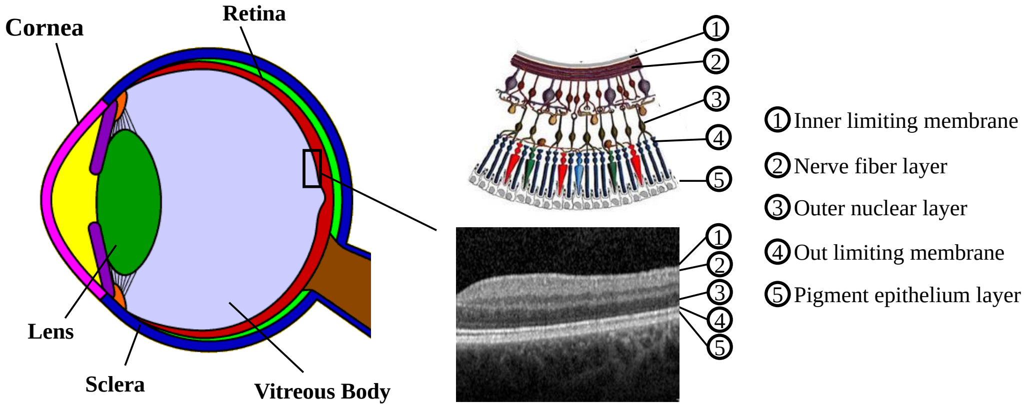


Figure 1.1: The structure of the human eye.

Subretinal injection is a form of vitreoretinal surgery. In this operation, conventionally, the surgeon is required to inject a microcannula into a specific area of the translucent retina to a certain depth. This area is a target where is normally defined pre-operatively by the ophthalmologist (see Fig. 1.2). The progress can be concluded to three steps for tool manipulation: approach, insertion, and retinal surgery. In other words, a tool needs to be positioned at the scleral incision, inserted into the eye, and manipulated around the incision point [12]. Three incisions which created by keratome always 3.5 mm away from the limbus and circled [13], are usually made in the sclera to provide an entrance for three tools at most: light source, surgery tool, and irrigation cannula [1, 14]. The irrigation cannula is always used in vitrectomy which injects liquid to maintain intraocular pressure (IOP). If IOP exceeds vascular occlusion pressure, it will cause temporary cessation of vascular perfusion, while low IOP will cause temporary meiosis and striate keratopathy, which greatly limit surgical visibility [15]. The light source is used to light up the intended area on the retina and so that the top view of the area can be obtained and analyzed through the microscope. The surgical tools include picks, micro tweezers, vitrectomy cutters, and other laser ablation devices that vary depending on the requirements of the procedure. Since the visual area is limited, the surgeon also needs to tilt the eye under the microscope to view more on the retinal. In other circumstances, the surgeon also needs to roll or tilt the light guide or the surgical tool to relocate or reposition the target area. The tools typically pivot at the sclera insertion point which is called the remote center of motion (RCM) in robotics, which is devised by Taylor et al [16], to reduce the damage to the tissue [17].

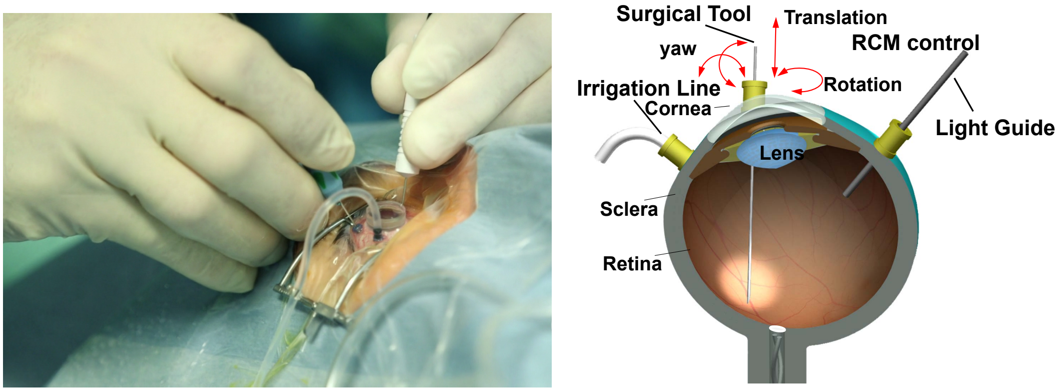


Figure 1.2: A conventional subretinal injection setup.

1.2 The current challenges

In this part, we are discussing some challenges that the surgeons are facing when they performing the vitreoretinal surgery manually.

1.2.1 Hand tremor

Even though some kinds of the ophthalmic surgery are very promising to treat some ophthalmic diseases, for example, to insert a needle into a vein on the surface of the retina to deliver some drug [18], it is unreliable or even impossible to perform clinically nowadays for their very delicate and sophisticated requirements. In other words, the challenges that the surgeon can not handle are high dexterity, long duration, and high precision motion [19]. With these problems, surgeon always with a great deal of experience but cannot operate because of physical limitations. The manual stability including tremor and dexterity is always influenced by fatigue, alcohol, and caffeine even for young microsurgeon [20, 21, 22]. So, it is necessary to govern their sleep, alcohol and caffeine consumption, and surgeons with anxiety or stress-related tremor may need to take some beta blockers to obtain a better surgical performance [23]. The research also shows that some suitable surgical postures and devices for resting hands and wrists are also helpful for improving hand steadiness [24]. Even though with these methods or strategies described above, the manual surgery precision still far from meets the requirement. The precision requirement of ophthalmic surgery varies depending on the specific operation. For the ILM peeling process, the surgeon can do it quite well, which means the 182 µm accuracy (the hand tremor RMS amplitude [25]) is enough. For the sub-retinal injection, the average thickness of retina is around 200 µm, therefore, 20 µm would be an acceptable position accuracy. For the retinal vein cannulation, the ideal position accuracy would be 20 µm, since the diameter of branch retinal veins is typically less than 200 µm. Therefore, the typical position accuracy for robotic eye surgery is considered to be roughly around 10 µm [26, 27, 28] to meet all potential operations.

1.2.2 Limited sensing

The microscope mounted above the eye of the patient can only provide the top view of the afflicted region. This can not provide enough information for the surgeon during the operation. Compared to other medical image technologies for example magnetic resonance imaging (MRI) or computed tomography (CT), OCT can achieve an axial resolution at 5-10 µm [29] which meet the requirement of vitreoretinal surgery and provide the cross-sectional image in vivo at the same time. Since 1996, after the first commercially OCT machine was marketed by Carl Zeiss, the

OCT technology has been quite matured in ocular disease treatment especially the diagnosis and evaluation progress [3]. In 2014, Carl Zeiss launched its Rescan 700 OCT systems which is a real time intraoperative Spectral Domain OCT (i-SD-OCT). This new product allows microsurgeon to see the surgical field in both a planar view and a cross-sectional view simultaneous in real time without pausing the surgery to look away from the surgery. The Rescan system from Zeiss is not approved for sale in some countries and still under tested in real clinical operation. Justis P. Ehlers from Cole Eye Institute of Cleveland Clinic used the Rescan system under a research protocol [30]. M. Pfau tested the system in 40 consecutive cases which concluded that i-OCT has the potential to improve the quality of posterior and segment surgery [31]. Besides the limited sensing from the imaging part which is developing fast for clinical level, the force feedback information exerted by microsurgical tools is another important factor [32]. The benefit from force feedback is that the interaction forces are reduced which contributes to the improvement of patient safety and decreases problems of tissue trauma [33]. The experiments showed that 75% of the forces imposed by the tissue on the instrument tip are less than 7.5 mN during retinal surgery which is below the surgeons' tactile sensitivity [34]. So it can be concluded that in the traditional ophthalmic surgery, the surgeons operate mainly using visual feedback with a little or without force interactions between retinal tissue and the surgical tool [34]. This impedes the improvement of ophthalmic surgical quality. It is also a challenge to integrate the high precision and sensitive force sensor into the tiny manipulator and make it high reliability and robustness in clinical operation.

1.2.3 Virtual fixture RCM

Virtual fixture including RCM control is an important technology to ensure the safety of surgery, improve the quality of minimally invasive surgery (MIS) and reduce the healing process after the surgery. During the surgery, the surgeon needs to change the virtual fixtures. The flexibility of virtual fixtures plays an important role in vitreoretinal surgery. Dewan concluded that during the vitreoretinal surgery, there exist five motion styles including free motion, surface following, tool alignment, targeting and insertion/extraction [35]. Besides the free motion, all other motions need constraints. Based on the observation of microscope images of surgeries performed by expert surgeons revealed that the position accuracy of the RCM should be at less than 3mm [1] to reduce the damage to the tissue. In the description of unaided surgery progress, it is a challenging task for the surgeon to maintain RCM control with delicate motion which needs a lot of experience. Since the robotic system also can perform the same flexible and delicate level as a human hand, it will be more effective and easier to realize this operative progress.

1.2.4 No available commercial solving solution

The famous robot surgical system called Da Vinci has been widely commercialized for some kinds of surgery like colorectal surgery and gastrointestinal surgery [36, 37]. There are also some attempts in the literature to perform ophthalmic surgery. In 2007, Tsirbas performed the corneal laceration in a porcine model using Da Vinci surgical robot and results showed that the robotic ocular microsurgery is technically feasible, but the high cost and how to integrate with the standard ophthalmic microsurgical instruments are limitations. [38] In 2008, J. P. Hubschman reported that the da Vinci Surgical System in the current design has two limitations: the inconvenient control for intraocular surgery and the inferior endoscope-acquired imaging system [39]. In 2015, Beheshti M concluded that da Vinci Surgical System was too bulky, and lacks the precision and dexterity required for delicate applications that require higher accuracy in ophthalmic surgery [40].

To overcome surgeon's hand tremor and to achieve dexterous motion and precise RCM control, many researchers in recent years have introduced robotic setups with high precision in different scales and design mechanisms [1, 14, 26, 27, 28, 41, 42, 43]. These robots can be classified into four main categories: 1) Hand-held surgical instruments, which have the benefits of tremor suppression and intuitive operation [44, 45]. 2) Cooperatively controlled systems, which focus on tremor filtering with a stable robotic arm, normally equipped with force sensors. Cooperatively controlled systems are offering intuitive operation because the surgical tool is held by the robot and surgeon's hand simultaneously. However, they lack motion scaling and are unable to execute motion profiles. They also introduce inertial and frictional forces that may limit their applications in dynamic tasks. 3) Teleoperation systems, which are often embodied within a console-based setting to provide tremor filtering and motion scaling [46]. These systems are known to be the most successful in commercial products e.g. da Vinci surgical robot (Intuitive Surgical Inc.). However, robots designed for general surgery apart from their significant footprints, do not have sufficient precision for retinal surgery [47]. 4) Magnetically controlled micro-robots, which could provide an alternative surgical approach but current systems are in a very early stage and lack most of the benefits of robot assistance [48], e.g. safety consideration and haptic feedback [49, 50]. We listed the outline with some typical robots for eye surgery shown as in Fig. 1.3.

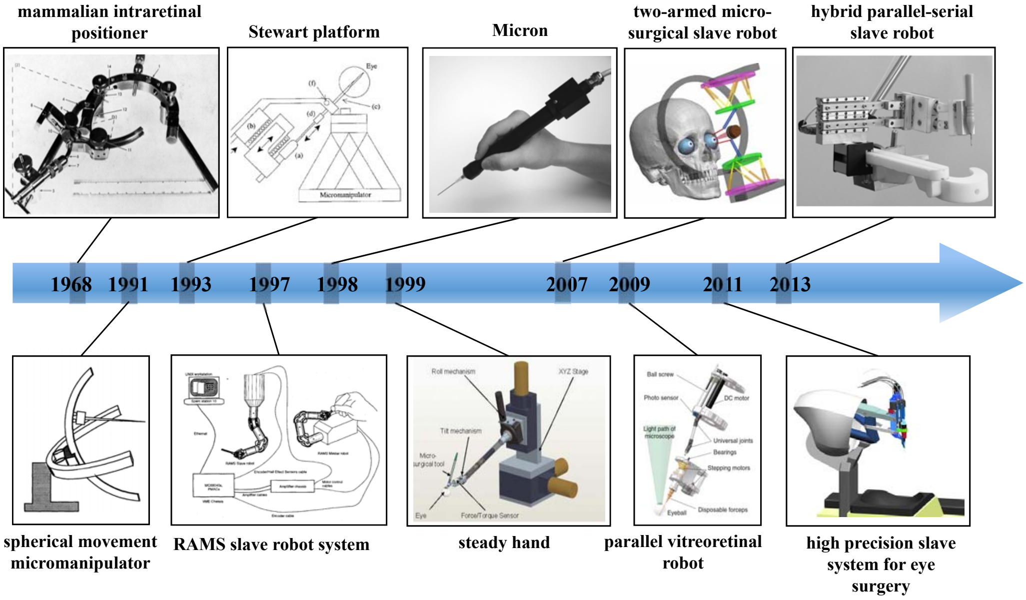


Figure 1.3: Some typical robots for eye surgery.

Over the years, ophthalmic robots are becoming mature to enter into the clinical trial. The Robotic Retinal Dissection Device (R2D2) with teleoperation functionality performed the world's first robotic subretinal injection in 2016 [48, 51] proving the feasibility and safety concept of the robot-assisted eye surgery. Different from the R2D2, in this thesis, we proposed a hybrid parallel-serial surgical robot designed for subretinal injection. The main motivation of using parallel-serial mechanism is that we are able to avoid using active rotational joints and consequently to reduce the robot dimensions significantly. With this system, besides addressing the challenges from the surgeon's hand control ability, we are also aiming to improve visual feedback during the injection.

Since the retina is a transparent tissue with limited illumination from conventional intraocular light sources, it is difficult for the surgeon to judge the insertion depth when the needle tip is under the retinal surface only from the traditional microscopic view. The needle tip position is also hard to estimate from the robot since the kinematics of the system is not suitable to accurately estimate the needle tip position due to a thin, long, and flexible needle body [52] with heterogeneous deformations. Therefore, additional imaging modality is necessary. 3D ultrasound and MRI imaging, which are typically used in non-microsurgical scenarios, do not have sufficient precision for subretinal interventions. These imaging modalities normally offer needle localization error of around 500 µm [53], while the retina with an average thickness of around 250 µm [54] requires operational accuracy with a maximum error of 25 µm.

A promising ophthalmic imaging modality that became well-known in recent years is optical coherence tomography (OCT). Other 3D medical imaging technologies, including computed tomography (CT) scans, fluoroscopy, magnetic resonance (MR) and ultrasound technology are already applied in the cardiac, brain, and thoracic surgeries, not only for diagnostic procedures but also as a real time surgical guidance [55, 56, 57]. However, these imaging technologies can hardly achieve the ideal resolution for ophthalmic surgery applications. For MRI-guided interventions with resolution in millimeters in breast and prostate biopsies, 18 gauge needle with a diameter of 1.27 *mm* is used, while for ophthalmic surgery, 30 gauge needle, which has the diameter of 0.31 *mm*, requires resolution of submillimeter [58]. As shown in Fig. 1.4, we can see the resolution distribution of different 3D image modalities.

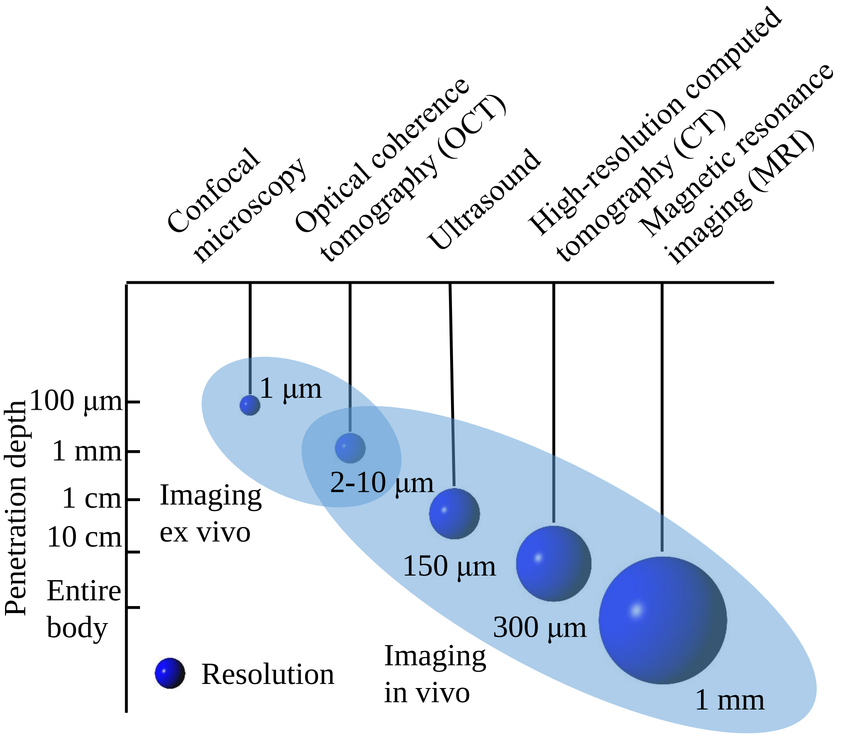


Figure 1.4: The resolution distribution of different 3D image modalities.

OCT is originally used for ophthalmic diagnosis and it is now modified for intraoperative procedures by offering a suitable resolution with non-invasive radiation that brings the minimum risk of toxication for the ocular tissue. The latest OCT machine has been developed to even give real-time cross-sectional images of the microstructural anatomies with the potential of observing information about the interactions between the surgical instrument and intraocular tissues [30]. The OCT Scanning and scanner coordinate system schematic is shown in Fig. 1.5. Fig. 1.5(a) shows 1D acquisition (A-scan). A single depth profile is acquired which measures backscattered intensity vs. axial dimension (depth). Fig. 1.5(b) shows 2D imaging (B-scan). The OCT beam is scanned in a transverse direction while A-scans (red arrows) are acquired. Fig. 1.5(c) shows 3D acquisition. Multiple B-Scans are acquired such that A-scans are sampled on a 2D grid in the transverse plane. Fig. 1.5(d) shows the one example of relationship for three types of scan on the retina, where C-scan refer to the scan of volumetric image. [59] The more detailed of OCT imaging principle can be referred [29].

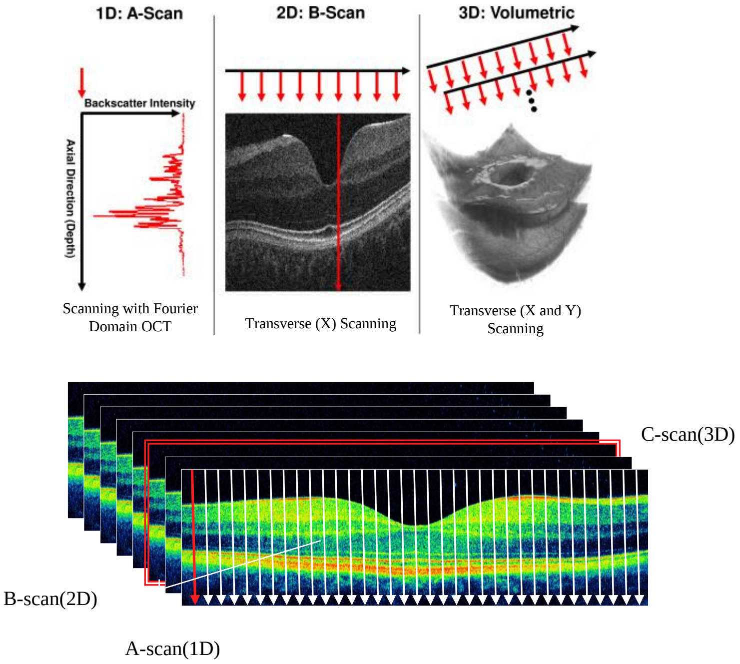


Figure 1.5: OCT Scanning and scanner coordinate system schematic.

1.3 Motivation

The image-guided robotic-assisted technology has been introduced recently to help the surgeon to increase and ensure the quality of surgery intraoperatively. The Da Vinci surgical robot introduced the stereo vision system to enhance the surgeon's perception during laparoscopic surgery. MRI/CT imaging-guided robotic system is also introduced to brain surgery. Ultrasound/CT imaging-guided robotic system is also applied in the Spinal puncture. VR surgery is typical microsurgery which microscope is essential for the traditional operation. However, the normal microscope image cannot provide strong enough information regarding the contact between the tissue and instrument. Hereby, we introduce the OCT volumetric images to navigate the needle performing the insertion. To ensure these procedures, we have several key points which need to be considered.

1.3.1 RCM performance

The RCM performance is the fundamental requirement to ensure surgical safety. The designed robot have to constrain the RCM fixture within a tolerance. This proposes that hardware and software should be co-designed in an accurate and timely manner.

1.3.2 Needle tip&pose estimation

The question of where the needle tip and how to estimation the pose of the needle is the most critical for the subretinal injection. Afterward, we could better understand how is the relationship between the needle and retina. Based on the observation of the OCT B-scan image of needle and retina, we can find that when the needle is above the retina, there is obvious geometrical information to difference the needle and retina. However, when the needle goes below the retina, it is hard to segment the needle from the retina since the needle is inside the retina and the intensity of them is similar.

Since the autonomous system is more and more popular in the surgical robot. Especially when we consider to totally free the surgeon's hand from the high precision and dexterity motion, making surgeons focusing more on the treatment plan and decision. These challenges are the prior problems we need to figure out when we want to realize the image-guided robotic-assisted vitreoretinal surgery and move forward to safety critical autonomous system.

1.4 Contribution

This thesis provides partial answers to the above-listed questions. The contributions of this thesis are to provide the robot and try to tackle the needle localization problem thus to path the way to a certain degree of autonomy for the subretinal injection. This thesis is subdivided into 5 main chapters to present those contributions as well as shown in Fig. 1.6.

1. Chapter 2 presents a robot system for subretinal injection integrated with intraoperative optical coherence tomography (OCT). The surgical workflow using this system consists of

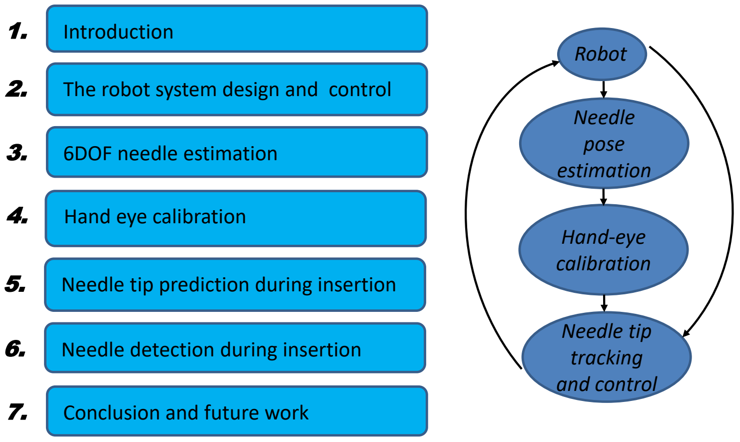


Figure 1.6: The overall structure of the thesis.

two main parts: the first part is the manual robot control part which is to aim the target before approaching the retinal surface while considering remote center of motion (RCM) constraint. When the injection area is precisely located, the second part which is needle injection mode will be activated. To ensure surgical safety, needle insertion depth will be estimated continuously from OCT images. A soft RCM control method is designed and integrated for the controller of our hybrid parallel-serial surgical robot. Safety and accuracy performance evaluation with 15 ms control loop shows RCM deviation error is within 1 mm. Experimental results demonstrated that the proposed system has the ability to improve surgical outcomes by helping the surgeon to overcome their physical limitation for enabling a better dexterous motion and furthermore enhancing their visual feedback for a better intraocular perception.

2. Chapter 3 presents a novel method to estimate the 6DOF needle pose specifically for the application of robotic intraocular needle navigation using OCT volumes. The key ingredients of the proposed method are (a) 3D needle point cloud segmentation in OCT volume and (b) needle point cloud 6DOF pose estimation using a modified iterative closest point (ICP) algorithm. To address the former, a voting mechanism with geometric features of the needle is utilized to robustly segment the needle in OCT volume. Afterward, the CAD model of the needle point cloud is matched with the segmented needle point cloud to estimate the 6DOF needle pose with a proposed shift-rotate ICP (SR-ICP). This method is evaluated by the existing ophthalmic robot on ex-vivo pig eyes. Quantitative and qualitative results are evaluated and presented for the proposed method.

3. Chapter 4 presents a flexible framework for hand-eye calibration of an ophthalmic robot with a microscope-integrated optical coherence tomography (MI-OCT) without any markers. The proposed method consists of three main steps: a) we estimate the OCT calibration parameters; b) with micro-scale displacements controlled by the robot, we detect and segment the needle tip in 3D-OCT volume; c) we find the transformation between the coordinate system of the OCT camera and the coordinate system of the robot. We verified the capability of our framework in ex-vivo pig eye experiments and compared the results with a reference method (marker-based). In all experiments, our method showed a small difference from the marker based method, with a mean calibration error of 9.2 µm and 7.0 µm, respectively. Additionally, the noise test shows the robustness of the proposed method.

4. Chapter 5 presents a novel approach to estimate the 3D position of the needle under the retina using the information from MI-OCT. We evaluated our approach on both phantom tissues and ex-vivo porcine eyes. Evaluation results show that the average error in distance measurement is 4.7 µm (maximum of 16.5 µm). We furthermore, verified the feasibility of the proposed method to track the insertion depth of needle in robot-assisted subretinal injection.

5. Chapter 6 demonstrates a robust framework for needle detection and localization in subretinal injection using MI-OCT based on deep learning. The proposed method consists of two main steps: a) the preprocessing of OCT volumetric images; b) needle localization in the processed images. The first step is to coarsely localize the needle position based on the needle information above the retinal surface and crop the original image into a small region of interest (ROI). Afterward, the cropped small image is fed into a well trained network for detection and localization of the needle segment. The entire framework is extensively validated in ex-vivo pig eye experiments with robotic subretinal injection. The results show that the proposed method can localize the needle accurately with the confidence of 99.2%.

6. Chapter 7 summarizes this thesis and discuss future research directions based on this thesis.

Chapter 2

Robot Design and RCM control

2.1 Introduction

In this chapter, a hybrid parallel-serial robot is introduced to assist ophthalmologists for performing the subretinal injection with enhanced precision and RCM control for further integration with intraoperative OCT imaging and proper visualization. Currently, to overcome the surgeon's hand tremor and to achieve dexterous motion and precise RCM control, in recent years many researchers have introduced robotic setups with high precision in different scales and design mechanisms [1, 14, 26, 27, 28, 41, 42, 43]. These robots can be classified into four main categories: 1) Hand-held surgical instruments, which have the benefits of tremor suppression and intuitive operation [44, 45]. 2) Cooperatively controlled systems [41], which focus on tremor filtering with a stable robotic arm, normally equipped with force sensors. Cooperatively controlled systems can be operated intuitively, due to the fact that the surgical tool is held by the robot and the surgeon's hand simultaneously. The use of force sensors in the system allows to have a motion "scaling" in the admittance control scheme. However, they introduce inertial and frictional forces that may limit their application in dynamic tasks. 3) Teleoperation systems, which are often embodied within a console-based setting to provide tremor filtering and motion scaling. These systems are known to be the most successful in the commercial products e.g. the da Vinci surgical robot (Intuitive Surgical Inc.). However, quite apart from their significant footprints, robots designed for general surgery do not have sufficient precision for retinal surgery [47]. 4) Magnetically controlled micro-robots could provide an alternative surgical approach, but their current systems are in a very early stage and lack most of the benefits of robot assistance [48], e.g. safety consideration and haptic feedback [49, 50]. We proposed a robot with two parallel units and one prismatic joint that unlike the mechanical hardware-based RCM mechanisms is able to realize precise and robust software-based RCM control. The benefit of our proposed method is that with such a mechanism a very compact robot is designed (the current functional prototype weights 315 grams and its volume is equivalent to the size of an average human hand) [46]. This would be easier for being integrated to the conventional ophthalmic operation room where already occupied by several machines. However, the main challenge from this configuration is that the RCM control has to be designed in the "soft" way which means that the robot controller should be able to synchronize all the joint positions in a proper and timely manner. Serious consideration is needed for safe and critical application scenario where the RCM point should be controlled within the clinical-grade tolerance so that the trauma on the sclera is minimized. To address this challenge, a lightweight RCM control algorithm was proposed with safety routine. The overall control loop of the system is within 15 ms and the RCM deviation is limited to 1 mm. Moreover, to resolve the unknown needle insertion depth during the injection, we intergraded the intraoperative OCT (from Lumera 700 Carl Zeiss microscope with Rescan 700 OCT engine) as additional feedback. The volumetric OCT images are captured and processed to track the needle tip position under the retina for providing needle insertion depth. 3D visualization is realized to offer enhanced visual feedback for the surgeon. Quantitative and qualitative experiments are performed to verify the feasibility and effectiveness of the system.

The rest of this chapter is organized as follows: Section 2.2 details the mechanical design of the robot, control as well as OCT integration. Section 2.3 presents the verification of the overall system including the RCM performance and needle insertion tracking with the robot. Section 2.4 concludes this chapter with the discussion and the presentation of future work.

2.2 Method and Materials

2.2.1 Robot Design

In order to design a compact and lightweight robot for intuitive integration to the ophthalmic clinical routine, we propose a platform based on the hybrid parallel-serial mechanism designed by prismatic and rotation joints and realized by using linear stick-slip piezo actuators. The driving force for choosing this mechanism is reducing backdriving effect caused by the gear or the tendon. This effect will be significantly reduced in our mechanism while the joint is directly driven by the motor. The detailed parallel mechanical unit is shown in Fig. 2.1. The piezo motor (SLC1750, SmarACT GmbH, Germany) M_1 and M_2 with the prismatic joints J_1 and J_2

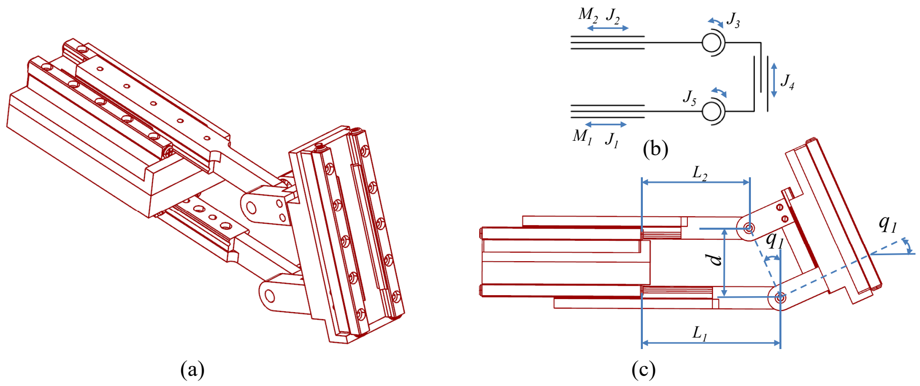


Figure 2.1: Parallel mechanical unit. (a) The oblique view of the unit. (b) The structure diagram of the unit. (c) The rotation effect with the front view of the unit with main dimensions.

are mounted in parallel. The rotation joints J_3, J_5 and prismatic joint J_4 are used to close the kinematics loop of the unit. The synergetic control of two prismatic joints can be transferred into a virtual prismatic q_0 and rotation joint q_1 which can be calculated as,

$$q_0 = L_1 \tag{2.1}$$

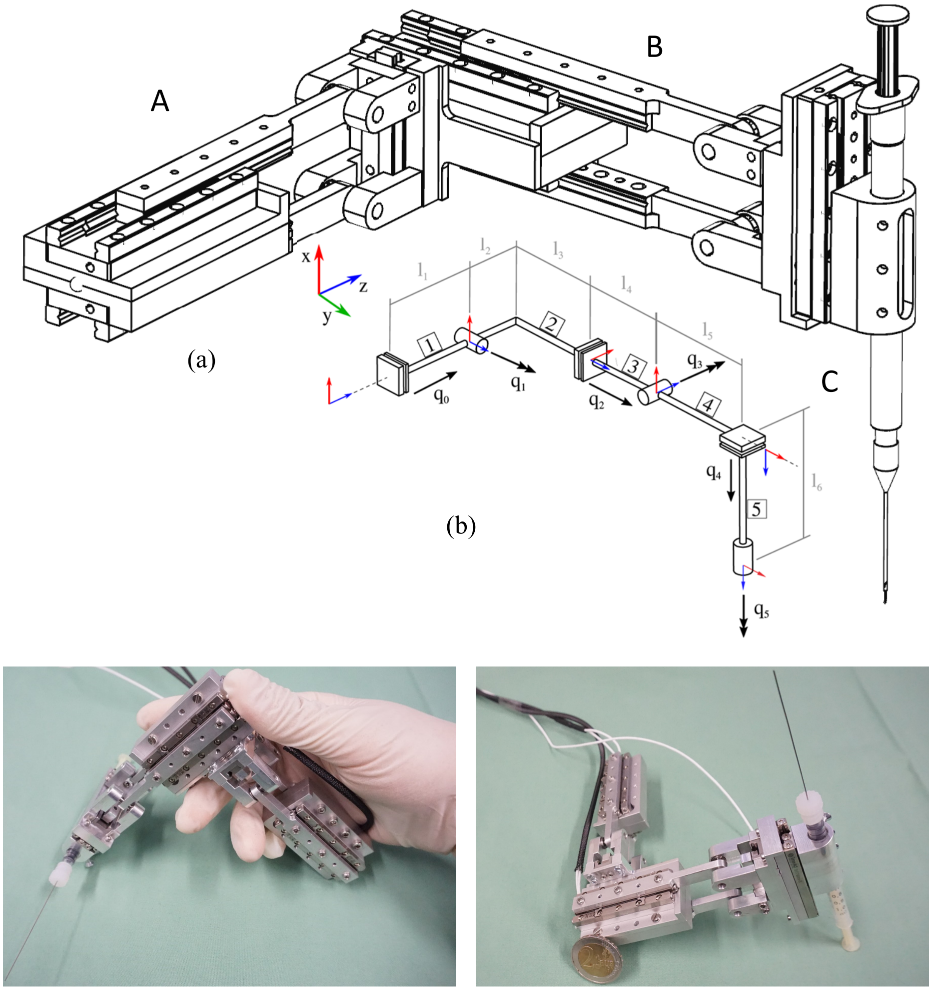
$$q_1 = atan(\frac{L_2 - L_1}{d}) \tag{2.2}$$

where L_1 is the position of joint J_1 from initial position, L_2 is the position of joint J_2 from initial position.

The robot consists of two parallel mechanical units and a needle slider, comprising a five degree-of-freedom(DoF) manipulator, shown in Fig. 2.2. Its Denavit-Hartenberg parameters are obtained via CAD model. A key feature is that its workspace is free of singularities as the mapping between joint vectors and needle poses is one-to-one. Due to this advantage and the plain forward kinematics, spatial coordinates can be stored as joint vectors for the purpose of a virtual fixture constraint.

2.2.2 Robot Control and RCM Constraint

The robot is designed to have three types of movement by a master control device. (a) The unconstrained movement: the robot is controlled by a master device without any motion constraint. In this control mode, the needle is placed outside of the eye, the robot is controlled to align the needle with the trocar to be introduced inside the eye. (b) The RCM control when needle is inside the eye: as soon as the needle tip is approaching the trocar entrance with the correct orientation, normally perpendicular to the scleral surface, the needle tip position



(c) (d)

Figure 2.2: The overall design of the robot. (a) The CAD model of the robot with two identical parallel mechanical units **A**, **B**. **C** is the final prismatic joint, used for decoupled injection. (b) The simplified kinematics structure of the robot. (c) and (d) The real entity of designed robot in different view.

is registered as the RCM point. The robot controller is switched to the RCM mode to reduce the incision point trauma. (c) The insertion mode for needle injection: this mode is enabled when the needle approaches very close to the retina before puncturing the tissue. During the injection, only the last degree of freedom, slider q_5, is enabled.

The robot movement type (a) and (c) can be directly executed using PID controller of the master interface. Here we mainly introduce the RCM control algorithm design for our specific designed robot. To achieve RCM motion, the joint parameters which are sent to the robot must

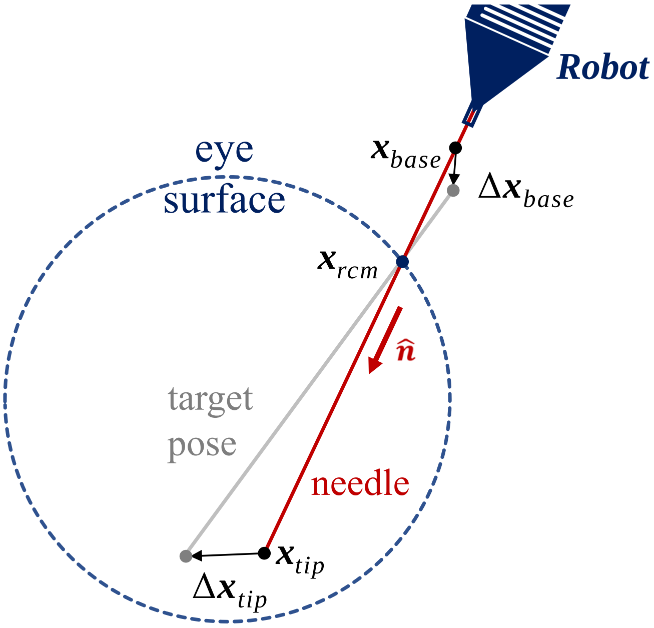


Figure 2.3: Schematic showing change of a straight needle pose as a result of the tip displacement in RCM mode.

be such that the end-effector (i.e. the needle) moves through a point in space with minimum deviation. If the operative part of the needle is represented by a unit vector $\hat{\boldsymbol{n}}$ connecting the needle tip $\boldsymbol{x}_{tip}$ to a needle "base" $\boldsymbol{x}_{base}$ (a point remaining outside the eye), the constraint is expressed as,

$$\Delta\boldsymbol{x}_{tip} \cdot \hat{\boldsymbol{n}} = \Delta\boldsymbol{x}_{base} \cdot \hat{\boldsymbol{n}} \tag{2.3}$$

$$\Delta x_{base} \times (\boldsymbol{x}_{rcm} - \boldsymbol{x}_{tip}) = \Delta x_{tip} \times (\boldsymbol{x}_{rcm} - \boldsymbol{x}_{base}) \tag{2.4}$$

where Δx_{base} and Δx_{tip} are the movement for base and tip point in a small time interval, Δx_{tip} is obtained from master control signal, $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_{tip} \\ \boldsymbol{x}_{base} \end{bmatrix}$ is the current position the base and needle tip point which can be calculated from the forward kinematics of the robot, $\boldsymbol{x}_{rcm}$ is the registered RCM point in space, as shown in Fig. 2.3. The equations state that the tip and the base velocity components along the needle must be equal (conservation of needle length), whereas those perpendicular to the needle are coplanar, in opposite directions, and with magnitudes proportional to the distance from the RCM point to the base point, and the RCM point to the tip point, respectively. Integrating these requirements when solving the inverse kinematics can be used to control the robot with a two-point kinematic scheme. The needle state is taken as a vector comprising the tip and the base positions, and the transnational part of the respective Jacobians is used to iteratively solve

$$\boldsymbol{J} \cdot \Delta \boldsymbol{q} = \begin{bmatrix} \Delta \boldsymbol{x}_{tip} \\ \Delta \boldsymbol{x}_{base} \end{bmatrix} = \boldsymbol{x}^{targ} - \boldsymbol{x} \tag{2.5}$$

for updating the joints difference $\Delta \boldsymbol{q}$, where $\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J_{tip}}(\boldsymbol{q}) \\ \boldsymbol{J_{base}}(\boldsymbol{q}) \end{bmatrix}$, $\boldsymbol{J_{tip}}(\boldsymbol{q})$ and $\boldsymbol{J_{base}}(\boldsymbol{q})$ are the Jacobians for $\boldsymbol{x}_{tip}$ and $\boldsymbol{x}_{base}$ respectively. $\boldsymbol{x}_{base}$ can be calculated based on $\boldsymbol{x}_{tip}$ and $\boldsymbol{x}_{rcm}$ with assumption of straight needle as follows,

$$\lambda(\boldsymbol{x}_{tip} - \boldsymbol{x}_{base}) = \boldsymbol{x}_{rcm} - \boldsymbol{x}_{tip} \tag{2.6}$$

where λ is the ratio factor which starts from 1 when the needle tip enter into the trocar. Afterwards, λ will change base on the movement of the needle tip and will be updated with a difference $\Delta\lambda$,

$$\Delta\lambda = \Delta \boldsymbol{x}_{tip} \cdot \widehat{\boldsymbol{n}} \tag{2.7}$$

A target state based on desired tip displacement can be constructed,

$$\boldsymbol{x}^{targ} = \begin{bmatrix} \boldsymbol{x}_{tip} + \Delta \boldsymbol{x}_{tip} \\ M(\boldsymbol{x}_{tip} + \Delta \boldsymbol{x}_{tip}, \boldsymbol{x}_{rcm}, \lambda) \end{bmatrix} \tag{2.8}$$

where $M(\boldsymbol{x}_{tip}, \boldsymbol{x}_{rcm}, \lambda)$ is a function refining from Eqn. 2.5 for calculating $\boldsymbol{x}_{base}$. A RCM deviation error, shown in Fig. 2.4, can also be derived by using the cross-product-area identity as,

$$e_{rcm} = \frac{1}{\|\widehat{\boldsymbol{n}}\|} \cdot \|(\boldsymbol{x}_{rcm} - \boldsymbol{x}_{tip}) \times (\boldsymbol{x}_{rcm} - \boldsymbol{x}_{base})\| \tag{2.9}$$

which is to be used for step-size control.

Employing forward kinematics to evaluate (2.12) between steps to ensure that the RCM control error is within an algorithm control tolerance ε_1 and ε_2 for needle tip and RCM point positions. Combining the aforementioned equations, Algorithm 1 shown below recursively updates joint target positions with given tip displacements. The tolerance ε_1 and ε_2 are depended on actuator precision and are a trade-off between the displacement of needle tip in small time interval and quality of virtual-fixture adherence. The algorithm will first check whether the target position can be reachable by the movement of joints, then a check on mid-step RCM deviation is done, resulting in recursively decreasing the step size or executing the update if e_{rcm} is within tolerance.

2.2.3 Master and Slave Control Design

The control signal from the master controller can be treated as $\boldsymbol{m} = [m_1, m_2, m_3, m_4, m_5]$, where m_1, m_2, and m_3 are the translation speed input, m_4 and m_5 are the rotation speed input. We

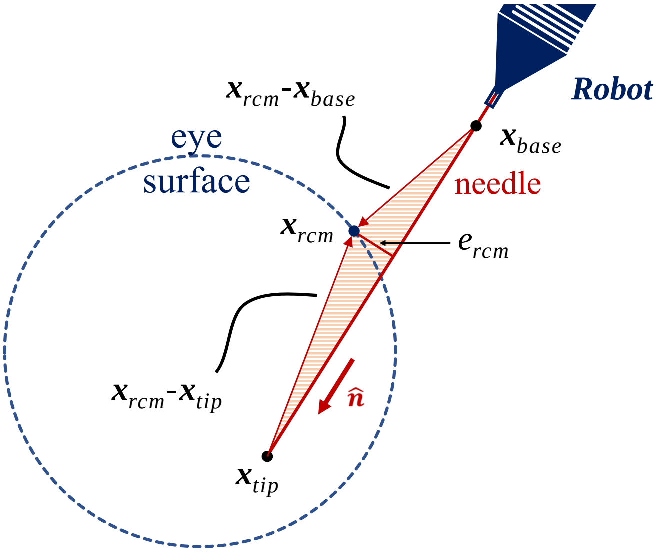


Figure 2.4: Forward kinematics and needle detection are used to obtain an estimation for the RCM control error (exaggerated for visual purposes here), proportional to the ratio between the highlighted area and the length of the needle's operative segment.

Algorithm 1 RCM Control Algorithm

INPUT1: $\Delta \boldsymbol{x}_{tip}$ - displacement of needle tip
INPUT2: $\boldsymbol{x}_{rcm}$ - fixed position of RCM point
OUTPUT: $\Delta \boldsymbol{q}$ - movement for each joint

```
function RCMSTEP($\Delta x_{tip}$, $x_{rcm}$)
    $\boldsymbol{q} \leftarrow getJoints()$, $\Delta\boldsymbol{q} \leftarrow \boldsymbol{0}$
    $\boldsymbol{x} \leftarrow fwdKinematics(\boldsymbol{q})$
    $\boldsymbol{x}^{targ} \leftarrow calcTarget(\boldsymbol{x}, \Delta\boldsymbol{x}_{tip}, \boldsymbol{x}_{rcm})$
    while $true$ do
        $\boldsymbol{J} \leftarrow calcJacobian(\boldsymbol{q})$
        $\Delta\boldsymbol{q} \leftarrow \Delta\boldsymbol{q} + \boldsymbol{J}^{\dagger} \cdot (\boldsymbol{x}^{targ} - \boldsymbol{x})$
        $\boldsymbol{x} \leftarrow fwdKinematics(\boldsymbol{q} + \Delta\boldsymbol{q})$
        if $||\boldsymbol{x}^{targ} - \boldsymbol{x}|| <= \varepsilon_1$ then
            if $e_{rcm} > \varepsilon_2$ then
                $\Delta\boldsymbol{q} \leftarrow$ RCMSTEP $(\Delta\boldsymbol{x}_{tip}/2, \boldsymbol{x}_{rcm})$
            else
                break
            end if
        else
            $\Delta\boldsymbol{q} \leftarrow \boldsymbol{0}$, break
        end if
    end while
    return $\Delta\boldsymbol{q}$
end function
```

use the joystick as the master controller. Due to the fact that there are some unintentional movements randomly in other directions while controlling the motion by the joystick, the input signal is taken from the maximum value in $\boldsymbol{m}$. The speed and acceleration for the translation and rotation of the needle need to be constrained within the range of motor dynamic character, shown as follows,

$$s = \begin{cases} max(\boldsymbol{m}), \text{if } max(\boldsymbol{m}) < \boldsymbol{V_{max}} \\ \boldsymbol{V_{max}}, \text{if } max(\boldsymbol{m}) \geq \boldsymbol{V_{max}} \end{cases} \tag{2.10}$$

$$s = \begin{cases} s, \text{if } s - s' < \Delta t \ \boldsymbol{A}_{max} \\ s' + \Delta t \boldsymbol{A}_{max}, \text{if } s - s' \geq \Delta t \ \boldsymbol{A}_{max} \end{cases} \tag{2.11}$$

where s is the input for slave robot, s' is the input for slave robot in the previous step, $\boldsymbol{V}_{max}$ and $\boldsymbol{A}_{max}$ is the maximum translation and rotation speed and acceleration constraints for the needle mounted on the robot when the RCM is set on the needle tip, Δt is the overhead for one control loop.

The intended difference of needle pose can be calculated as $\Delta t\boldsymbol{m}$. It should be considered that different instruments with different loads may influence the motor dynamic character. Moreover, the different RCM point location on the needle with the same rotational speed will lead to different speed for the piezo motor. In order to ensure that each of the motor joints are running on the safety critical mode within the fault tolerance, the position of joints obtained from the optical encoder sensors inside the piezo motor are used to check the deviation of the RCM. This deviation is calculated as,

$$d_{rcm} = |\boldsymbol{x}_{rcm} - \boldsymbol{p}_{rcm}| \tag{2.12}$$

where p_{rcm} is the RCM position calculated from the joint position with forward kinematics. The $\boldsymbol{V}_{max}$ will be reset online if the deviation exceeds the potential tolerance.

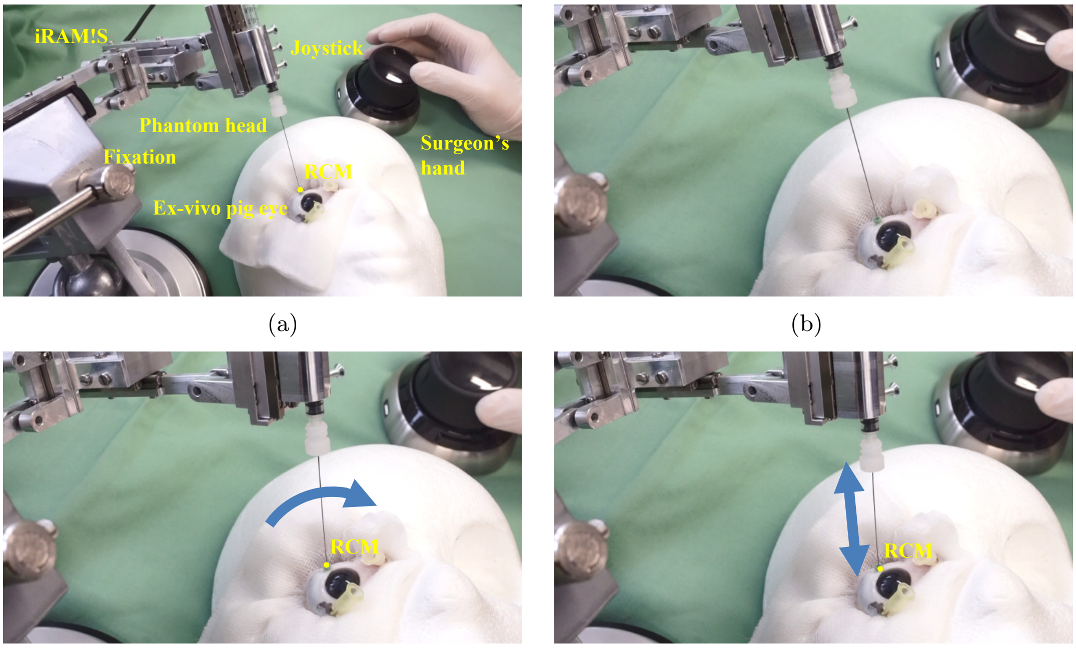
$$\boldsymbol{V}_{max} = \begin{cases} \boldsymbol{V}_{max}, \text{if } d_{rcm} < \kappa\varepsilon \\ max(\boldsymbol{m}), \text{if } d_{rcm} \geq \kappa\varepsilon \end{cases} \tag{2.13}$$

where $\kappa \in [0,1]$ is the security coefficient, ε is the tolerance which is defined by clinic requirement.

2.3 Experiments and Results

We performed two sets of experiments with the proposed system. The first experiment is designed to verify the accuracy and safety performance of RCM, which is critical for any minimally invasive teleoperation system. The second experiment is the needle tracking during insertion when the needle is reconstructed and the needle tip is predicted under the retina to give the reference of the needle depth during the injection.

2.3.1 RCM control Performance Evaluation



(a) (b)

(c) (d)

Figure 2.5: The subretinal injection trail on the ex-vivo pig eye for busy mode control loop test. (a) The overall setup. The RCM control point will be registered with a button triggered on the joystick when the needle is inserted into the trocar. The trocar is placed on the sclera on the ex-vivo pig eye. (b) The needle is inserted into the ex-vivo eye. (c) The needle is pivoted around the insertion point to relocate the target area. (d) The final joint is only activated for high accuracy injection.

The performance of the RCM control is one of the most critical features to be evaluated for the system. The accuracy of the RCM control is directly related to the movement of each control loops. The control loop is decided by two factors, one is the control loop time and the other one is the motor speed. The control loop time is critical since it will influence the position accuracy of the robot and also adds the extra delay of human hand-eye reaction loop. The low interface of PID speed&postion servo control is running on the firmware of FPGA board with 10KHz frequency. The joystick is connected to a host computer with 2.8GHz i7-7700HQ processor and 16GB memory. To ensure real-time property of the proposed system, we implemented the control framework in C++ in Ubuntu 16.04 with real-time patches. The host computer communicates with FPGA low control board in Ethernet connection to update the joint position and send speed&position command. In order to test the control loop time of the designed system, we performed the experiments in a time period of 750s with 50000 control

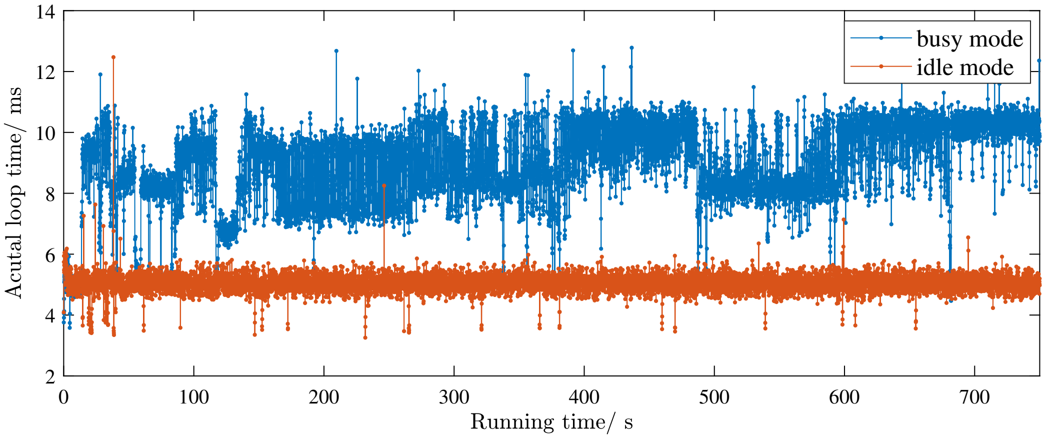
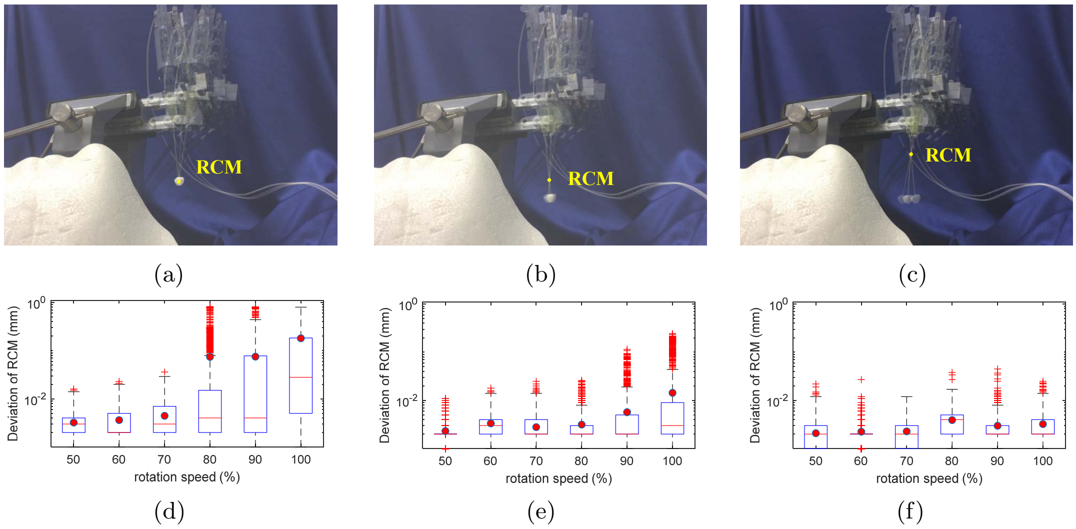


Figure 2.6: The actual loop time with two modes test under 750s.

loops under two condition: 1) idle mode which there is no input signal from the joystick; and 2) busy mode with continuous input signal from the joystick. The busy modes simulates the subretinal injection on the ex-vivo pig eye shown as Fig. 2.5. Fig. 2.6 shows that the idle mode control loop time is less than busy loop time because during the idle mode no signal is updated and the motor control function is not called. The time for both of the two control loop modes is less than 14 ms which means we can fix the control loop time as 15 ms. This is significantly less than the surgeon's hand-eye reaction time which is around 200 ms [60].

Since each part of the robot is manufactured by precision Computer Numerical Control (CNC) machine, the manufacturing accuracy of the parts can be considered within 5 µm. Therefore, we could infer the accuracy of RCM by calculating the forward kinematics with the joint positions of the robot. The optical encoder integrated into our piezo motors has a resolution of 0.05 µm. The algorithm control tolerance ε_1 and ε_2 are set to 1 µm to ensure the accuracy from algorithm calculation aspect. We tested various RCM positions along the needle with different λ to see the correlation of the motor speed and the RCM control deviation. The maximum rotation speed for the needle in RCM control is set as 10° per second. The RCM control error was set as 3 mm with the suggestion from the medical doctor in previous reference [61]. However, due to the fact that the needle and the trocar diameter are continuously decreasing for minimally invasive surgery, we set the error tolerance ε in Eqn. 2.13 as 1 mm and the safety coefficient κ as 0.8. Fig. 2.7 shows the time-lapse photography for RCM control with different λ and whisker plots for RCM deviation using the proposed control method. The needle is pivoted around the RCM with 30°. Normally, a larger rotation speed under the same



(a) (b) (c)

(d) (e) (f)

Figure 2.7: The time-lapse photography for RCM control with different (a) $\lambda = 1.0$, (b) $\lambda = 0.8$, and (c) $\lambda = 0.6$, while (d), (e), and (f) are the deviation *d_rcm* for (a), (b), and (c), respectively. The yellow point in (a), (b), and (c) are the RCM control points. The whiskers show the minimum and maximum recorded change of the distance while the first and third quartile show the start and the end of the box. Band, red dot, and cross represent median, mean, and outliers of the recorded changes respectively.

λ will lead a larger RCM deviation since we fixed the control loop time to a certain value. Even though the pivoted speed is the same, the varies of λ will contribute to the different speed on the prismatic joint. This can be seen that with the same rotational speed, with a larger λ we can observe larger deviation. The maximum deviation is restricted within 1.0 mm to ensure the safety and accuracy of RCM control.

2.4 Summary

In this chapter, we presents our novel system for assisting ophthalmologists to perform the difficult task of subretinal injection using soft RCM with 5 DoF robot with the hybrid parallel-serial mechanism. The safety and accuracy of RCM is evaluated on the robot with tissue phantom and ex-vivo pig trials. The evaluation results show that the single control loop of the system is executed in 15ms and the RCM control accuracy can be ensured within 1mm, both are within clinical range. With the help of the robot, surgeons can put more effort to be a decision maker during the operation rather than spending too much time for improving their surgical skills to cope with hand tremor and dexterous motion.

2.5 Related Publication

1. **Mingchuan Zhou**, Qiming Yu, Kai Huang, Simeon Mahov, Abouzar Eslami, Mathias Maier, Chris P. Lohmann, Nassir Navab, Daniel Zapp, Alois Knoll, M. Ali Nasseri: A Hybrid Parallel-Serial Robot System Design and Preliminary Evaluation, accepted. IEEE Transactions on Industrial Electronics, 2019.

Chapter 3

6DOF Needle Pose Estimation

3.1 Introduction

In robot-assisted ophthalmic surgery, the accurate 6DOF pose information of needle is essential for image guided operation which can enhance the visual perception for the surgeon. However, it is inaccurate to estimate 6DOF pose of needle tip part via forward kinematics due to the very thin, long and relatively flexible needle body [62]. Many researches have addressed on instrument pose estimation in 2D microscope images [63]. However, 2D images from single microscope are insufficient to estimate the 6DOF movements during the operation. Traditional navigation solutions such as optical tracking or electromagnetic tracking are not applicable as they usually have an accuracy in the range of 200 to 1400 µm. To overcome the above mentioned challenges, Probst et al [64] firstly proposed to use a stereo-microscope to detect and localize the instrument in robot-assisted vitreoretinal surgery, which can achieve a precision of 100 µm. However, vitreoretinal surgery presents unique challenges such as strong illumination changes, blur and even higher accuracy requirements, which limits the stereo-microscope method in some scenarios, e.g. subretinal injection and RVC. Optical Coherence Tomography (OCT), which is originally used for ophthalmic diseases diagnosis for its suitable resolution, has been developed to present real-time image interactions between the surgical tool and intraocular tissue. Microscope-mounted intraoperative OCT solution developed by Carl Zeiss Meditec ($RESCAN700$), firstly described in clinical use in 2014 [30], can share the same optical path with the microscope and give real-time cross section information of the target scan area, which is an ideal imaging modality for ophthalmic surgery. Taking advantage of OCT image modality, we have the chance to obtain the 3D scan of target area intraoperatively.

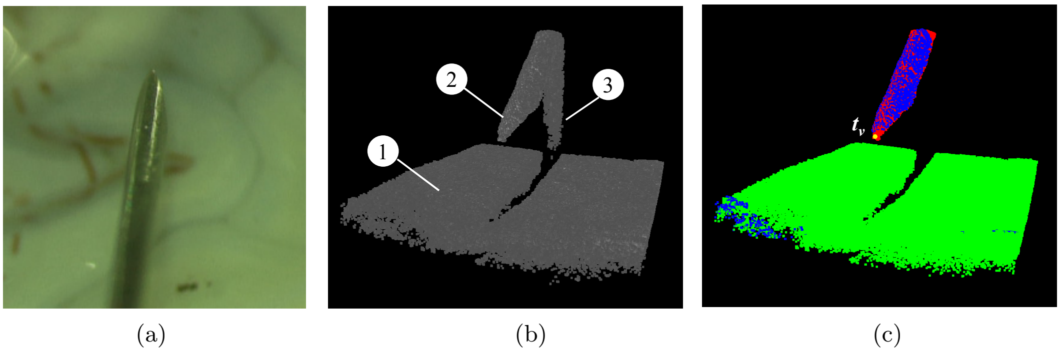
Estimating 6DOF pose of an object from incomplete point cloud draws much attention in computer vision with significant applications. Kehl et al [65] introduced a light-weight 3D tracking with 6DOF pose estimation. However, their method can not be directly applied in our study since the iterative closest point (ICP) with 6DOF parameters heavily relies on the initial guess from the object viewpoint features, e.g. clustered viewpoint feature histogram (CVFH), and the geometrical feature of the needle is cylindrical and has a bevel shape at the needle tip, which does not have strong features. This will lead to a local optima result and may not be suitable for safety-critical surgical applications.

In this chapter, we introduce a modified iterative closest point to estimate the 6DOF pose of the needle directly from the OCT volume data. The main premise for the proposed method is that the actual dimensions of needle are within the range which obeys the standard for manufacturing medical devices (ISO 9626:2016), typical ±6.4 μm in diameter and ±1° in bevel angle [66]. The method consists of two main parts. The first part is a robust needle segmentation method introduced to get the 3D needle point cloud in OCT volume. Due to the infrared light source of OCT and geometrical feature of the needle, the segmentation result will be robust to illumination variation and speck reflection. The second part is a shift-rotate ICP (SR-ICP) to estimate the 6DOF pose of the segmented needle point cloud. Using the geometrical features of the needle, the 6DOF pose is reduced to a 2DOF optimization problem, which can dramatically decrease the chance of local optima. Furthermore, different from the typical methods which use object viewpoint features to start the initial guess, we propose to align the CAD model tip to the visual needle tip in the OCT volume. This initial guess is very close to the global optima. To validate the proposed method we compare the result with brutal grid search (GS) and standard ICP. The experiment performed on the ex-vivo pig eyes demonstrated that the proposed method is qualified in 6DOF needle pose estimation for the vitreoretinal surgery application. Especially, the position accuracy can be controlled within 10 μm with 95% confidence which meets the most of surgical requirements.

The rest of this chapter is organized as follows: Section 3.2 briefly presents related work. The proposed method is introduced in Section 3.3. Then Section 3.4 gives the experimental evaluations to prove the effectiveness of our proposed method. Section 3.5 concludes this chapter with discussion and future work.

3.2 Related Work

Many studies have been carried out with significant progress in the needle tracking through microscopic images [67, 68, 69]. These work got satisfactory results using either color-based or



(a) (b) (c)

Figure 3.1: The needle segmentation. (a) The needle placed above the retina in microscope image. (b) The 3D OCT volume with needle and retina inside (① denotes retina, ② denotes needle, and ③ denotes needle reflection). (c) 2D detected ellipses are colored in blue and used for voting of needle segmentation. The red cluster is the voted needle cluster. The green cluster is the retina. t_v denotes the visual needle tip.

geometry-based features. However, due to the limitation of single 2D microscopic images, these methods can not provide enough information to estimate 6DOF needle pose.

Other 3D medical imaging technologies, including computed tomography (CT) scans, fluoroscopy, magnetic resonance (MR) and ultrasound technology are already applied in the cardiac, brain, and thoracic surgeries, not only for diagnostic procedures but also as a real time surgical guidance [55, 56, 57]. However, these imaging technologies can hardly achieve the ideal resolution for ophthalmic surgery application. For MRI-guided interventions with resolution in millimeters in breast and prostate biopsies, 18 gauge needle with the diameter of 1.27 mm is used, while for ophthalmic surgery, 30 gauge needle, which has the diameter of 0.31 mm, requires resolution of submillimeter [58].

Some researchers used OCT for distance estimation between the surgical tool and eye tissue. Song *et al.* [45] developed a robotic-surgical tool with an OCT probe integrated to estimate the one-dimensional distance between tool tip and tissue. Yu *et al.* [70], and Liu *et al.* [71] applied the OCT probe to assist robotic ophthalmic surgery. However, they focused on integrating the OCT into the surgical system, rather than estimating the instrument pose information. Weiss *et al.* [72] developed an algorithm to track the needle 5DOF pose without needle rotation information. Gessert *et al.* [73] proposed a 3D convolutional neural network (CNN) to segment and estimate the pose of a small marker geometry from OCT volumes directly with a mean error of 14.89±9.3 μm. They used a special marker instead of needle with more geometrical features. The benefit of deep learning is that it can archive the end-to-end pose estimation while reducing the progress of geometrical modeling. However, the drawback is that a big amount of data is required to get a well-trained network, meanwhile, the different type of needles, e.g. with different diameter, contribute to an even more data set preparation and network tuning overhead.

Different from deep learning approach, in this chapter, based on the geometric features of beveled needle and its imaging characteristic in OCT cube, a modified ICP algorithm is proposed and evaluated to obtain the 6DOF pose of needle for ophthalmic surgery.

3.3 Method

The method section contains two parts. The first part is to segment the needle point cloud from the scan area. In the second part, we propose the segmented needle point cloud for needle pose estimation.

3.3.1 Needle Segmentation

The OCT volume is obtained by setting the OCT engine into C-scan mode. One C-scan is constituted by multi B-scan gray images which is cross-sectional image of the scan area. The original B-scan gray image is transformed into a binary image by the adaptive thresholding method B-scan [74]. We eliminate the noise inside the binary image by applying a median filter and a Gaussian filter. A voting mechanism is used to specify whether a point in the B-scan image belongs to the needle body or not. Since the needle body part in the B-scan image is a half ellipse, which can be considered as a strong feature, the ellipse fitting is applied to the topmost contours of each B-scan [74]. In order to filter any fitted ellipse E_i detected other than the needle body, we constrain the ellipse's minor axis m_e to a value less than m_t, where m_e is related to the diameter of needle. The contour with fitted ellipse E_i can also be overlapped by a bounding box $B = (Bx_i, Bz_i, w_i, h_i)$ shown as Fig. 3.2, where Bz_i and By_i are the left corner of the bounding box, w_i and h_i are the width and height of the bounding box, respectively. The bounding box location information can be used to remove the needle reflection part in the original OCT volume. The needle reflection part is generated because needle reaches out of OCT imaging range while blocking the imaging path shown as in Fig. 3.1(b). The E_i will also be filtered if the upper edge of corresponding bounding box B_w reaches close to the image range. The E_i is represented as,

$$E_i = \begin{cases} \varnothing, \text{ if } m_e \geq m_t \text{ or } \overline{Z} - Bz_i \leq \mu \\ E_i, \text{ if } m_e < m_t \text{ or } \overline{Z} - Bz_i > \mu \end{cases} \tag{3.1}$$

where $\overline{Z}$ is the OCT imaging range in Z direction, μ is the tolerance between the upper edge of corresponding bounding box and image range in Z direction.

The OCT volume is represented as a point cloud structure P (see Fig. 3.1(b)) with (x_i, y_i, z_i, b_i), where (x_i, y_i, z_i) is the position of point from P, and b_i is the boolean value to identify whether the point belongs to the needle or not. Afterwards, we differentiate between objects (needle and retina) in the point cloud with euclidean cluster extraction [75]. Then we can obtain a set of euclidean point clusters. The cluster that has the greatest number of points with $b_i = 1$ (i.e. most voting, blue color in Fig. 3.1(c)) will be treated as the needle. To make sure that the visual needle tip is located in the first B-scan of the segmented needle, the B-scan direction can be adjusted manually or automatically to match with the needle insertion direction [74]. The yellow point (shown in Fig. 3.1(c)) is the centroid of the needle cluster in this slice, which represents the needle visual tip t_v.

3.3.2 Needle 6DOF Pose Estimation

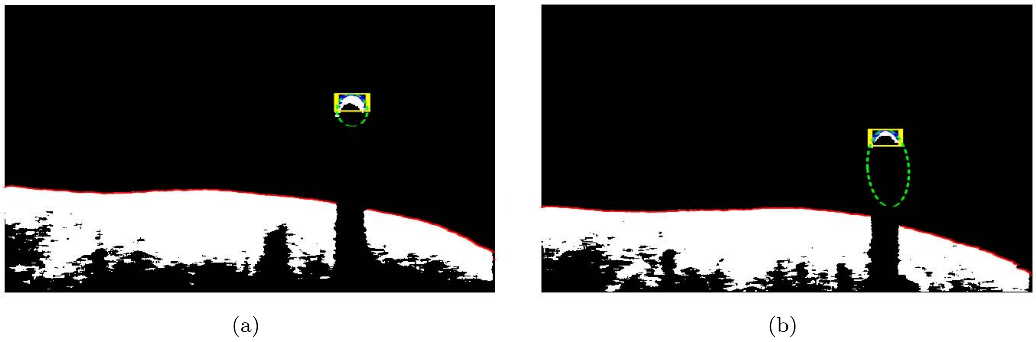

(a) (b)

Figure 3.2: The ellipse fitting in B-scan. The images show a 2 mm × 3 mm area. (a) The correct fitting of ellipse. (b) The fitting of ellipse with large deviation. The fitting of ellipse (green dash line) and bounding box (yellow solid line). The correct fitting ellipse is tangent with the bounding box.

The needle tip part can be treated as a rigid body and the needle pose in OCT can be shown as in Fig. 5.4(a). The needle 6DOF can be defined by the needle tip point position, the needle center axis direction n and the rotation angle φ. In order to localize the space line L which is the center axis of the needle, the center of the ellipse e_i in each B-scan need to be calculated. The straightforward way is to fit the needle pixel with a ellipse equation, however, this half ellipse contains noise which could easily lead to a large deviation to the ground truth, see Fig. 3.2. Therefore, we investigate the geometrical relationship of the ellipse and the bounding box which

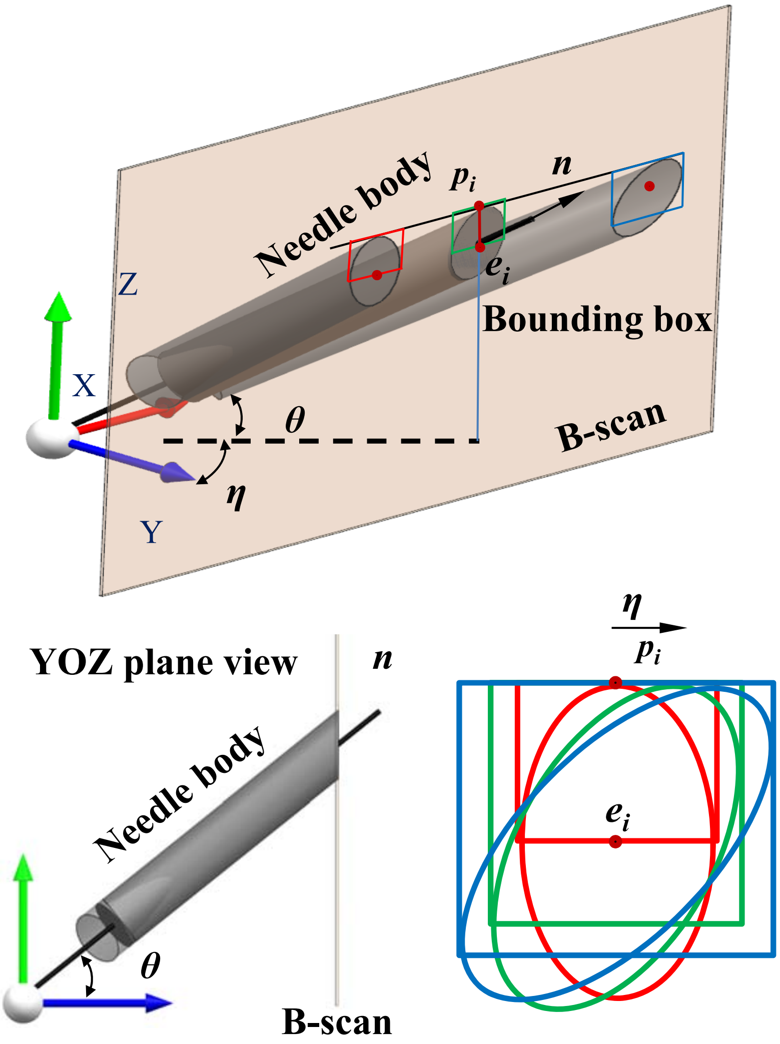
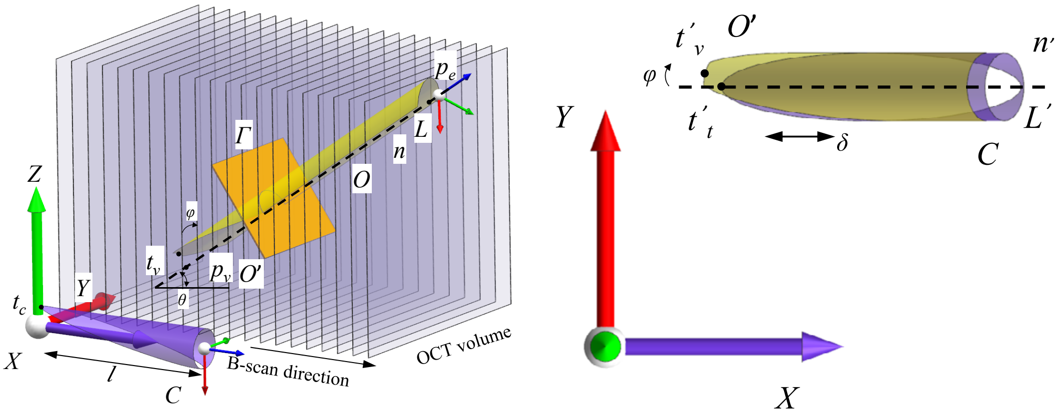


Figure 3.3: The geometrical relationship of the ellipse and the bounding box.

is robust and less computation overhead. As shown in Fig. 3.3, η is the angle between n and YOZ, and θ is the angle between n and XOY. From the YOZ plane view, we can find that with η changing, the distance between p_i and e_i remains the same, where e_i is the center of ellipse and p_i is the middle point for the top border of bounding box. The distance of p_ie_i can be obtained by known needle diameter d and θ. Then we can calculate the e_i as follows,

$$e_i = (Bx_i, By_i, Bz_i - \frac{d}{2\cos\theta}) \tag{3.2}$$

To analyze the needle pose in 3D OCT volume, we can refer Fig. 5.4(a). Considering that the target needle point cloud has the fixed axis L which we could calculate from e_i by RANSAC line fitting. The source needle point cloud can be slid along and rotated around L. Therefore, the needle pose can be determined by L, δ and φ, where δ is the shift distance along this axis, and φ is the rotation around this axis.



(a) (b)

Figure 3.4: (a) The illustration of needle pose in OCT volume. (b) The initial guess of SR-ICP is shifted and rotated value of C to match t'_v and t'_t on the projection.

In order to make the needle in OCT and the CAD model the same length while keep most of the point information (bevel part), which helps to determine the 6DOF pose, we crop the OCT point cloud to have a consistent length with the CAD model by using split plane Γ, shown in Fig. 5.4(a). The CAD needle point cloud and OCT needle point cloud are respectively denoted as $C = \{c_i\}_{i=1}^{N_c} \in \mathbb{R}^3$ and $O = \{o_i\}_{i=1}^{N_o} \in \mathbb{R}^3$. The cropped OCT needle point cloud $O' = \{o'_i\}_{i=1}^{N_{o'}} \in \mathbb{R}^3 \subseteq O$ is cropped by Γ, which can be calculated as:

$$\Gamma : n_x(x - x_v - ln_x) + n_y(y - y_v - ln_y) + n_z(z - z_v - ln_z) = 0 \tag{3.3}$$

where $n = (n_x, n_y, n_z)$ denotes the unit direction vector for space line L, $t_v = (x_v, y_v, z_v)$ denotes the visual needle tip in the OCT needle point cloud, l denotes the length of needle for matching. l is selected the by minimum value between the length of bevel part for needle CAD model and the length of needle in OCT volume, which could be calculated as,

$$l = \min(l_c, l_o) \begin{cases} l_c = d/\cos(\beta) \\ l_o = \|p_v p_e\| - d\tan\theta/2 \end{cases} \tag{3.4}$$

where p_v is the projection point of t_v on the center axis line L, and p_e is the intersection point of line L and the last B-scan plane. p_v is calculated by,

$$p_v = p_a + \frac{p_a t_v \cdot p_a p_b}{p_a p_b \cdot p_a p_b} \times p_a p_b \tag{3.5}$$

where p_a and p_b are any two different points on the line L. t_v is obtained from needle segmentation.

As previously mentioned, we obtain the adjusted target needle point cloud O' from OCT images and the source point cloud C from CAD model, as well as the fixed needle center axis L. In the following, we introduce the SR-ICP with 2DOF (shift along and rotate around L). We achieve the 2DOF by constraining the center axis of C by the center axis of O'. The proposed matching algorithm is consecutively minimizing the following functions,

$$C(i) = \underset{j \in \{1,...,N_c\}}{\operatorname{argmin}} \|(A_{k-1}c_i + t_{k-1}) - o'_i\|_2^2, \text{ for all } i \in \{1, ..., N_{o'}\} \tag{3.6}$$

$$(A_k, t_k) = \underset{A,t}{\operatorname{argmin}} \frac{1}{N_{o'}} \sum_{i=1}^{N_{o'}} \|(Ac_i + t) - c_i\|_2^2 \tag{3.7}$$

$$s.t. \begin{cases} A = (\cos\varphi)I + (\sin\varphi)[n]_\times + (1 - \cos\varphi)(n \otimes n), \\ t = p_v - \delta n \end{cases} \tag{3.8}$$

Equation 3.6 finds the correspondence $C(i)$ ($i = 1, ..., N$ where N is the number of correspondence) between CAD and OCT point cloud. Equation 3.6 minimizes the Euclidean distance between the cropped needle points in OCT volume and the transformed points of the needle CAD model with constrains of Equation 3.7. A and t are the affine and translational matrices. A_k and t_k are the desired affine and translational matrices at iteration k. Equation 3.8 is the constraint for angle φ rotating around L and shifting distance δ along L. The rotation is clockwise and the shift is starting from the viusal tip p_v. I is the identity matrix. $[n]_\times$ is the cross product matrix of n, where can be calculated as,

$$[n]_\times = \begin{bmatrix} 0 & -n_z & -n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix} \tag{3.9}$$

$\otimes$ is the tensor product, where can be calculated as,

$$n \otimes n = \begin{bmatrix} n_x^2 & n_x n_y & n_x n_z \\ n_x n_y & n_y^2 & n_y n_z \\ n_x n_z & n_y n_z & n_z^2 \end{bmatrix} \tag{3.10}$$

The Eqn. 3.7 is typical hyperparameter optimization problem which can be solved by gradient-based and grid search methods. The grid search method is exhaustive searching through a manually specified subset of hyperparameter space, which can avoid the local minima but at the cost of high computation overhead. Gradient-based method, by contrast, is much less in computation overhead but can lead to a local minima.

The proper search space of hyperparameter and initial guess are critical for the performance of both methods. As shown in Fig. 5.4 (b), L', n', t_t', and t_v' are the projection of L, n, t_t, and t_v on plane XOY, respectively. t_t is the needle CAD model tip after transformation from the original position. Due to the fact that the OCT scan has a certain resolution and the optical coherence feature will miss the very tip point reflection in the OCT images, the visual needle tip t_v is not position of real needle tip but the distance between t_v and the real needle tip is close. The projection of actual needle tip on XOY is also close to t_v'. Therefore, we could use t_v' as the initial guess, φ and δ could be adjusted to make the two points t_v' and t_t' overlaping with each other. The output of rotation angle and shift distance value are denoted as φ_0 and δ_0, which are the initial search values for SR-ICP. The search space can be box restricted with reference of the initial guess.

$$(\delta_0, \varphi_0) = \underset{\delta,\varphi}{\operatorname{argmin}} \|(At_t + t)_{XOY} - t_v'\|_2^2 \tag{3.11}$$

where $(At_t + t)_{XOY}$ donates the projection of $(At_t + t)$ on XOY plane. The Levenberg-Marquardt algorithm is introduced to solve Eqn. 3.11. The gradient-based method with Levenberg-Marquardt algorithm (LM based SR-ICP) and grid search (GS based SR-ICP) method with several iterations are programmed to solve Eqn. 3.7 to obtain the optimized A and t. The search area is constrained in a box area with $\delta \in [\delta_0 + \triangle\delta, \delta_0 - \triangle\delta]$, $\delta \in [\varphi_0 + \triangle\varphi, \varphi_0 - \triangle\varphi]$, where $\triangle\varphi$ and $\triangle\delta$ are the regulation parameter for adjusting the searching range.

With the desired output affine and translation value A_k and t_k, the estimated needle tip position can be calculated as,

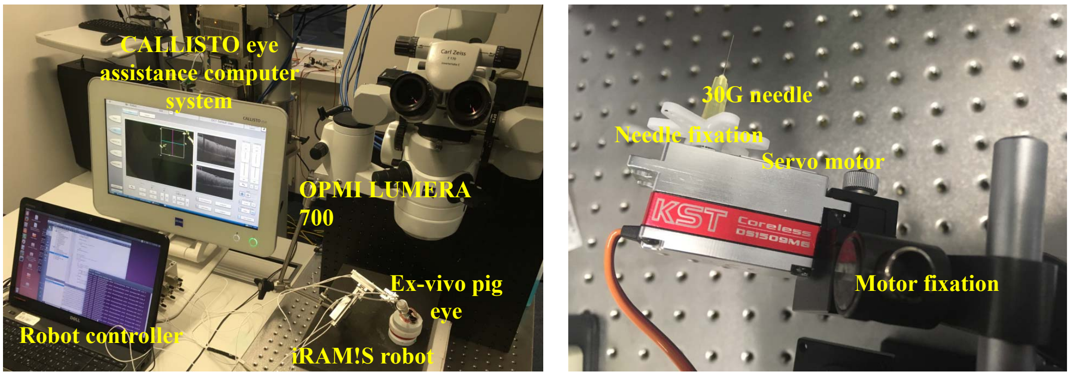
$$t_t = A_k t_c + t_k \tag{3.12}$$

where t_c is the tip position for needle CAD model in the original position which can be calculated by,

$$t_c = (d/2, 0, 0) \tag{3.13}$$

3.4 Experiments and Results

The experimental setup is depicted in Fig. 3.5. The OCT engine is set to the maximum speed available with 27000 A-scans per second in cube resolution of 128×512×1024 in the corresponding scan range of 3 mm×3 mm×2 mm. The movement of needle can be controlled



(a) (b)

Figure 3.5: (a) The experiment setup with iRAM!S eye surgical robot. (b) The needle rotation experiment setup with a servo motor.

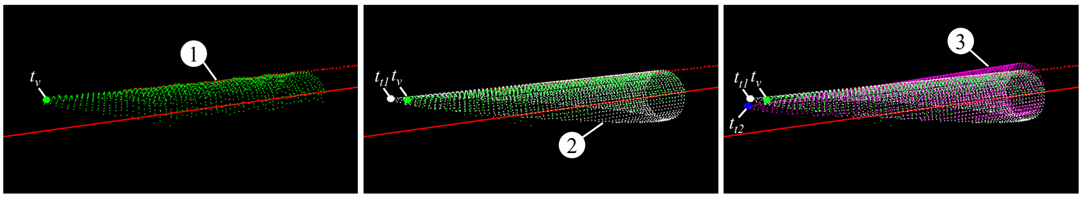
by the surgical robot named iRAM!S [76] with piezo motors' accuracy of 1 µm (SmarACT GmbH, Germany), as shown in Fig. 3.5(a). Since there is no roll control of needle in the robot, we use a servo motor (Security GmbH, Germany) with 1° accuracy to estimate the performance of different rotation of needle, see Fig. 3.5(b). The framework is implemented on the *CALLISTO* eye assistance computer system with Intel Core i7 CPU of 200 ms processing time for needle segmentation to obtain the needle point cloud with parallel programming in C++ with OpenCV 3.4 and Point Cloud Library (PCL) 1.8 library. The experiments contains two parts, (1) the performance analysis of SR-ICP (GS and LM based) and normal ICP (standard algorithm from PCL 1.8 library), and (2) the accuracy validation of needle movement and needle rotation for SR-ICP.

Fig. 3.6 shows needle point cloud matching results for the LM based SR-ICP. As we can see in Fig. 3.6(c), SR-ICP resulted in a better matching (white) compared to the normal ICP method (purple). The normal ICP trends to mismatch with the point cloud without the constraint of the needle center axis. Fig. 3.7 shows the performance of three methods with different point cloud sample size. As soon as the larger sample size is selected, the sparser point cloud is obtained. The sum of distance is the all over distance for the correspondence with the optimized A_k and $\tilde{t_k}$ which can be calculated as,

$$\sum_{i=1}^{N_{o'}} \|(A_k c_i + t_k) - c_i\|_2^2 \tag{3.14}$$

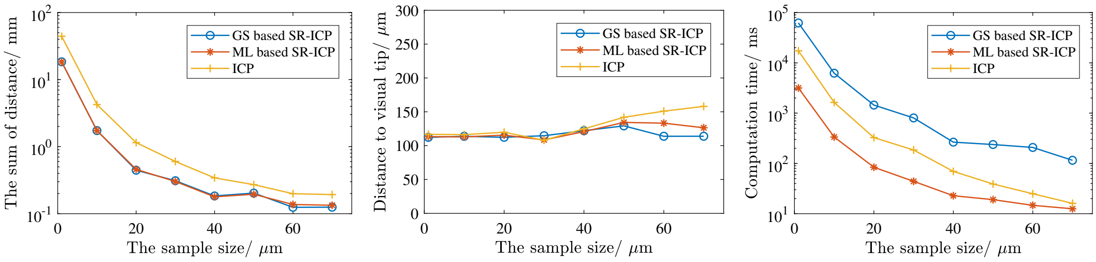
We can see that the ML based SR-ICP and GS based SR-ICP perform better than ICP in matching with smaller sum of distances for corresponding points (see Fig. 3.7(a)). In Fig. 3.7(b), the distance between estimated and visual needle tip does not have significant difference, which means that the estimated needle tip would be similar for all three methods. When the sample size increases to 60 µm, the ICP trends to be far away from the visual tip than the ICP which indicates that the SR-ICP has a better performance when the point sample size is large.

Fig. 3.7(c) shows that LM based SR-ICP has the best computation performance. When the sample size is 10 µm, the computation time is 335.4 ms for the LM based SR-ICP, which is 4.9 and 18.6 times faster than the ICP and GS based SR-ICP, respectively.



(a) (b) (c)

Figure 3.6: (a) The cropped OCT needle point cloud (① denotes needle point cloud with green color, the red points are the middle point for the top border of bounding box, and the red line is the fitted axis line L). (b) The LM based SR-ICP matching result (② denotes the transformed CAD needle point cloud using SR-ICP in white color). (c) The comparison of LM based SR-ICP matching and ICP result (③ denotes the transformed CAD needle point cloud using ICP in purple color). t_v denotes visual needle tip with white dot, t_{t1} and t_{t2} denotes the estimated needle tip via SR-ICP with green dot and ICP with blue dot, respectively.



(a) (b) (c)

Figure 3.7: (a) Sum of distances between corresponding points calculated by three methods with change of sample size. (b) The distance between estimated and visual needle tip calculated by three methods with change of sample size. (c) The computation time by three methods with change of sample size.

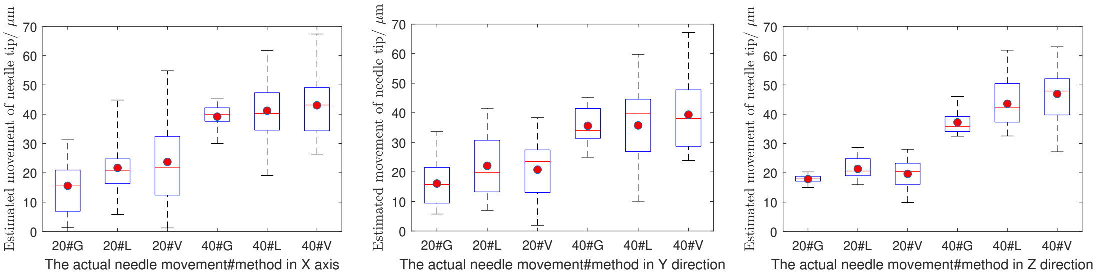
In the second part of the experiment we demonstrate the effectiveness of the proposed method. The point cloud sample size is set as 10 µm considering the algorithm processing speed. As the experiment setup in Fig. 3.5, the robot is controlled to move the needle in X, Y, and Z direction with 20 and 40 µm and each movement is repeated 25 times. The robot is also controlled to yaw and pitch around the needle tip as the RCM control point with virtual fixture control [77]. The yaw and rotation angle is set from -45° to 45° with 15 ° as the step. The servo motor is controlled to rotate the needle from -45° to 45° in 15° increments and each angle is collected 10 times. As we can see from Fig 3.8, the mean error of needle tip distance estimation is 3.1 µm (maximum 18.7 µm) for LM based SR-ICP, 2.4 µm for GS based SR-ICP (maximum 29.9 µm), and 2.6 µm for visual tip (maximum 34.8 µm). We can find that the mean error accuracy does not have significant difference, while the SR-ICP has the lower maximum value compared to visual tip method. The main reason is that the visual tip is directly obtained from the images which varies more depending on the resolution of the OCT scan in each direction. The standard deviation is 6.0 µm for GS based SR-ICP, 10.1 µm for LM based SR-ICP, and 10.7 µm for visual tip, which demonstrates that the GS based SR-ICP have the most steady output results. This verifies that the GS based SR-ICP is slow but more robust for single evaluation. In order to further analysis the performance of these method, we put all the movement error data in processing method with different evaluation metrics and list them in Table 5.1.

Table 3.1: Descriptive statistics for needle tip position error (in micron).

	AE	ME	RSME	0.95CI
$\bar{e_G}^1$	-3.1	18.7	6.7	[-14.6, 9.2]
$\bar{e_L}^1$	0.9	29.9	10.1	[-24.4, 21.6]
$\bar{e_V}^1$	2.2	34.8	10.9	[-16.1, 25.0]
$\bar{e_G}^3$	9.5	3.7	4.1	[-7.8, 1.5]
$\bar{e_L}^3$	0.8	15.2	5.0	[-9.6, 9.8]
$\bar{e_V}^3$	1.9	17.2	4.7	[-7.3, 10.9]
$\bar{e_G}^5$	-2.8	6.8	3.1	[-6.0, 0.3]
$\bar{e_L}^5$	1.9	9.4	3.4	[-3.9, 7.7]
$\bar{e_V}^5$	234.1	230.0	4.1	[-3.7, 9.6]

The evaluation metrics includes average error of estimated value to the ground truth (AE), the error of maximum estimated value and actual value (ME), the root mean square error (RSME),and the 95% of estimated confidence interval (0.95CI). $\bar{e_G}^n$, $\bar{e_L}^n$ and $\bar{e_V}^n$ denote randomly selecting n estimated error values from 150 data samples with the methods of LM based SR-ICP, GS based SR-ICP, and visual tip, respectively, and averaging these values. The 0.95CI for $\bar{e_G}^n$, $\bar{e_L}^n$ and $\bar{e_V}^n$ can be significantly narrowed down with the increasing of n. The advantage of the GS based SR-ICP method is that the processing speed is faster with the benefit of averaging more data in the same time duration. Table 5.1 shows that with more data samples averaging, the accuracy of the needle position can be increased. When 3 data samples are averaged, 10 µm needle tip position accuracy can be achieved with confidence of 95% for

GS based SR-ICP and LM based SR-ICP.



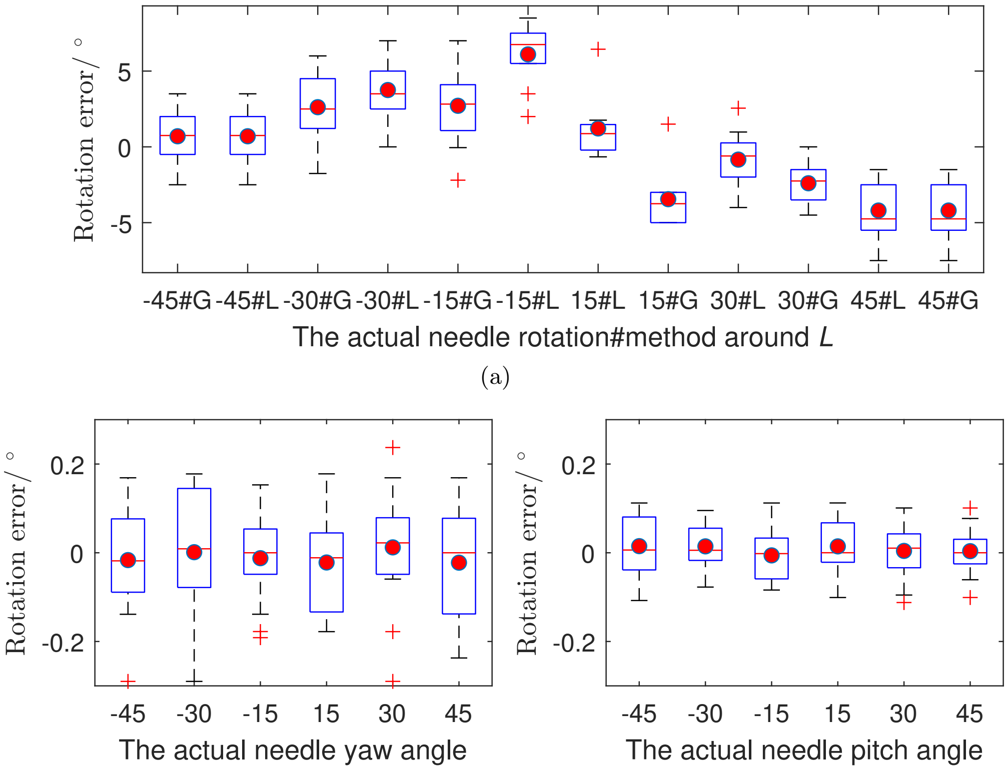
(a) (b) (c)

Figure 3.8: The actual movement for the needle tip and the estimated movement calculated by the Grid based SR-ICP, LM based SR-ICP, and visual tip with 20 and 40 μm in (a) X direction, (b) Y direction, and (c) Z direction. The whiskers show the minimum and maximum recorded change of the distance while the first and third quartile show the start and the end of the box. Band, red dot, and cross represent median, mean, and outliers of the recorded changes respectively.

Fig. 3.9 shows the performance of needle rotation. The needle roll angle around L can be calculated by the methods of GS based SR-ICP and LM based SR-ICP. The yaw and pitch angle is calculated by the equation of line L, thus the result of yaw and pitch angle are the same for the GS based SR-ICP and LM based SR-ICP. The mean error of the estimated needle roll angle is 0.367°(maximum 7.5°) and 0.0835°(maximum 8.5°) for the GS based SR-ICP and LM based SR-ICP, respectively. The performance of roll angle estimation between two methods is similar with the consideration of the servo motor error is 1°. The mean error is 0.010° (maximum 0.290°) for yaw angle, and 0.008° (maximum 0.112°) for pitch angle. The above two parts of experiment demonstrate that the SR-ICP has a better estimation result compared to the normal ICP algorithm. There is no significant difference between the GS based SR-ICP, LM based SR-ICP, and visual tip regarding the performance for relative position accuracy of the needle tip. However, visual tip method only provide needle tip information without needle rotation information. LM based SR-ICP have the advantage of similar optima output ability with GS based SR-ICP but much less computational overhead.

3.5 Summary

This chapter presents a novel approach to estimate the 6DOF needle pose directly from the OCT volume. A SR-ICP algrithm is proposed to estimated the incompleted needle point cloud after the segmentation of needle from OCT volume. Two solving strategies are used to solve the optimum problem that is GS based SR-ICP and LM based SR-ICP. Both of the methods



(a)

(b) (c)

Figure 3.9: (a) Sum of distances between corresponding points calculated by three methods with different sample size. (b) The distance between estimated and visual needle tip calculated by three methods with change of sample size. (c) The computation time by three methods with change of sample size.

have the ability of estimating the needle tip position in accuracy of 10 μm with confidence of 95%, which meets the positioning accuracy requirement for most surgical applications in vitreoretinal surgery. The mean error of roll angle, yaw angle, and pitch angle are 0.0835°, 0.010°, and 0.008°, respectively. Our future work will focus on path planning and trajectory design in a way that the drug can be delivered directly and precisely to the designed area with 6DOF needle navigation.

3.6 Related Publication

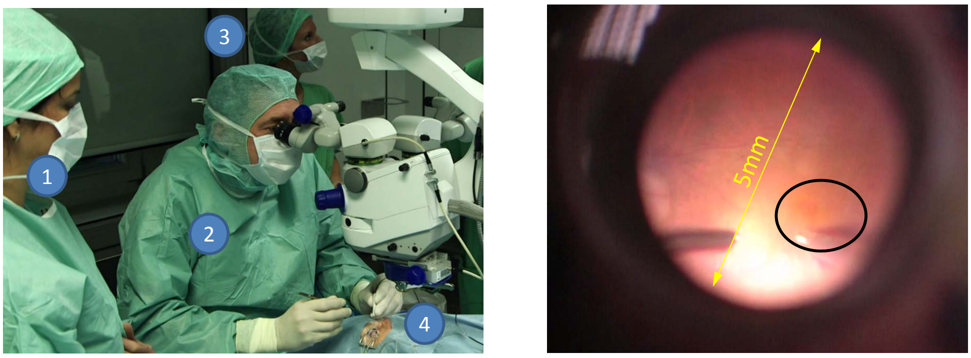
1. **Mingchuan Zhou***, Xing Hao*, Abouzar Eslami, Chris P. Lohmann, Nassir Navab, Alois Knoll, M. Ali Nasseri. 6DOF Needle Pose Estimation for Robot-assisted Vitreoretinal Surgery. IEEE Access.

Chapter 4

Hand Eye Calibration

4.1 Introduction

In this chapter, we will introduce the hand eye calibration with the micro precision robot (hand) and the OCT engine (eye).



(a) Ophthalmic operation

(b) Microscope view of focal region

Figure 4.1: (a): A conventional setting in an ophthalmic operation room: a surgeon (2); two assistants (1) and (3); the patient (4). (b): The focal region in the microscopic view. The operation area is marked by an ellipse.

As we mentioned in the previous chapter, ophthalmic surgery is a typical microsurgery with a delicate and complex workflow, which needs critical surgical skills and considerations (Fig. 5.2). In addition to conventional en-face images from the ophthalmic microscopes, OCT is currently known to be a proper imaging modality for visualizing the micro structural anatomy of the eye [78], as it offers suitable resolution and non-invasive radiation. Compared to the stereo-microscope vision system proposed by Probst et al. [64], the drawbacks of OCT are expensive hardware, acquisition and processing time, while the advantages are (a) the imaging quality will not be influenced by the illumination, which in some cases (e.g. vitreoretinal surgery, an operation on the posterior segment of the eye) is not easy to control, and (b) the OCT can provide information inside the tissue, which allows surgeons to analyze the anatomical structures. The latest OCT machines have been developed to even give real-time information of the interior and posterior segment of the eye as well as the interactions between surgical instruments and intraocular tissues [30].

Some research groups integrated the OCT probe into surgical instruments for navigation in head surgery [79] and ophthalmic surgery [80]. However, to localize the global location of a probe, this may need preoperative Computer Tomography (CT) or Magnetic Resonance Imaging (MRI). These extra preoperative images and modified surgical instruments may increase the complexity of the application.

Instead of using OCT integrated instruments, we propose using the microscope-integrated OCT (MI-OCT) for RAS navigation. In MI-OCT, the OCT engine is integrated inside the microscope head and shares the same optical path with the microscope in order to capture volumetric images, thus the OCT probe avoids integration into surgical instrument. During the planned MI-OCT guided RAS, the surgeon will move the microscope head and adjust its view to locate the focal region (Fig. 4.1(b)). Usually, the microscope imaging area is enough to cover the focal area, thus the microscope head could be fixed. Following this fixation, the surgeon will manipulate the robot and move the needle close to the focal area. Therefore, hand-eye calibration is essential for such applications e.g. keeping a safe distance between needle tip and retina or performing path planning for targeted drug delivery.

In this chapter, we address the problem of hand-eye calibration for a surgical robot with the MI-OCT camera. For this purpose, we use a surgical robot developed by [76]. In this setup, head and eye fixations are used to reduce the risk of movement caused by the patient. The imaging device and the robot should also be rigidly connected to the operation bed following the configuration of a da Vinci surgical system. These setups are necessary for the RAS surgery and will usually require only a one-time calibration. To the best of our knowledge, this is the first time that proposes such a calibration framework. The main contributions of this chapter are summarized as follows:

- We propose a flexible framework for hand-eye calibration of an ophthalmic robot and OCT camera without using additional markers;
- A robust method for needle tip segmentation and localization in OCT volume is proposed;

- Clinical grade accuracy performance (9.2 µm) of the calibration method is demonstrated by experimental evaluations on ex-vivo pig eyes.

The remainder of the chapter is organized as follows. In the next section, we briefly present related work. The proposed hand-eye calibration framework is described in Section 4.3. In Section 4.4, the performance of the proposed method is evaluated. Experimental results are discussed in Section 4.5. Finally, Section 4.6 concludes this chapter and presents the future work.

4.2 Related Work

Hand-eye calibration is an active research topic in robotics and has been investigated by many researchers. In particular, the seminal work of Tsai et al. since 1980s [81]. There exist several off-the-shelf calibration solutions that implement Tsai's method (e.g. ViSP library [82], which is designed for visual servoing applications). These existing off-the-shelf calibration solutions can be easily integrated into standard camera vision systems and allow fast prototyping. However, they require separate measurements and extra equipment (calibration grids, markers,...). In some medical scenarios, it is not practical to place the calibration objects or markers inside the tissue or bodily organ. In addition, medical imaging modalities (e.g OCT and ultrasound) have different imaging principles, which leads to a different camera calibration process.

In the surgical environment, hand-eye calibration is normally transferred into a Procrustean problem by the [83] and has been modified to meet with different surgical requirements in various 3D medical imaging scenarios, e.g. two-camera stereoscopic vision and ultrasound. To determine the unknown fixed transformation between the robot and imaging coordinates, a ball-tip stylus or chessboard is used as a calibrator for hand-eye calibration [84]. However, these markers may influence the conventional surgical workflow as well as cause complication with sterility. Thompson et al. [85] designed a tracking collar for rigid laparoscopes hand-eye calibration. The method can achieve 0.85 mm error, and the calibration process can be finished within 3 minutes. In another work, 3D markers were used by Bergmeir et al. [86] to realize ultrasound probe calibration with a robotic arm. In terms of none marker research work, Sarrazin et al. [87] used ultrasound volume registration data to achieve calibration that does not require 3D localizers. Pachtrachai et al. [88] proposed a calibration method where an instrument with known geometry is used instead of an additional calibrator. The calibration was achieved without compromising sterility. Francisco et al. [89, 90] used constraints based on the needle orientation in 2D/3D ultrasound image to calibrate the ultrasound probe.

Taking the above research work and requirements of ophthalmic surgery into account, we can conclude that the following considerations need to be taken into account for a qualified calibration method: (a) external calibration objects are best avoided to reduce surgical complication; (b) small computational overhead can make the method easier for online implementation; (c) the calibration accuracy needs to be kept within clinical tolerance. Unlike the aforementioned surgical applications, the intended operation area for ophthalmic surgery is very small and the precision requirement is much higher. The precision requirement of ophthalmic surgery varies depending on the specific operation. For the ILM peeling process, the surgeon can do it quite well, which means the 182 µm accuracy (the hand tremor RMS amplitude [25]) is enough. For the sub-retinal injection, the average thickness of retina is around 200 µm, therefore, 20 µm would be an acceptable position accuracy. For the retinal vein cannulation, the ideal position accuracy would be 20 µm, since the diameter of branch retinal veins is typically less than 200 µm. Therefore, the typical position accuracy for robotic eye surgery is considered to be roughly around 10 µm [26, 27, 28] to meet all potential operations.

Due to the properties of microsurgeries, it will be complicated to properly place calibration markers during surgery, specifically for the posterior parts, where the markers could move inside the eye. To overcome the high accuracy requirement challenge, we propose using the robot's precise movement, in three orthogonal directions, and the detected needle tip position, to obtain the transformation matrix between the camera and the robot coordinate system.

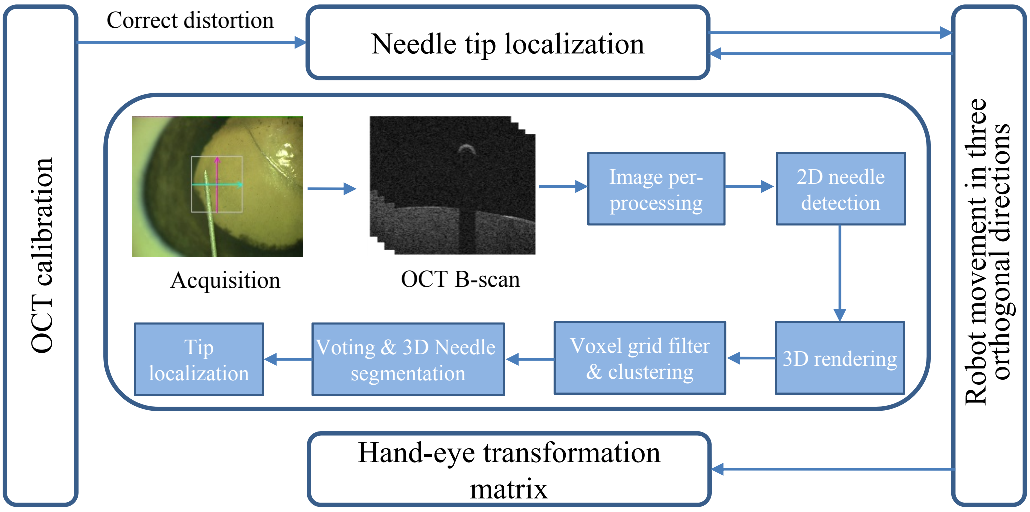


Figure 4.2: The framework of the proposed method.

4.3 Method

The proposed framework is shown in Fig. 6.1. The OCT calibration parameters are obtained in advance to eliminate the distortion caused by the imaging device. With micro-scale displacements controlled by the robot, we move the needle in a trajectory that contains three orthogonal directions. Following each movement, an OCT volume is captured, and the needle tip is detected and accurately segmented even with the presence of image noise. After finishing all the movements defined by the trajectory, we calculate the transformation between the coordinate system of the OCT camera and the coordinate system of the robot. The details of each step are presented as follows.

4.3.1 OCT calibration

Our setup contains an OPMI LUMERA 700 with integrated RESCAN 700 OCT (Carl Zeiss Meditec AG., Germany). The OCT engine is spectral domain OCT, which has improved scanning speed compared to time domain OCT. The OCT engine has a wavelength of 840 nm and a scanning speed of 54000 A-scans per second. Due to the fact that the intended surgical area for ophthalmic operation is usually very small, we set the internal mirror galvanometer, which can deflect a light beam by an electric current, in the OCT to obtain the scan area 3.01 mm×3.10 mm×2.60 mm with the highest resolution of 128 B-scans, each with 512 A-scans. B-scan refers to a two dimensional, cross-section scan and A-scan is one dimensional scan.

The simplified principle of an OCT scanner is shown in Fig. 4.3. The cross-sectional profiles of the flat surface is curved, because the two scan mirrors M_x and M_y have different scan radius which arch the beaming path. During the scanning in the X- and Y- direction, M_y and M_x are fixed, respectively. Therefore, there are geometric distortion effects imposed onto the OCT volume by two mirror galvanometers, which are considered to be decoupled from each other. We refer to the method which is proposed by Van der Jeught et al. [91] to correct the OCT distortion. The glass slider (26 mm×76 mm×1 mm) with flatness error of ±3 µm is used as the flat surface (Brand GmbH, Germany). The topmost surface of the mirror is recognized with the threshold filter in each B-scan image. The radius of the topmost surface is fitted by the analytical equation of a circle, thereby the virtual pivot center (x_c, z_{xc}) of the mirror galvanometer is calculated by averaging all fitted circle centers for the B-scans in X-direction (B_x). We repeat the same process to obtain the virtual pivot center (y_c, z_{yc}) with all the B-scans in Y-direction (B_y). The distortion correction equation is reported below,

$$\begin{cases} x' = R_x(x,z)sin(\theta(x,z)) + x_c \\ y' = R_y(y,z^*)sin(\varphi(y,z^*)) + y_c \\ z' = R_y(y,z^*)cos(\varphi(y,z^*)) + z_{yc} \end{cases} \tag{4.1}$$

where the point (x', y', z') is the corrected result of the point (x, y, z) in the original volume, $R_x(x, y)$ denotes the distance between virtual pivot center (x_c, z_{xc}) and (x, z), $sin(\theta(x, z))$ denotes the angle of polar coordinate system with the virtual pivot center of M_x, z^* which equals $R_x(x, z)cos(\theta(x, z)) + z_{xc}$ is the corrected z value for the distortion causing from M_x scanning mirror, $R_y(y, z^*)$ denotes the distance between virtual pivot center (y_c, z_{yc}) and (y, z^*), $\varphi(y, z^*)$ denotes the angle of the polar coordinate system with the virtual pivot center M_y.

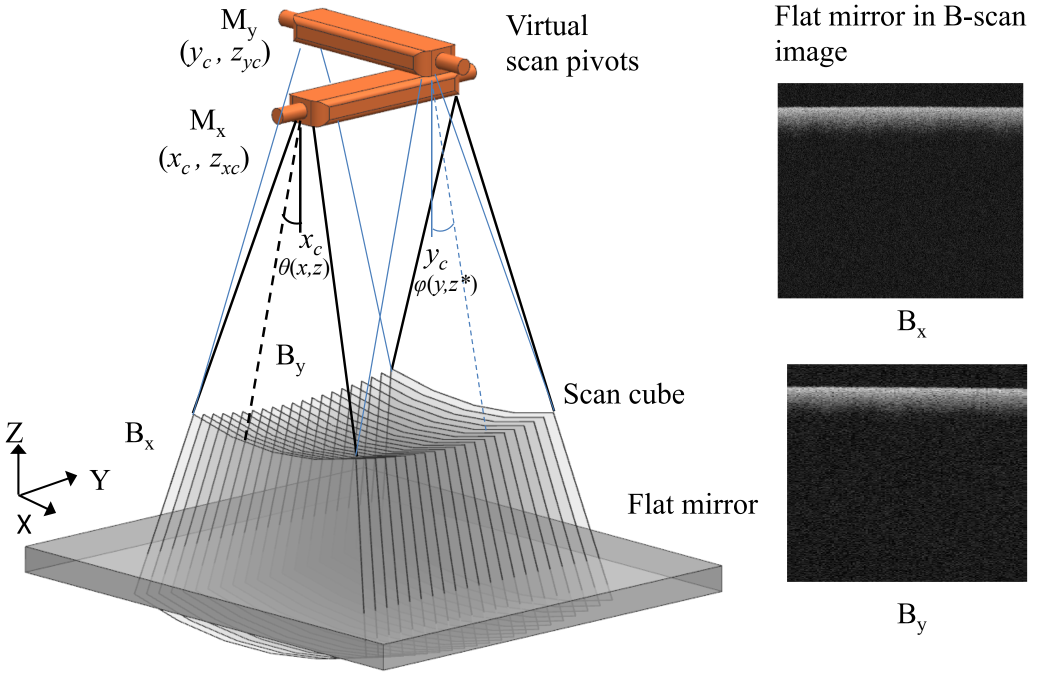
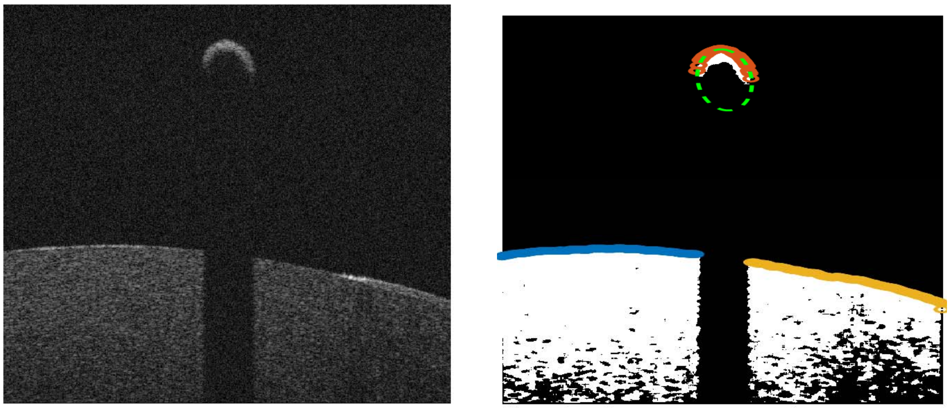


Figure 4.3: The distortion of OCT scan and the cross-sectional B-scan images.

4.3.2 Needle tip localization

Instead of using 3D markers for hand-eye calibration, in our approach, we directly utilize a robust and precise method to localize the needle tip in OCT images. The actual tip is invisible due to the light diffraction and its optical property, therefore, this detected needle tip is the smallest tip part in the OCT images. The original B-scan gray image (Fig. 4.4(a)) is transformed into a binary image by the thresholding method. The threshold value is adaptively defined, based on the statistical measurements of each B-scan [74]. We eliminate the noise inside the binary image, by applying a median filter followed by a Gaussian filter. The sizes of the median and Gaussian filters are much less than the visible needle tip. Therefore, the smallest visible tip part will not be missed as a result of the de-noising procedure.



(a) Orginal B-scan image

(b) Needle detection

Figure 4.4: The original B-scan image (a) and needle detection using ellipse fitting for topmost of contours (b). The ellipse with green dash line is the detected ellipse for needle body. The contours are marked with different colors for different connection areas.

A voting mechanism is used to specify whether a point in the 2D image belongs to the needle body. Since the needle body part in the B-scan image is a half ellipse, which can be considered as a strong feature, the ellipse fitting is applied to the topmost contours of each B-scan. The topmost contours are obtained by scanning each column of pixels going from top to bottom (Z-direction), starting from the leftmost column and moving to the right (X-direction). The connected contour pixels (blue, yellow, and red in Fig. 4.4(b)) are considered as one group and fitted to an ellipse equation. In order to filter any ellipse detected other than the needle body (e.g. blue&yellow eye tissue), we constrain the ellipse's minor axis to a value lower than m_e, where m_e is defined based on the diameter of the instrument (i.e. needle). The fitted ellipse is shown as a green color in Fig. 4.4(b). Finally, an additional property (Boolean value) is added to each pixel in B-scan. The Boolean value is true if the pixel belongs to the fitted ellipse. Otherwise, it is false.

In order to render the images in 3D, the OCT volume is represented as a point cloud structure $\mathcal{P}$ (Fig. 4.5(a)) with (x_i, y_i, z_i, b_i), where (x_i, y_i, z_i) is the position of point $p_i \in \mathcal{P}$, and b_i is the Boolean value to identify whether the point belongs to the needle. To reduce the excessive amounts of data and computation overhead, a voxel grid filter with a leaf size of 0.02 mm is used. Afterwards, we need to differentiate between objects in the point cloud. For this, we use Euclidean cluster extraction. The clustering step uses a Kd-tree structure for finding the nearest neighbors of every point in the cloud. Two points p_i and p_j belong to two different clusters c_i and c_j if:

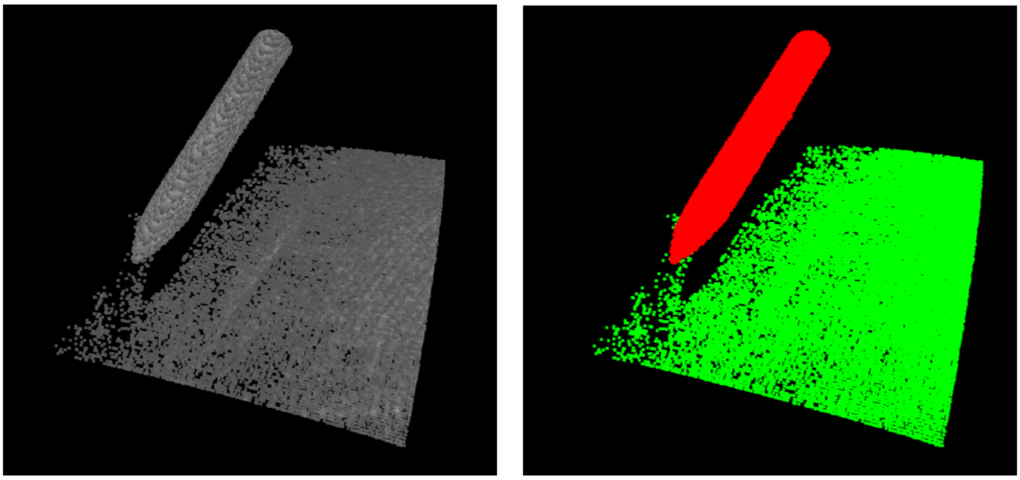
$$min \parallel p_i - p_j \parallel_2 \geq t \tag{4.2}$$

where t is the maximum imposed distance threshold. The result of these processing steps (Fig. 4.5(b)) is a set of euclidean point clusters $C = \{(c_1)...(c_n)\}$, where each segmented cluster c_i is at least t far away from the other cluster. At this point, we have two different clusters, but we do not know which one is the needle. To segment the needle (see Algorithm 2), the cluster that has the greatest number of points with $b_i = 1$ (i.e. most voting, blue color in Fig. 4.5(c)) will be treated as the needle (Fig. 4.5(d)). In order to localize the needle tip, the B-scan direction can be adjusted manually or automatically to match with the needle insertion direction [74]. The alignment will not add extra complexity since it only costs several seconds manually with the scanner control panel of RESCAN 700 OCT. As a result, the needle tip will always be located in the first B-scan (slice) of the segmented needle. The yellow point (Fig. 4.5(d)) is the centroid of the needle cluster in this slice, which represents the needle tip. Finally, the needle tip position is calibrated by Eqn. 6.8 to correct for the distortion in the OCT volume.

4.3.3 Hand-eye calibration

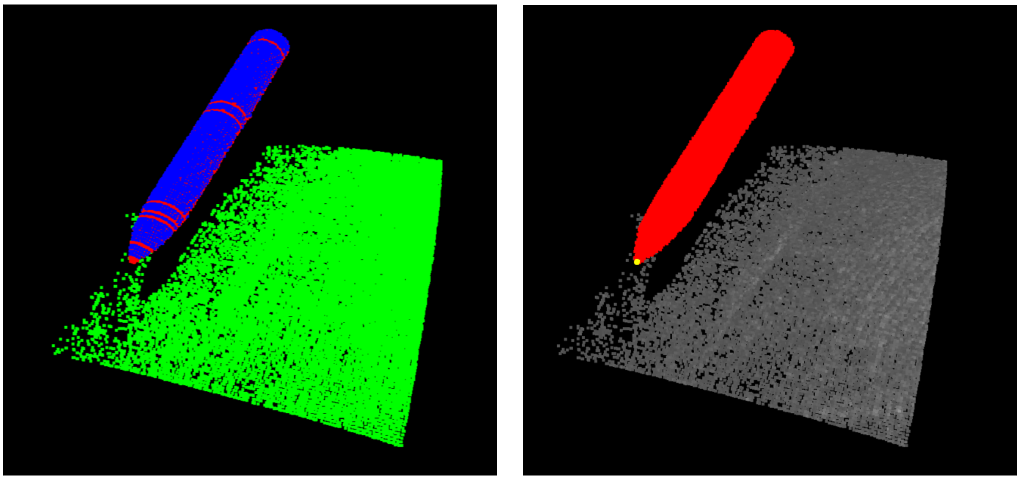
The hand-eye calibration is usually defined by solving a system of homogeneous matrix equations in the form $AX = ZB$, where X and Z are rigid transformations defined by the configuration used. The calibration usually involves eye-in-hand and eye-to-hand configurations. The camera is said to be eye-in-hand when rigidly mounted on the robot end-effector, and it is said to be eye-to-hand when it observes the robot within its work space. The latter is our setup, and since our needle is attached directly to the end effector of the robot, and given the unique kinematics of the robot [76], we can find $A' = AX$. We then move the robot with micro-scale displacements and formulate the system as follows:

$$\begin{bmatrix} A'_i \\ \vdots \\ A'_j \end{bmatrix} = \begin{bmatrix} R & \mid & T \end{bmatrix} \times \begin{bmatrix} B_i \\ \vdots \\ B_j \end{bmatrix} \tag{4.3}$$



(a) 3D OCT volume

(b) Segmented point clusters



(c) Voting results

(d) Needle tip localization

Figure 4.5: (a) The needle with ex-vivo pig eye tissue in 3D OCT volume. (b) The points are clustered and marked with different colors. (c) 2D Detected ellipses from Fig. 4.4(b) are colored in blue and used for voting. (d) Needle segmentation and needle tip localization (the needle tip is marked by a yellow point).

where $A' = \{A'_i, ..., A'_j\}$ is the 3D needle position in the robot coordinate system and $B = \{B_i, ..., B_j\}$ is the corresponding needle position in the camera coordinate system. $R \in SO_3$ and $T \in \mathbb{R}^3$ are the rotation matrix and translation vector that transform the needle positions from the OCT camera coordinate system to the robot coordinate system. The transformation matrix is represented as,

$$\begin{bmatrix} R & | & T \end{bmatrix} = \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix} \tag{4.4}$$

Algorithm 2 Segmentation of needle body in 3D

INPUT: C - Euclidean point clusters $\{(c_1)...(c_n)\}$

OUTPUT: c_i - Cluster with the greatest number of votes

```
procedure SegmentNeedle($C$)
    $maxVote = 0$, $maxIndex = 0$, $vote = 0$
    for each cluster $i$ in $C$ do
        for each point $j$ in $i$ do
            if $b_j == 1$ then
                $vote = vote + 1$
            end if
        end for
        if $vote > maxVote$ then
            $maxVote = vote$
            $maxIndex = i$
        end if
        $vote = 0$
    end for
    Return $C(maxIndex)$
end procedure
```

We apply two different methods to solve Eqn. 4.3: in the first method (SVDT), we calculate the rotation and translation parameters separately based on singular value decomposition [92]; and in the second method (QT), we use unit quaternions [93]. A linear Kalman filter[94] can be used on the needle tip position, for noise reduction, before passing it to either one of the above methods. This filter can predict the future state of a system, so it is very useful for the prediction of the position of the needle.

4.4 Experiment

The experimental setup with an ex-vivo pig eye is depicted in Fig. 4.6. The OPMI LUMERA 700, with integrated RESCAN 700 OCT engine, is fixed on the optical table to reduce the influence of ambient vibration. The *CALLISTO* eye assistance computer system (Carl Zeiss Meditec AG., Germany) is utilized to simultaneously display the microscopic (en-face) image and B-scan images. The iRAM!S eye surgical robot with 5 DoF is mounted on the adjusted bracket. Piezo motors (SmarACT GmbH, Germany), used in the robot, provide 1 µm of accuracy using PID control with integrated incremental optical encoders. A 30G needle, typical for ophthalmic injection with a diameter of 0.31mm, is mounted on the end effector of the robot. The needle tip is fixed in its initial position very close to the ex-vivo pig eye to simulate an ophthalmic operation. The ex-vivo pig eye is fixed with pins into a rubber support. The

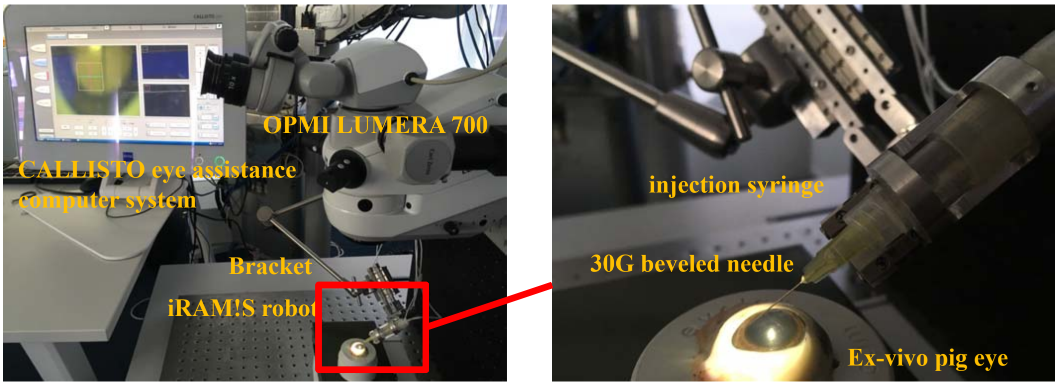


Figure 4.6: The experimental setup for hand-eye calibration with ex-vivo pig eye.

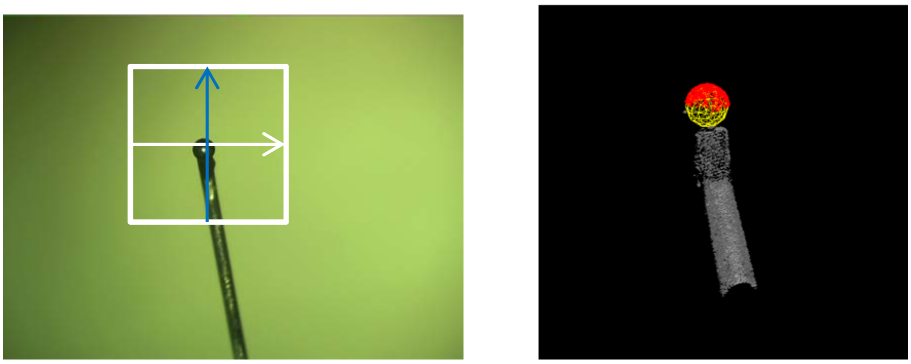
program is executed on the laptop with an Intel Core i7 CPU in Ubuntu 16.04 with Point Cloud Library 1.8 and OpenCV 3.3 in C++.

We are aiming to intraoperatively calibrate the robot and OCT camera with micro-scale movements. In calibration trajectory #1, the needle is moved 10 steps, each being 20 µm, first in the X-, then Z-, then Y- direction. Trajectory #2, is a zigzag, in which the needle moves 5 steps, each being 20 µm, in the X-, Z- and Y- direction. The same trajectory is repeated once again. After each step is finished, the OCT engine is triggered in order to capture an image volume. The parameters for the OCT calibration are listed in Tab. 4.1. These parameter were used to calculate the corrected needle tip position.

Table 4.1: The OCT calibration parameters (mm)

x_c	z_{xc}	y_c	z_{yc}
1.489	151.563	1.068	428.541

To evaluate the accuracy of our calibration method, we compared it with a reference method (marker-based). The reference method is created by attaching a 0.5mm diameter steel ball (Chrome AISI 52100 Grade 10) with a diameter accuracy of 2.5 µm on the end of the needle. The steel ball has a half sphere surface in the OCT volume, which makes it detectable using RANSAC sphere fitting, see Fig. 4.7.



(a) The ball marker in microscopic view (b) The ball marker detected in 3D

Figure 4.7: The steel ball attached to the tool tip in microscopic image (a) and the detected steel ball in 3D OCT volume using RANSAC sphere fitting (b).

4.5 Results

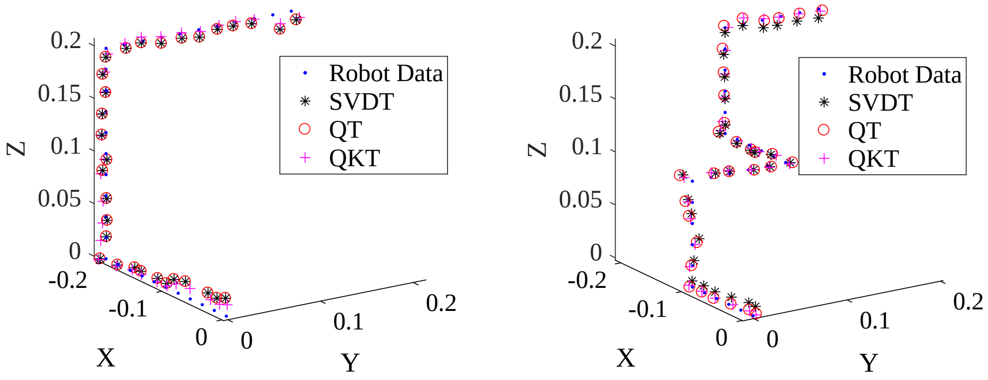
Fig. 4.8 and Fig. 4.9 show the trajectories of the marker based method and our proposed calibration method, respectively. The trajectories of both methods are transformed from the OCT coordinate system to the robot coordinate system with a transformation matrix derived from the methods described in Section 4.3.3: SVDT, QT and quaternion method with linear Kalman filter (QKT). As shown in the figures, all three methods resulted in a trajectory that overall fits well with the trajectory in the robot coordinate system (i.e Robot Data in Fig. 8 and Fig. 9), even though some outliers exist. The average computation time for detecting and transforming one marker ball center and one needle tip position is 1.5s and 0.6s, respectively.

The calibration error for a point is calculated as,

$$e_i = \sqrt{(x_i^R - x_i^{R'})^2 + (y_i^R - y_i^{R'})^2 + (z_i^R - z_i^{R'})^2} \tag{4.5}$$

where (x_i^R, y_i^R, z_i^R) is the (needle tip/marker ball center) position in the robot coordinate system and $(x_i^{R'}, y_i^{R'}, z_i^{R'})$ is the (needle tip/marker ball center) transformed from the OCT coordinate system.

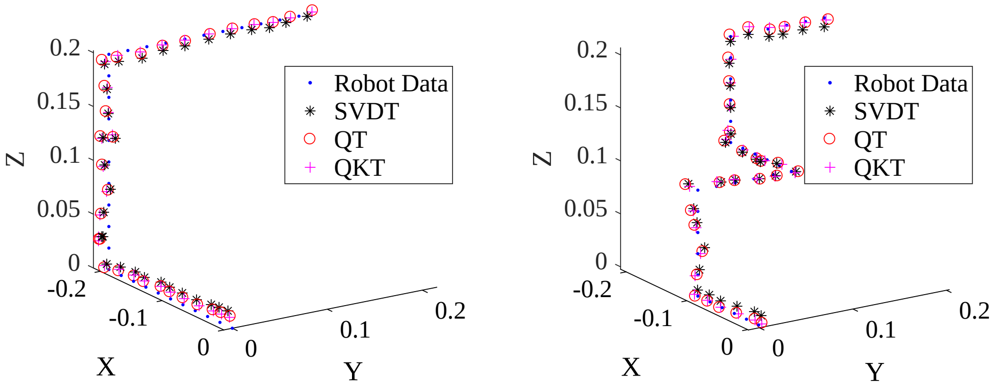
Fig. 4.10 gives the error e with box-and-whisker plots for each method and trajectory. The maximum whisker length is specified as 1.0 times of the interquartile range. Data points beyond the whiskers are treated as outliers. The whiskers are showing the minimum and maximum recorded error while the first (25%) and third (75%) quartile are shown at the bottom and top edge of the box. Bands and red dots represent the median and the mean of the error data,



(a) Trajectory#1 of marker ball center

(b) Trajectory#2 of marker ball center

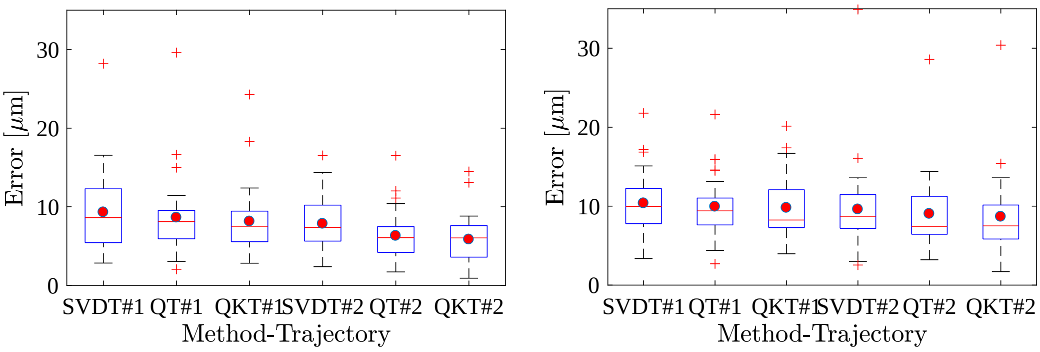
Figure 4.8: The marker ball center trajectory#1 (a) and trajectory#2 (b) in robot coordinate system and the data points transformed from OCT volumes using QT, SVDT, and QKT.



(a) Trajectory#1 of needle tip

(b) Trajectory#2 of needle tip

Figure 4.9: The needle tip trajectory#1 (a) and trajectory#2 (b) in robot coordinate system and the data points transformed from OCT volumes using QT, SVDT, and QKT.

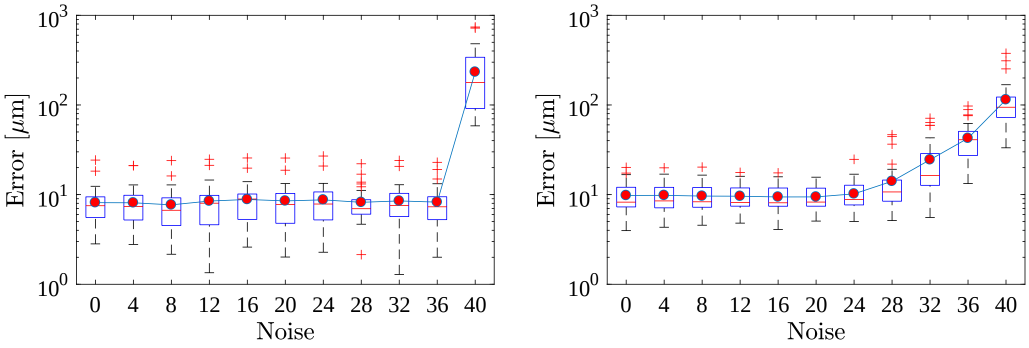


(a) Marker ball calibration error

(b) Needle tip calibration error

Figure 4.10: The calibration error for two trajectories with each method using steel ball(a) and needle tip (b).

respectively. From the figure, we can see that all the methods maintain a low error without a significant difference. For the marker based method, QKT outperforms SVDT & QT slightly, with an average error of 0.8 µm and 1.2 µm for trajectory#1 and #2, respectively. For the needle tip method (i.e our proposed method), QKT outperforms the other two methods slightly with an average error of 0.4 µm and 0.6 µm for trajectory#1 and #2, respectively. It also shows that the performance gap between the marker based method and the needle tip method is very small with an average error of 7.0 µm and 9.2 µm, respectively.



(a) Marker based method

(b) Needle tip method

Figure 4.11: The relationship between the calibration error for each point and standard deviation sigma of the Gaussian noise for marker based method (a) and needle tip based method (b).

In order to test the robustness of the proposed method, we carry out a noise test for the QKT method (i.e. method with the best error), trajectory#1. A zero-mean Gaussian noise is added to each B-scan image from standard deviation $\sigma = 0$ to $\sigma = 40$ by steps of 4. The relationship between the calibration error and the degree of noise added for both methods is shown in Fig. 4.11. We can see that the error for the marker based method (Fig. 4.11(a)) remains stable before $\sigma = 36$, but changes dramatically for $\sigma = 40$. This is due to the fact, that with too much noise added in the image, the sphere fitting cannot work correctly. For the needle tip method (Fig. 4.11(b)), the error remains stable before $\sigma = 24$, and steadily increases with more noise. This is because we are detecting shifted versions of the actual needle, due to the noise attached to the tip. From the above results, we find that our proposed method shows a suitable accuracy performance within a clinical tolerance. The method also showed robustness to noise influence.

4.6 Summary

We proposed a flexible framework for calibrating an ophthalmic robot with an intra-operative MI-OCT. No markers were introduced in the calibration process, which makes our method easier to integrate into the current operation room, by avoiding instrument modification and sterilization. The calibration error of 9.2 µm meets the requirement of ophthalmic operations. Although micron-level displacements in the calibration process may cause safety concerns, the relative position between needle tip and eye tissue surface can be estimated to ensure safety. Accordingly, systems-theoretic safety assessments [95] could be a potential path for completing, refining, adapting, and customizing the whole setup and framework.

In future work, besides the needle tip information, the extra geometric constraints from the needle [90] or even the full geometric information of needle (i.e. CAD model of the needle) could be a potential improvement to increase the accuracy and robustness of the method. Nevertheless, this chapter firstly verify the feasibility and accuracy performance of hand-eye calibration between OCT camera and ophthalmic robot, which motivates future work of path planning for ophthalmic injection and OCT based visual servoing control to achieve an enhanced RAS.

4.7 Related Publication

1. **Mingchuan Zhou***, Mahdi Hamad*, Jakob Weiss, Abouzar Eslami, Kai Huang, Mathias Maier, Chris P. Lohmann, Nassir Navab, Alois Knoll, M. Ali Nasseri. Towards Robotic

Eye Surgery: Marker-free, Online Hand-eye Calibration using Optical Coherence Tomography Images. IEEE Robotics and Automation Letters (also presented at IEEE-IROS 2018), 2018.

Chapter 5

Geometrical Based Approach for Needle Localization under Retina

5.1 Introduction

Subretinal injection is a typical vitreoretinal surgery with a delicate and complex workflow, which needs critical surgical skills and considerations. In this thesis, we will take this as an example of vitreoretinal surgery to explore its possibility of autonomy. Despite its difficulties, the subretinal injection is an important ophthalmic procedure since it enables delivery of the drug directly into the retinal layers to provide more effective treatments [48]. As an example, Fig. 5.1 shows a subretinal injection of Tissue Plasminogen Activator (tPA) to dissolve the clotted blood under the retina. This injection is performed for a subretinal hemorrhage case. Subretinal hemorrhage results from various eye diseases such as Retinal Vessel Occlusion (RVO); Aged-related Macular Degeneration (AMD); and Diabetic Retinopathy (DR). Recently, subretinal injection has also been proposed to deliver the stem cell and gene cargo under the retina for curing AMD which is one of the leading causes of blindness in developed countries (15% incidence rate with people age over 65 years old).

The main challenges of manual subretinal injection are the low position ability of human hand and limited visual feedback from conventional microscope images. The reported hand tremor of an ophthalmic surgeon is 182 µm RMS in amplitude [25] while the acceptable precision for subretinal injection is 25 µm with 250 µm average thickness of the retina. Furthermore, the en-face view from conventional microscope is not able to provide position information of the needle tip under the translucent retina. Consequently, subretinal intervention is a demanding operation requiring a surgeon with excellent physical condition and abundant clinical experience, only a limited number of surgeons around the world fulfill these capabilities.

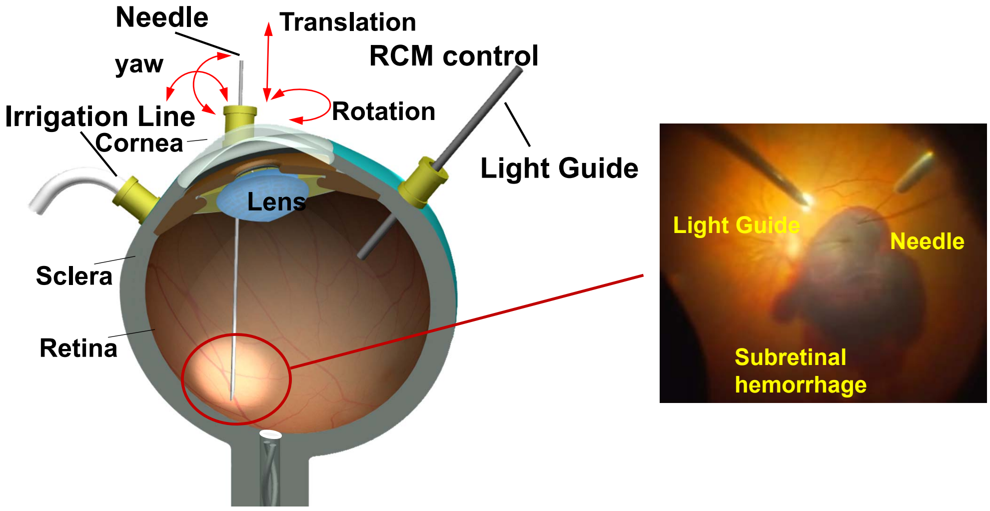


Figure 5.1: The setup for subretinal injection in a subretinal hemorrhage case. A conventional subretinal injection setup. The incision ports are made by trocars at the sclera in a circle 3.5mm away from the limbus to provide entrance for surgical tools: light source, instrument, and irrigation line [1]. The light source is used to illuminate the intended area on the retina, allowing its planar view to be analyzed by surgeon through the ophthalmic microscope. The irrigation line is used for liquid injection to maintain appropriate intraocular pressure.

Recently, robot-assisted surgery (RAS) setups are known as the solution for reducing the work intensity, increasing the surgical outcomes and prolonging the service time of experienced surgeons in ophthalmic surgery [27, 28, 41, 45, 46, 96]. Over the years, these setups are getting closer to clinical trials. In September 2016, surgeons at Oxford's John Radcliffe Hospital performed the world's first robotic eye surgery. The eye surgical robot named Robotic Retinal Dissection Device (R2D2) with 10 µm accuracy was used in the clinical trials for subretinal injection [48]. They used the microscope-integrated Optical Coherence Tomography (MI-OCT) (RESCAN 700, Carl Zeiss Meditec AG., Germany) to enhance the visual feedback during needle insertion under the retina. Although intraoperative MI-OCT setups show several benefits in ophthalmic applications for operations both by human and robot, all the current intraoperative MI-OCT setups are adjusted manually and do not have the automatic needle detection and positioning function which extends the surgery time and also distracts the surgeon's attention. In parts of the procedure where OCT is needed the most, it is distracting the most. For instance, during the needle insertion, the surgeon needs to pay attention to a lot of information e.g. the tool pose and position information, tool shadow information for estimation the distance of the tool tip to retina, and also the B-scan image from X or Z direction for estimation the needle tip depth under the retina (see Fig. 5.2). During the injection, another assistance may need to manually localize the needle tip by adjusting the scan window position.

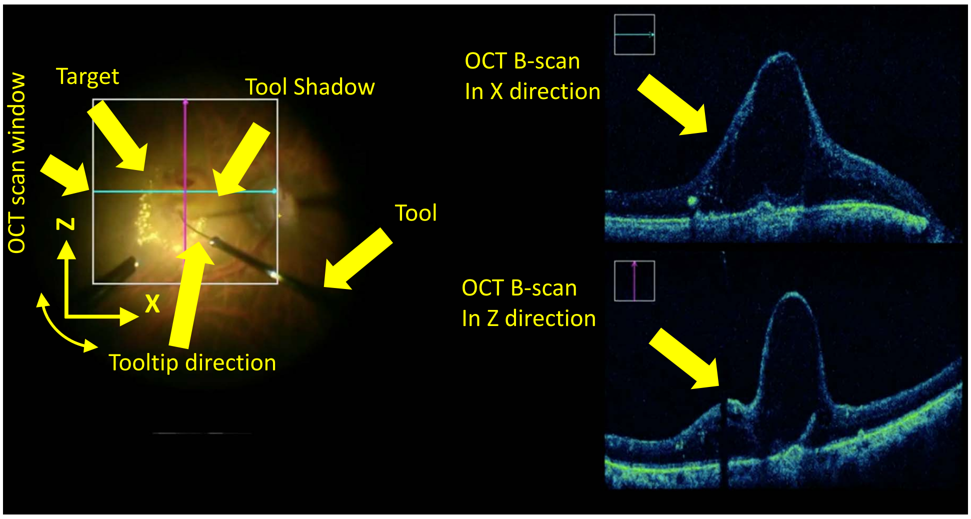
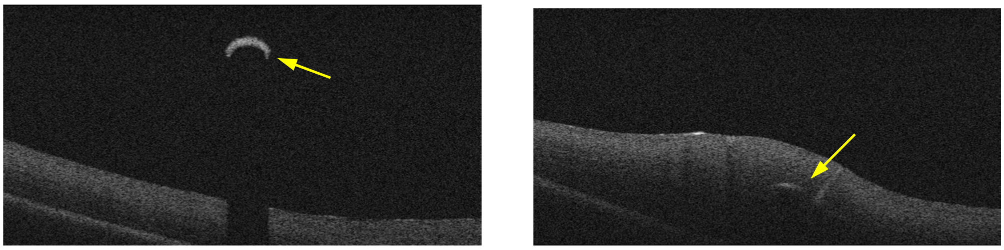


Figure 5.2: The current OCT scan setup. The surgeon needs to pay attention to several information points labeled with the yellow arrow.

OCT is originally used for ophthalmic diagnosis and it is known for having suitable resolution with non-invasive radiation that brings minimum risk of toxicating the ocular tissue. The latest OCT machines has been developed to even give realtime information of the microstructural anatomies and the interactions between surgical instrument and intraocular tissues [30].

Currently there are two methods to intraoperatively deploy OCT data for intra-ocular needle localization. The first method proposed by Cheon et al. [80] is to detect the distance of the target surface from the OCT probe using A-scans. Afterwards, the insertion depth can be calculated considering the distance of the needle tip and OCT probe is known. The limitation of this method is that the needle tip needs to be perpendicular to the target surface. Furthermore, the integration of OCT probe into the needle makes the instrument more complicated in process and maintenance. Instead of integrating OCT probe in surgical instrument, the microscope-integrated intraoperative OCT, which share the same path with microscope to capture volumetric images, is used to estimate the needle tip position. The benefit of this method is to avoid using modified needles but utilizing conventional subretinal cannulas. Due to the optical property of the metallic instruments, the needle above the retina can have a clear upper surface and a shadow on the retina in B-scan, as shown in Fig. 5.3(a). Making use of this feature, Roodaki et al. [74] estimated the distance between surgical tool tip and the underlying tissue. However, when the needle tip is inside the retina, it is hard to reliably distinguish it from the retina because of having similar intensity and unclear shadow, an example of which is shown in Fig. 5.3(b). Therefore, it is difficult to estimate the position and orientation of the needle tip when it is inside the retina.



(a) Needle above the retina

(b) Needle inside the retina

Figure 5.3: The needle fragment and retina in OCT B-scan images. Needle fragment is labeled by arrow.

To address this issue, this work proposes a novel approach to localize the needle tip within the retinal surface using 3D OCT images. The main premise of this chapter is that the needle tip distortion during the insertion is ignored, due to the fact that the interaction force between retinal tissue and the surgical tool is normally below 15 mN [97]. With this assumption, our approach consists of needle and retinal surface reconstruction, Computer Aided Design (CAD) model based calibration, and distance measurement of a needle tip to retinal surface. To precisely model the shape of the needle which is inserted inside the eye, CAD model of the needle is reconstructed and registered to 3D point cloud of the needle obtained from the OCT images when it is still above the retina. Afterwards, the needle tip is predicted following its insertion into the retina. The distance of the needle tip to retina is computed and the retinal surface is also reconstructed to give a better perception for the surgeon. Our approach is evaluated on both tissue phantom and ex-vivo porcine eyes. Evaluation results show that the maximum error in distance measurement of needle tip to retinal surface is within 16.5 µm, and this error achieves the clinical tolerance for subretinal injection. We also verified the feasibility of our proposed method to control the insertion depth in robot-assisted subretinal injection.

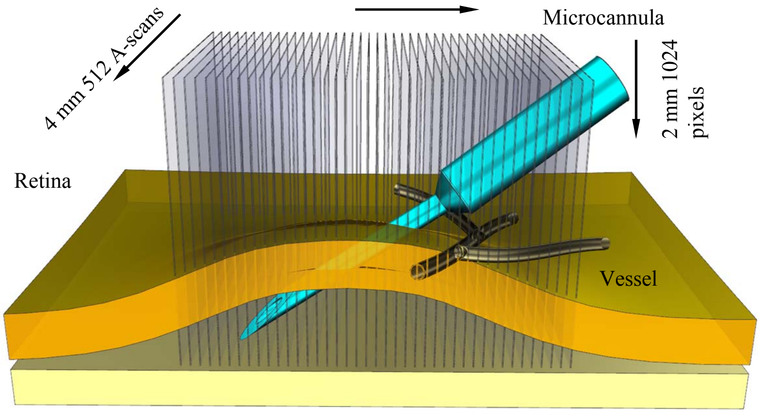
The rest of this chapter is organized as follows. Section 5.2 briefly describes the related work. The background information is presented in Section 5.3. Section 5.4 details the method. Section 5.5 gives the accuracy performance of proposed method and verifies the concept of insertion depth measurement with iRAM!S eye surgical robot [46]. Section 5.6 concludes this chapter with the discussion and the presentation of future work.

5.2 Related Work

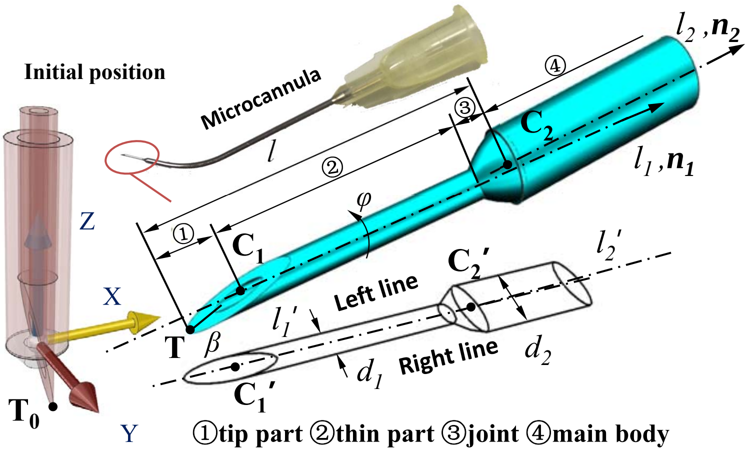
Needle localization have been intensively researched in various 3D medical imaging technologies, e.g. CT scans, fluoroscopy, MRI or ultrasound. However, these imaging technologies can hardly achieve the ideal resolution for subretinal injection. As an example, for MRI-guided interventions with millimeter resolution in breast and prostate biopsies, 18G (Ø=1.27 mm) needle is used, while for subretinal injection, 32G (Ø=0.235 mm) needle requires resolution of submillimeter [58]. Moreover, due to the needle fragments in OCT images it contains more geometrical information which determines the accuracy of needle localization and consequently the localization method for other 3D imaging technologies can not be directly used in OCT images.

Having the benefit of suitable resolution and radiationless imaging mechanism, the OCT imaging modality now is popular not only in the ophthalmologic diagnostics but also in the intraoperative operation to enhance the visual feedback for the surgeon. Edwards et al. [48] performed the first subretinal injection with the robotic assistance under MI-OCT guidance which focuses on the human clinical trials and safety assessment. However, the OCT images could be processed to localize the needle segment instead of continues manually tracking the needle tip, which can further improve the system accuracy and reduce the surgical time without creating distraction for the surgeon. To track and localize the needle in vitreoretinal surgery, the possible method currently is based on either microscope or MI-OCT [98] providing appropriate precision. The single microscope camera was first introduced to track the needle pose in ophthalmic surgery using either color-based or geometry-based features [68]. However, localization of the needle in 3D with only a single view is not practically possible. Probst et al. [64] introduced a deep learning based method to localize the forceps in stereo microscopic images. This method has the advantage of low device cost, however, it faces challenges including the illumination varies and difficulty of detection instrument underlying tissue.

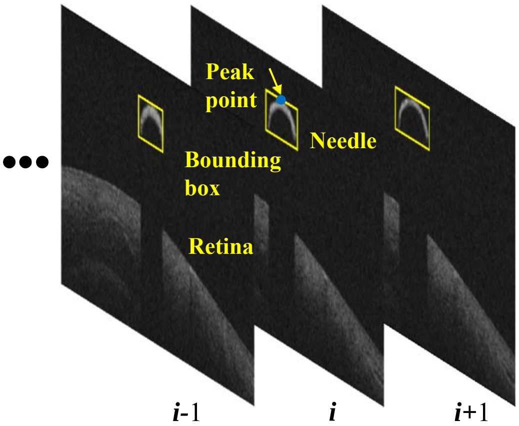
Despite the importance of the application, little work has been done for needle localization in OCT images. Some researchers deployed OCT probe on the instrument for distance estimation between the surgical tool and eye tissue. Song et al. [45] developed a robotic surgical tool with an integrated OCT probe to estimate single dimensional distance between surgical tool and eye tissue for membrane peeling purposes. Yu et al. [70], and Liu et al. [71] deployed OCT probe to assist robotic ophthalmic surgery. They focused on integrating the OCT probe on the ophthalmic microforceps. Their method ensures the instruments constantly keeps its safe



(a) OCT scan setup



(b) The structure of microcannula



(c) The needle fragment in each B-scan

Figure 5.4: (a): OCT scan setup for subretinal injection. OCT scans a cube with a resolution of 128×512×1024 pixels in 4 mm× 4 mm × 2 mm. (b): The structure of the microcannula and its projection. The pose and position of the microcannula in space can be represented by the space transformation of CAD models from initial position. (c): The needle points are enveloped by a bounding box in each B-scan image.

distance from the retina. To estimate the needle insertion depth into the retina, Cheon et al. [80] developed a depth-locking handheld instrument which can control the insertion depth with OCT sensing. The integration of OCT probe into the instrument can simplify the acquisition setup avoiding the necessity of imaging through the cornea and crystalline lens. However, it makes the instrument more complicated in process and maintenance. Moreover, the mechanism of the proposed method need a perpendicular direction of insertion with respect to the target surface which may restrict the flexibility of the operation. Zhou et al. [99] introduced a fully conventional neural network to segment the needle in volumetric OCT images when the needle is above the tissue. Weiss et al. [72] introduced a method to estimate the 5 DoF needle pose for navigation of subretinal injection. Gessert et al. [73] introduced a 3D convolutional neural network to directly estimate the 6D pose of the marker from the OCT volume. They used a marker with obvious geometrical features instead of a surgical instrument to simplify the problem. All of these methods focus on the needle localization above the tissue.

The microscope-integrated OCT shares the same optical path with microscope which has been used in ophthalmic operating theater and proved having a positive impact on the surgical outcome [30]. With such an imaging principle, OCT images are captured without modification of instruments, which is obviously easier to apply in the current operating theaters and surgical workflows. Roodaki et al. [74] estimated the distance between surgical tool and the underlying tissue using geometrical information of instrument. This method is designed for those vitreoretinal operations that the instrument is operating above the retina (e.g. ILM peeling). In order to overcome the challenges of estimating the insertion depth for subretinal injection, this chapter propose a precision needle tip localization method using 3D OCT images. The microscope-integrated OCT is adopted to capture the information thus without modification of conventional instruments.

5.3 Background

This section introduces the background of the proposed framework which includes the OCT image acquisition, the geometrical structure of a commercially available microcannule for subretinal injection and the image preprocessing.

5.3.1 OCT Image Acquisition

Our setup contains an OPMI LUMERA 700 with integrated RESCAN 700 intraoperative OCT engine, and *CALLISTO* eye assistance computer system (Carl Zeiss Meditec AG.). The

CALLISTO eye assistance computer system is utilized to display the microscopic image and two cross OCT images at the same time, and a foot panel is connected to the system for helping to relocate the OCT scan area. All the images are acquired on a tissue phantom and during ex-vivo experiments with porcine eyes using conventional subretinal cannulas. The soft cheese is used as a simple tissue phantom because it has the similar OCT imaging result as ocular tissues and it is readily available. In order to acquire sufficient visual information for needle reconstruction and calibration, we reprogrammed the original iOCT device and the movement of its internal mirror galvanometers to obtain the highest resolution with 128 B-scans, each with 512 A-scans (4 mm×4 mm) [100]. Each A-scan has 1024 pixels for 2 mm depth information in tissue, see Fig. 5.4(a).

5.3.2 Geometrical Structure of the Needle

The conventional microcannula used for subretinal surgery is shown in Fig. 5.4(b). The microcannula is characterized to four parts: the tip part, the thin part, the joint and the main body. Its parameters consist of 20G (Ø=0.908 mm) main body, 32G (Ø=0.235mm) thin part with length of 3 mm and 15° beveled angle (Eagle Labs Inc.). The thin part and the main body are considered as the shape of a cylinder. The tip part and the thin part are represented by the bevel center point $\mathbf{C_1}$, space vector $\mathbf{n_1}$, diameter d_1, and bevel angle β. The main body is represented by body start point $\mathbf{C_2}$, space vector $\mathbf{n_2}$ and diameter d_2. Assume l_1 and l_2 are the axis of thin part and the center axis of the main body respectively, and l_1' and l_2' are their corresponding projection on XOY plane. $\mathbf{C_1'}$ and $\mathbf{C_2'}$ are the projection of $\mathbf{C_1}$ and $\mathbf{C_2}$, respectively. The left line and right line are the borderline of the thin part projection on XOY plane. The pose and position of microcannula in space can be represented by a CAD model, see Fig. 5.4(b).

Due to the manufacturing defects and unexpected forces causing deformations during needle transportation and sterilization, we cannot precisely model the needle according to the specifications provided by the manufacturer. For example, we cannot simply assume that the vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are the same. Only a 5 degree difference between these vectors can cause an error of 250 µm in the measurement of the tip. Hence, the parameters for the thin part and the main body need to be represented separately and calibrated to achieve the desired precision for ophthalmic surgery. Following the proper needle calibration, the tip of microcannula $\mathbf{T}$ is determined by transforming CAD models (see Fig. 5.4(b) translucent CAD models) from the initial needle tip position $\mathbf{T_0}$. In the inial position, C_1 and C_2 are at the origin point of coordinate system, meanwhile l_1 and l_2 are coincide with z axis. Therefore, we can calculate $\mathbf{T_0}$ as follows,

$$\mathbf{T_0} = (\frac{d_1}{2}, 0, \frac{d_1}{2\tan(\beta)}) \tag{5.1}$$

5.3.3 OCT Image Preprocessing

The main goal of the image preprocessing is to segment the needle fragment from tissue in each B-scan image. Based on the needle shadow principle from our previous work [99, 101], the needle and retinal surface can be segmented with dissolution of noise influence in each B-scan image when the needle is still above the retina. The position of the needle points can be calculated from the index of OCT B-scan and pixels position in current OCT B-scan image. See Fig. 5.4(c), in every B-scan, the needle points can be enveloped by the bounding box $\mathbf{B}_i = (Bx_i, By_i, Bz_i, w_i, h_i)$, $(i = 1, ..., n)$ where $(Bx_i,\ By_i,\ Bz_i)$ is the position of the top left corner of bounding box, w_i and h_i are the width and height of the box respectively, and n is the total number of bounding boxes.

5.4 Method

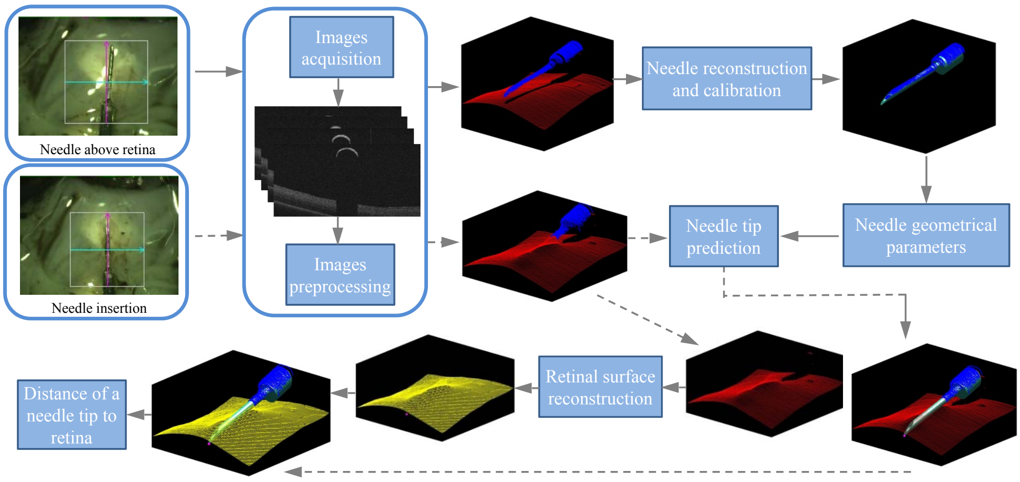


Figure 5.5: The framework of the method.

In this section, we present our approach, an overview of which is depicted in Fig. 5.5. Based on the OCT image preprocessing mentioned in the previous section, the needle point cloud and retinal surface point cloud can be segmented before the needle injects into the retina. The point cloud of the needle is then classified into four parts corresponding to the structure of microcannula. With these parts of point cloud, the needle CAD model is calibrated preciously with its geometrical parameters when the needle is above the retina. During the needle insertion, the CAD model of the needle is registered to the needle point cloud which remains above the surface. The tip of the needle is predicted using geometrical space transformation from its initial position. Simultaneously, the retinal surface is reconstructed to give the surgeon or the automatic system a better perception and finally the distance between the needle tip and the retinal surface is calculated for surgical reference. In the following, details of different components are discussed.

5.4.1 Needle Reconstruction and Calibration

Needle reconstruction and calibration is designed and processed when the needle is above the retina. The main geometrical parameters of needle are calculated to correct the needle CAD model for high precision intention. With the OCT image preprocessing, we can calculate the position of the needle and the surface points together with corresponding bounding box Bx_i. The variation of w_i with Bx_i reflects the geometrical information of the microcannula, where w_i increases along the tip and the joint, and remains constant along the thin part and the main body. Hence, $\mathbf{B}_i$ is divided in four segments with three break points. The bounding boxes are filtered as outliers if they have a large deviation in peak point position or w_i. Thereafter, piecewise linear regression with least square fitting is applied to find the break points as follows,

$$f(i) = \begin{cases} k_1 Bx_i + c_1, i \in [1, o] \\ c_2, i \in (o, p]; \\ k_2 Bx_i + c_3, i \in (p, q] \\ c_4, i \in (q, n] \end{cases} . \tag{5.2}$$

$$\{\bar{o}, \bar{p}, \bar{q}\} = \underset{o,p,q}{argmin} \frac{1}{n} \sum_{i=1}^{n} (f(i) - w_i)^2 \tag{5.3}$$

Eq. 6.8 gives the shape pattern of the needle projection, where k_1, k_2, c_1, c_2, c_3, and c_4 are the parameters after fitting and o, p, and q are the indices to divide the points into groups. Eq. 5.3 minimizes the algebraic distance between w_i and corresponding fitting value $f(i)$, and gives the optimal indices $\bar{o}$, $\bar{p}$ and $\bar{q}$. Subsequently, break points can be calculated from solving the equations from adjacent lines.

The points in group of indices $i \in (\bar{q}, \bar{n}]$ enveloped by the bounding box in each B-scan can be fitted by an ellipse. Through the ellipse center points, l_2 can be obtained with RANSAC fitting with equation $\frac{x-x_0}{n_x} = \frac{y-y_0}{n_y} = \frac{z-z_0}{n_z}$, where (x_0, y_0, z_0) is a point on l_2. The error threshold for RANSAC that defines the maximum deviation is set as 0.05 mm. Assuming the equation of

l_2': $y = k_0x + b_0$, the left side line and right side line for the edges of the main body can be represented as $y = k_0x + b_1$ and $y = k_0x + 2b_0 - b_1$, respectively. Both two lines are obtained by fitting data sets (Bx_i, By_i) and $(Bx_i, By_i + w_i)$, $i \in (\bar{p}, \bar{q}]$ respectively. The diameter of the main part d_2 is calculated by the distance between the left and right lines.

Introducing the breakpoints into l_2', we obtain point $\mathbf{D_2'}$ $(d_{2x}, d_{2y}, 0)$, see Fig. 5.6(a). $\mathbf{C_2'}$ has a certain shift to $\mathbf{D_2'}$ that is decided by the slope of l_2'. We can calculate the three parameters to represent the main body part as,

$$\begin{cases} \mathbf{n}_1 = (n_x, n_y, n_z) \\ d_1 = \dfrac{2|b_1 - b_0|}{\sqrt{1 + k_0^2}} \\ \mathbf{C_1} = (d_{2x} - \triangle x, d_{2y} - \triangle y, \dfrac{n_z}{n_x}(d_{2x} - \triangle x - x_0)) \end{cases} \tag{5.4}$$

where $\triangle x = \sqrt{\frac{k_0^2 d_2^2}{4(1+k_0^2)}}$ and $\triangle y = \sqrt{\frac{k_0^4 d_2^2}{4(1+k_0^2)}}$ are the shift distance from $\mathbf{D_2'}$ to $\mathbf{C_2'}$ in X and Y axis, respectively.

Similarly, we obtain l_1, $\mathbf{n}_1$, and d_1 which determine the pose of the thin part. Due to the fact that the needle bevel is upwards, we can reconstruct the needle tip part with the intersection of the bevel plane and thin part. The bevel plane is estimated by the RANSAC plane fitting [102] with the point cloud of tip part, see Fig. 5.6(a). The error threshold for RANSAC plane fitting that defines the maximum deviation is set as 0.01 mm. The bevel center point $\mathbf{C_1}$ can be calculated as the intersection point of the bevel plane Γ and the l_1. The tip part rotation φ and bevel angel β can be calculated as,

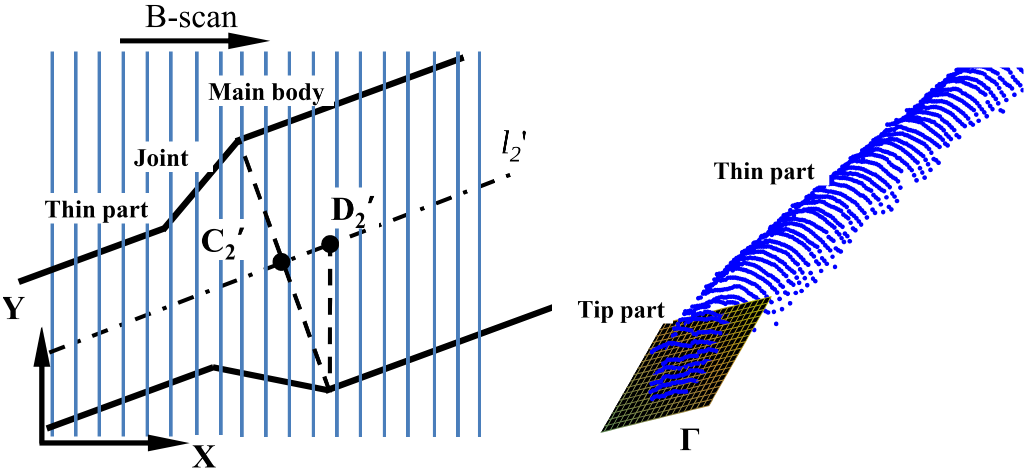
$$\begin{cases} \varphi = \arcsin(||\mathbf{n}_\Gamma \times \hat{\mathbf{n}}_\mathbf{z}||) \\ \beta = \arccos(||\mathbf{n}_\Gamma \times \mathbf{n_1}||) - \dfrac{\pi}{2} \end{cases} \tag{5.5}$$

where $\mathbf{n}_\Gamma$ is the unit normal of Γ in positive z direction and $\hat{\mathbf{n}}_\mathbf{z}$ is the unit normal in positive z direction. The length of the thin part l is calculated by the distance between point $\mathbf{C}_2$ and $\mathbf{T}$.

5.4.2 Needle Tip Prediction

Based on aforementioned progress, the structure of the microcannula can be determined. Moreover, these configurations are used to calibrate and reconstruct the needle CAD model, see Fig. 5.7. When the needle is above the retina, the needle-tip $\mathbf{T}$ can be calculated by geometrical space transformation of needle-tip $\mathbf{T}_0$ from initial position.

$$\mathbf{T} = \mathbf{T_0} R_z(\varphi) R(\hat{\mathbf{n}}_\mathbf{z}, \mathbf{n_1})\, T(\mathbf{C_1}) \tag{5.6}$$



(a) The projection of needle

(b) Needle bevel reconstruction

Figure 5.6: The diagram for the calculation of C_2' in XOY plane (a) and the needle bevel reconstruction (b).

where R_z is the rotation matrix around Z axis with angle φ, $R(\mathbf{\hat{n}_z}, \mathbf{n_1})$ is the rotation matrix from $\mathbf{\hat{n}_z}$ to $\mathbf{n_1}$, $T(\mathbf{C_1})$ is the translation matrix.

During the injection, the tip part will be inside the tissue but the main body will still remain above the retina. In this situation, $\mathbf{C}_2$ and $\mathbf{n}_2$ are used to estimate the position and pose of the calibrated needle CAD model, see Fig. 5.8(a) and 5.8(b). The renewed position of the needle tip $\mathbf{T}^*$ can be updated by,

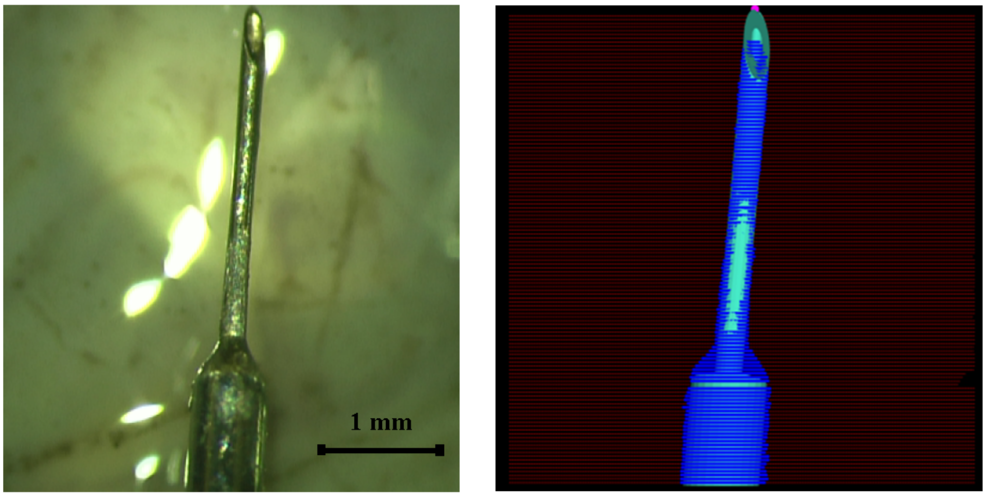
$$\mathbf{T}^* = \mathbf{T}R(\mathbf{n_2}, \mathbf{n_2^*})\,T(\mathbf{C_2^*} - \mathbf{C_2}) \tag{5.7}$$

where $\mathbf{n_2^*}$ and $\mathbf{C_2^*}$ are the corresponding updated value for $\mathbf{n_2}$ and $\mathbf{C_2}$. $R(\mathbf{n_2}, \mathbf{n_2^*})$ is the rotation matrix from $\mathbf{n_2}$ to $\mathbf{n_2^*}$. $T(\mathbf{C_2^*} - \mathbf{C_2})$ is the translation matrix.

5.4.3 Retinal Surface Reconstruction and Distance of the Needle Tip to Surface

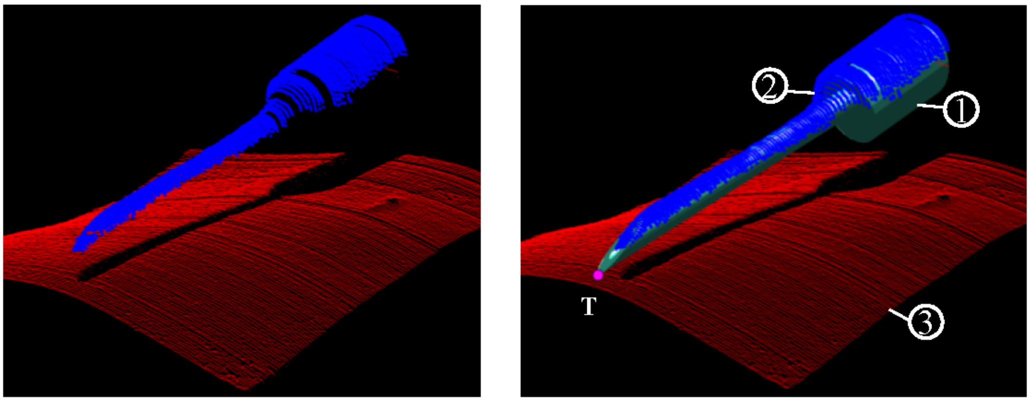
The retinal point cloud is sampled uniformly along XY grid by a factor decided by considering surface quality and processing time. The outliers in the point cloud are filtered by a statistical technique used in [103]. To approximate the missing points, which are lost due to the needle shadow and outliers, natural neighbor interpolation method is used.

The surface reconstruction is performed by a similar method used by Hoppe et al [104]. Given a set of points $\mathbf{X} = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}$ representing a surface M, a tangent plane is associated with each point $\mathbf{x_i}$, represented by a center $\mathbf{o_i}$ and a unit normal $\mathbf{\hat{n}_i}$. A neighborhood $Nbhd(\mathbf{x_i})$



(a) Microscopic en face view

(b) OCT en face view



(c) OCT oblique view

(d) Needle in OCT

Figure 5.7: The result of needle reconstruction and calibration. The needle in microscopic en face view (a) and the reconstructed and calibrated needle CAD model in OCT en-face view (b) are presented. The oblique view of OCT point cloud (c) and needle tip **T** prediction (marked by magenta piont) shown in (d) can give more information to guide the operation. ① represents CAD model of needle, ② represents the needle points cloud, and ③ presents the retinal surface points cloud.

is defined using k points in $\mathbf{X}$ nearest to $\mathbf{x_i}$, where $\mathbf{o_i}$ and $\hat{\mathbf{n}}_\mathbf{i}$ are calculated as the centroid, and the unit eigenvector associated with the lowest eigenvalue in the covariance matrix of $Nbhd(\mathbf{x_i})$ respectively. Here, k=20 is used for the estimation of tangent plane. Then, a signed function d is defined to estimate the distance between a point $\mathbf{p}$ and the surface M as,

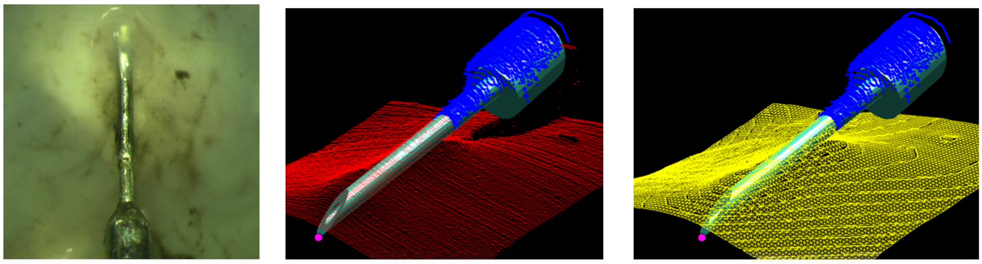
$$d(\mathbf{p}) = (\mathbf{p} - \mathbf{o_i}) \cdot \hat{\mathbf{n}}_\mathbf{i} \tag{5.8}$$

The zero set of d will be the estimate of the surface. The extraction the iso-surface is done by the contouring algorithm, marching cubes [105]. The reconstructed retinal grid and the needle inside retina are shown in Fig. 5.8(c).

The distance of needle tip $\mathbf{T}$ from the retinal surface in z-axis can be calculated using d with modified parameters,

$$d(\mathbf{T}) = (\mathbf{T} - \mathbf{o_i^{XOY}}) \cdot \hat{\mathbf{n}}_\mathbf{z} \tag{5.9}$$

where center $\mathbf{o_i^{XOY}}$ is calculated using k nearest points to $\mathbf{T}$ in the projection of XOY plane.



(a) Microscopic en face view (b) Before retinal surface reconstruction (c) Retinal surface reconstruction

Figure 5.8: The needle is injected inside retina in microscopic image (a). Oblique view of the point cloud for OCT cube (b). Needle tip prediction based on CAD model and reconstructed retinal surface (c).

5.5 Experiments and Results

Our experimental setup is depicted in Fig. 5.9. The iRAM!S eye surgical robot is mounted on an adjustment bracket and the motion of robot is controlled by the robot controller. The OPMI LUMERA 700 with integrated RESCAN 700 intraoperative OCT engine is fixed on the optical table to reduce the influence of ambient vibration. The foot panel is used to relocate the OCT scan area. The OCT is set to operate with the maximum available scanning speed (27000 A-scans per second), in cube resolution of 128×512×1024 pixels. The implementation

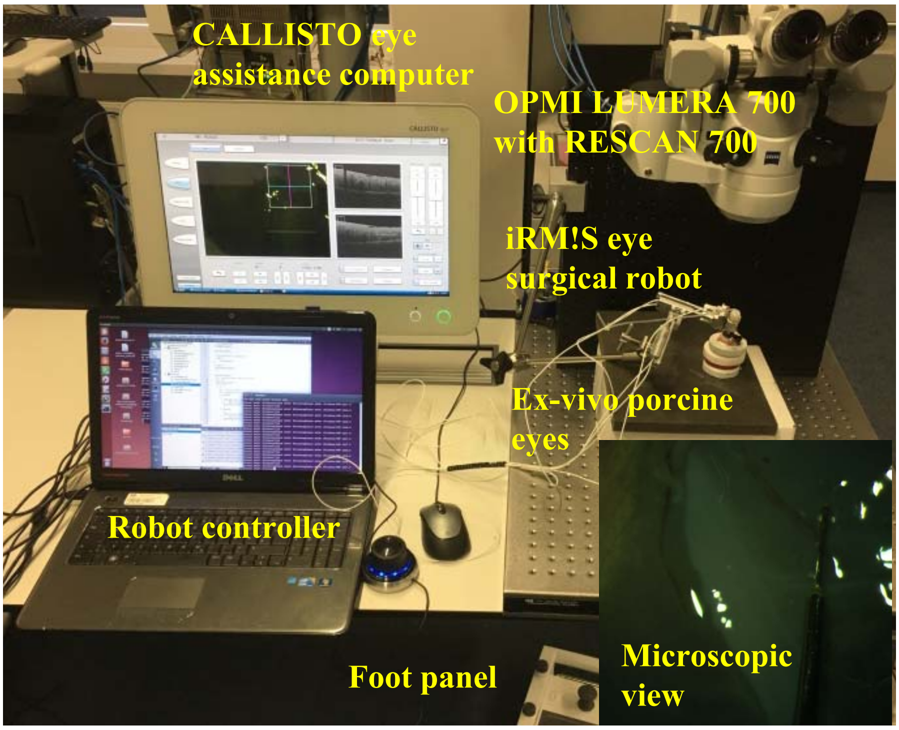


Figure 5.9: The experimental setup.

of our approach is executed on the *CALLISTO* eye assistance computer system with an Intel Core i7 CPU and a NVIDIA GeForce GTX 980i GPU with average speed of 0.4 s for each cube with parallel programming. We performed two types of experiments: 1) test-retest reliability and accuracy performance of proposed method, and 2) insertion depth tracking with the robot. Both experiments are conducted with both tissue phantom and cadaver porcine eyes.

5.5.1 Test-retest Reliability and Accuracy Performance

In this section, we conducted test-retest reliability experiments to verify the reconstruction and calibration results, and the accuracy performance of retinal surface to needle tip distance measurement using a manual micromanipulator. We estimate the variation of measurements referring to the intrinsic dimensions, the diameter of the thin part d_1, the length of the thin part l, and the diameter of the main body d_2. The actual values for these parameters are obtained by a digital caliper with a resolution of 10 μm after 10 times averaging of repeated measurements. We capture 80 image cubes with various injection postures above the retinal surface in ex-vivo porcine eyes.

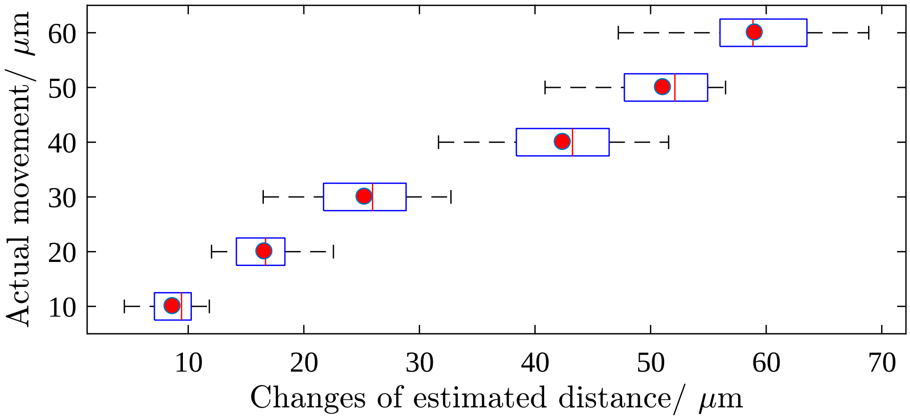
Table 5.1: Descriptive statistics for three intrinsic dimensions (in micron).

	Mean	Actual	AE	ME	RSME	0.95CI
d_1	234.0	230.0	4.0	6.8	4.2	[231.5-236.4]
d_2	671.8	670	1.7	5.7	3.0	[666.5-675.6]
l	3204.6	3200	4.6	45.7	21.5	[3152.4-3239.4]
$\bar{d_1}^5$	234.1	230.0	4.1	5.2	4.1	[233.1-234.9]
$\bar{d_2}^5$	671.6	670	1.6	4.7	1.8	[670.1-673.1]
$\bar{l}^5$	3204.8	3200	4.8	21.2	7.7	[3190.3-3216.1]

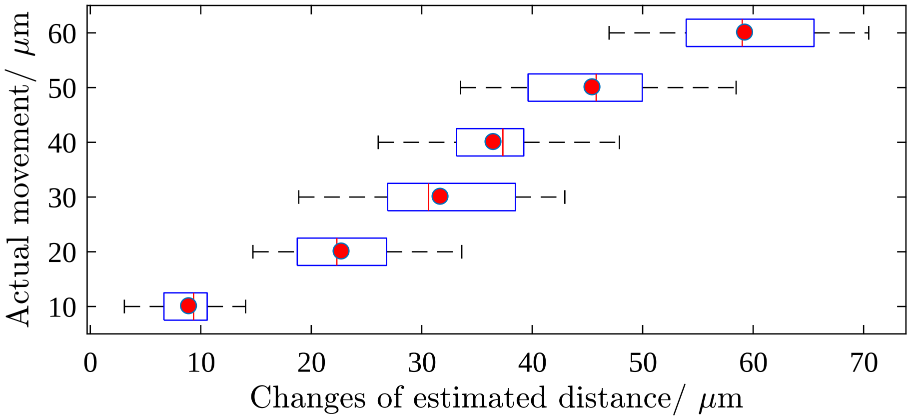
We randomly select 5 estimated values and average these values from each dimension for 80 times to obtain $\bar{d_1}^5$, $\bar{d_2}^5$ and $\bar{l}^5$, respectively. Several statistical metrics are used to evaluate deviation of these measurements shown in Tab. 5.1 including: the average error of estimated value to the ground truth (AE), the error of maximum estimated value and actual value (ME), the root mean square error (RSME), and the 95% of estimated confidence interval (0.95CI). All the mean errors are within 5 μm which indicate that the proposed method has the capability to achieve accurate reconstruction and calibration results. Due to the fact that the needle is placed along the B-scan direction with scan resolution of 31.3 μm and the number of tip part points is relatively less than the thin part and main body part, l_1 has a relatively larger maximum error and RSME than the other two parameters. This could be improved by averaging parameters with several OCT cubes to achieve a higher accuracy. The 0.95CI for $\bar{l}^5$ can be significantly narrowed down compared to single evaluation of l.

The accuracy of the distance between retinal surface and the needle tip is tested with a 3-axes HS-6 manual micromanipulator (Märzhäuser Wetzlar GmbH) with 5 μm resolution on both tissue phantom and ex-vivo porcine eyes. The needle is fixed on a micromanipulator and the micromanipulator is fixed under the microscope on the optical table. The needle tip is moved up and down to change the distance of needle tip to the retinal surface. Before the movements, five OCT cubes are captured to calibrate and reconstruct the needle. Afterwards, in total 16 measurements are captured and modeled to predict needle tip for each movement. The results are presented in Fig. 5.10.

From the figures, one can find following observations: Firstly, the maximal error for all measurements is 16.5 μm and the maximal error for the mean estimated distance is 4.7 μm. The 0.95CI for overall movement error is [-11.9, 9.0] and [-14.0, 10.9] for tissue phantom and ex-vivo porcine eye, respectively. Secondly, the results of the estimated distance vary due to the vibration from the environment but the mean estimated distance value is consistent with the



(a) Tissue phantom



(b) Ex-vivo porcine eye

Figure 5.10: The actual movement of the needle that is mounted on a micromanuipulator and the change of estimated distance for the needle tip obtained from OCT cube. The whiskers show the minimum and maximum recorded change of the distance while the first and third quartile show the start and the end of the box. Band and red dot represent median and mean of the recorded changes respectively.

actual input movement. Thirdly, the variance of distance measurement with the porcine eyes are larger than with tissue phantom. This is likely due to the fact that retinal surface is more curve and bumpy with more intention for deformation which makes the distance measurements more sensitive to the movement of the needle tip. In conclusion, our approach can cope with the ideal 25 µm error tolerance as the average thickness of the retina is around 250 µm.

5.5.2 Primary Experiment of Injection Depth Tracking

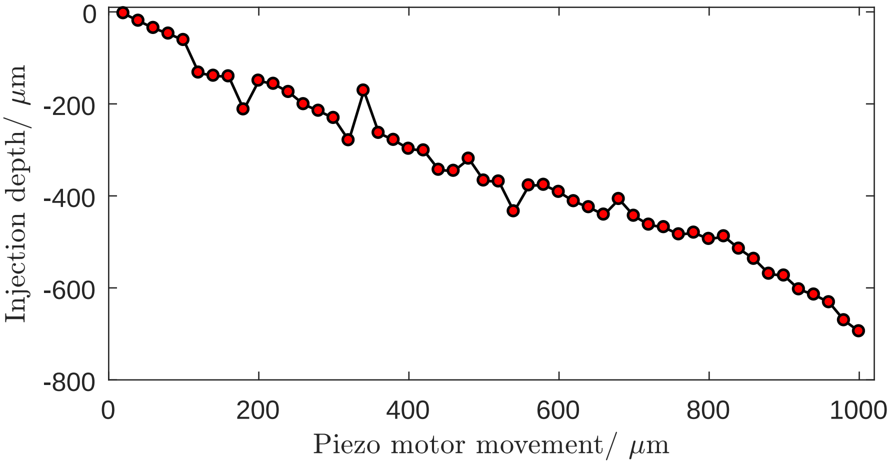


Figure 5.11: The needle insertion tracking using iRM!S surgical robot on the ex-vivo porcine eyes.

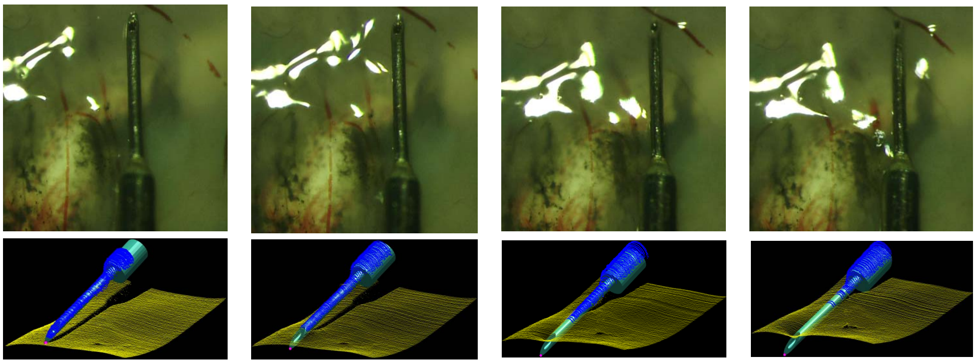


Figure 5.12: The needle insertion using iRAM!S surgical robot on the ex-vivo porcine eyes.

The iRAM!S eye surgical robot with 5 DoF is utilized to perform the primary injection experiments (see Fig. 5.9 for experimental setup). Piezo motors (SmarACT GmbH) provide 1 µm accuracy by using PID controller with integrated incremental optical encoders. The robot is adjusted manually by the adjustment bracket to make the needle tip approach to the retina. During the injection, only the motor that is holding the syringe is enabled to control the needle positioning. At the beginning we captured five cubes for needle reconstruction and calibration. Thereafter, the motor is programmed to move exact 20 µm for each step and the OCT cube is captured after each movement. Fig. 5.11 shows the typical tracking for the distance of needle tip to retinal surface. From this image it can be observed that there is an approximately liner increase between motor movement and insertion depth. The interaction of the needle and retinal tissue is very complicated that makes stochastic deformation of the retina. The fluctuation of the insertion depth during tracking may be as a result of distance measurement error and retina deformation.

Fig. 5.12 shows a few sequences of the insertion progress. The reconstruction results appropriately match with the needle tip status in en face image from the microscope. In the microscopic image, the needle tip position is significantly hard to estimate, while with our proposed method the needle tip is visualized and also the tip position and injection depth can be precisely calculated.

5.6 Summary

This chapter presents our approach to localize the needle tip inside the retina for the first time using OCT images. The presented evaluations show high accuracy distance measurement of the needle tip to retina which has mean error of 4.7 µm (maximum error 16.5 µm) for both tissue phantom and ex-vivo porcine eyes. The experiments further show the feasibility of our approach for precise needle tip localization with OCT images, given the ideal 25 µm error tolerance for subretinal injection. However, this approach is based on the 32 G needle which we have an assumption that the needle is not going to deform during insertion. This may not be realistic because the needle used in the clinic room is decreasing its diameter for reducing the trauma on the retina. In the next chapter, we will propose another approach that can potentially solve this problem.

5.7 Related Publication

1. **Mingchuan Zhou**, Kai Huang, Abouzar Eslami, Hessam Roodaki, Daniel Zapp, Mathias Maier, Chris P. Lohmann, Alois Knoll, and M. Ali Nasseri. Precision needle tip localization using optical coherence tomography images for subretinal injection. In 2018 IEEE

International Conference on Robotics and Automation(ICRA), May 2018.

Chapter 6

Deep Learning based Approach for Needle Localization under Retina

6.1 Introduction

The previous chapter proposed an algorithm to predict the needle tip under retina using the geometrical information of the needle. The main premise is that the needle does not have large deformation during the insertion. However, the thin needle can bring better clinic performance for reducing injection trauma to the retina. The latest clinical subretinal microcannula used in clinic operation is 40-41 gauge (Ø=0.16-0.15mm) while previously available cannulas were only 32G (Ø=0.235mm). Furthermore, soft material for the needle tip part e.g. teflon is used instead of metal. Therefore, the premise does not meet very well with the thinner needle used in the current clinical operation which introduces significant deformation during puncture and retinal insertion.

In order to have a more universal framework with the variety of needles compatibility, and to improve the efficiency of the robot-assisted subretinal injection and furthermore, to give the surgeon enhanced information about the needle tip under the retina, we propose a novel framework to robustly localize needle segment under the retina using volumetric OCT images. The contributions of this chapter are in two aspects: first and foremost, we take advantage of the needle information above the retina in B-scan images to coarsely predict the needle position under the retina and crop the original high-resolution image into a small region of interest (ROI). A convolutional neural network (CNN) is adopted to train on the cropped ROI images with the benefit of foreground and background balance, moreover, cropped images significantly reduce noise and reflection from the original B-scan images to enable facilitating very high accuracy results. Secondly, we create original and cropped datasets with 3811 OCT B-scan images in several ex-vivo pig eyes. The original image trained model and cropped image trained model are compared and evaluated. The results show that the cropped image trained model can localize the needle accurately with a confidence of 99.2%.

The remainder of the chapter is organized as follows: in the next section, we briefly present related work. The proposed method is described in Section 6.3. In Section 6.4, the performance of the proposed method is evaluated and discussed. Finally, Section 6.5 concludes this chapter and presents the future work.

6.2 Related work

Since the instrument pose estimation in ophthalmic surgery has been introduced in the previous chapter, here we analyze the related work in the state of art for object detection.

In order to have a versatile method to localize the needle under the retina to solve the problem from geometrical based method, we propose detection and localization of the needle directly in B-scan images from OCT volume. Therefore, the problem is transformed into the object detection task. The object detection is rapidly developing specifically because of the advancements in deep learning technology and improvement in computational power. The two-stage detectors and one-stage detectors are the dominant object detectors in modern objection detection [106]. Recently, Lin et al. proposed a one-stage detector RetinaNet [107] to surpass the two-stage detector in accuracy and remain the speed advantage of one-stage detector with focal loss principle. Based on these advancements, we propose a robust needle localization framework for subretinal injection. The prosperity of our method is having the capability to localize the deformed needle under retina without geometrical feature which can be more versatile and feasible in the current and future clinic application.

6.3 Method

The overall framework contains two parts: the first part is the robust ROI crop; and the second part is the deep learning based needle localization as shown in Fig. 6.1. The volumetric images are captured by the MI-OCT by scanning the injection area. The original B-scan images are processed with the ROI crop method to significantly reduce the size of the image from 512×1024 to 120×240. Furthermore, cropped images are fed into the RetinaNet for needle

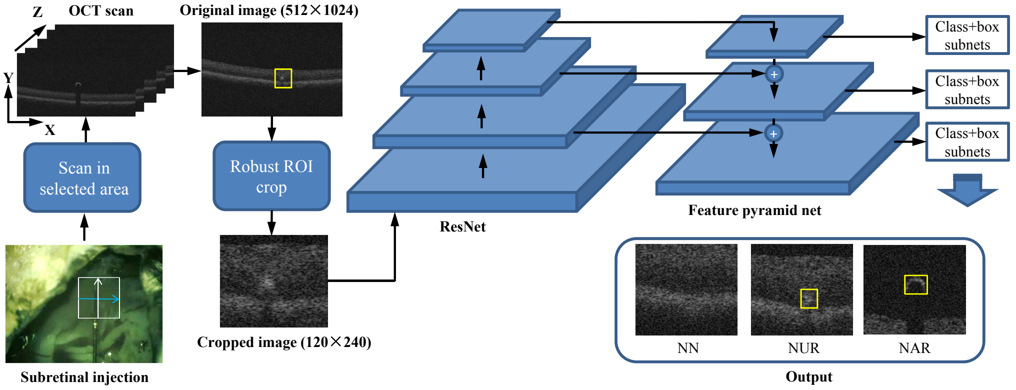


Figure 6.1: The proposed framework. ResNet, feature pyramid net, and class+box subnets constitute RetinaNet. The output has three categories, no needle (NN), needle above retina (NAR) which means needle has no interaction with retinal tissue, and needle under retina (NUR) which means needle has interaction with retinal tissue.

segment detection and localization.

6.3.1 Robust ROI Crop

The microscope used in this work is an OPMI LUMERA 700 with integrated RESCAN 700 OCT. The OCT engine has a wavelength of 840 nm and a scanning speed of 54000 A-scans per second. Due to the fact that the intended surgical area for subretinal injection is usually very small, we set the OCT engine to obtain the scan area 3 mm×3 mm×2 mm with the highest resolution of 128 B-scans, each with 512 A-scans. The benefit of high resolution is that it can reserve as much information as possible. Each of the B-scan images can potentially contain a needle segment. The needle bevel part cross-section is a spot of pixels while the needle body part cross-section is a half ellipse. When the needle is placed above the retina, the needle segment is isolated from the tissue and most of the time it creates a shadow on the retina. When the needle is inserted into the retina, some of the needle segments are under the surface layer reducing the clarity of the needle shape and making its position hard to be distinguished in B-scan image, especially with the presence of reflection and noise. Fig. 6.2 shows some examples of B-scan image with needle or without needle segment. The following points are taken into considerations in the ROI crop algorithm design: (1) the needle is a continuous object in the OCT volumetric image. (2) the very thin needle could be bent during the injection. (3) the needle tip is the most important segment to be localized. (4) most of the needle segments in OCT B-scans are above the retina since the injection depth is relatively shallow (less than 250

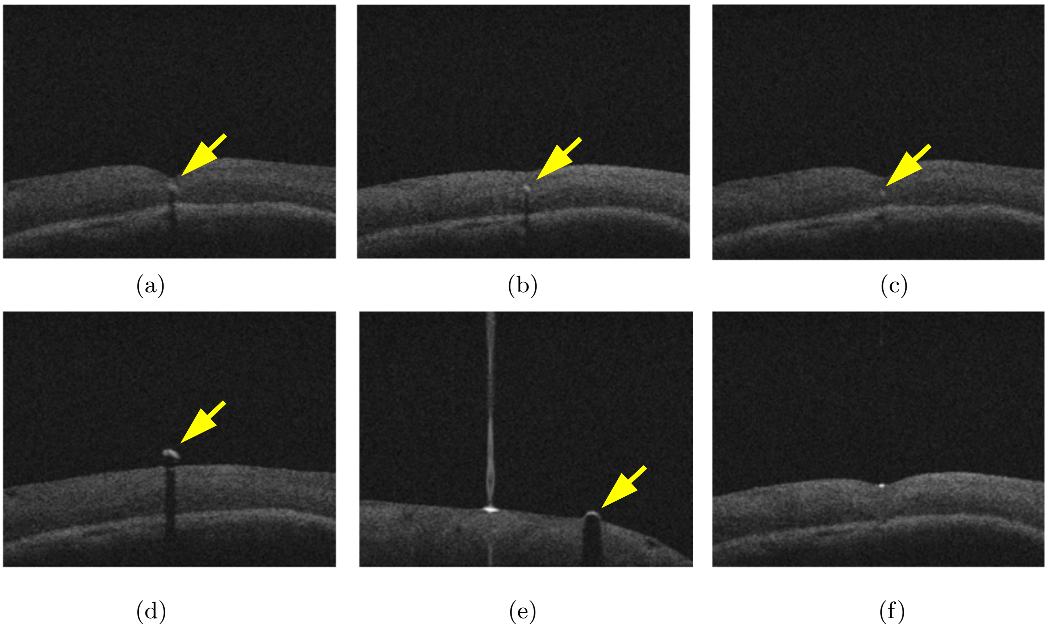


Figure 6.2: (a), (b), (c), (d), (e) are examples with needle segment above and under retina (yellow arrow is used to localize the needle segment position). (f) is the one of confusion examples without needle but noise or reflection.

µm) compared to the OCT imaging range in depth. Based on point (4) we can apply the needle and retinal surface detection algorithms based on ellipse detection algorithms described in [75], the points in each B-scan that are used to fit the ellipse can be covered with a bounding box B_i. Moreover, we use the middle point on the upper bottom edge $M_i = (Mx_i, My_i, Mz_i)$ to fit the space polynomial curve which predicts the needle segment location instead of using the ellipse center because the center of the fitted ellipse will not be accurate. A second order polynomial is chosen because higher orders result in rippling which effects uncharacteristic of the actual needle deflection. The needle model $N(x, y, z)$ is shown as,

$$N(x,y,z) \begin{cases} x(z) = a_2 z^2 + a_1 z + a_0 \\ y(z) = b_2 z^2 + b_1 z + b_0 \end{cases} \tag{6.1}$$

where a_2, a_1, a_0 and b_2, b_1, b_0 represent the parameters to be identified for the $N(x, y, z)$. The inlier dataset can be defined as,

$$inliers = \{e_i < \varepsilon\} \tag{6.2}$$

where ε is the threshold for inlier tolerance, e_i is the distance between point M_i and point N in Euclidean space which can be calculated as,

$$e_i = \sqrt{(Mx_i - x(Mz_i))^2 + (My_i - x(My_i))^2} \tag{6.3}$$

The point in M_i will be categorized as outliers if e_i is larger than ε. The rest points are treated as consensus set and a cost function C will be calculated for all of these points in M_i shown as follows,

$$C = \sum_{i=1}^{K} f(e_i) \tag{6.4}$$

where $f(e_i)$ equals e_i when $e_i < \varepsilon$, otherwise equals ε. The M-Estimator Sample Consensus (MSAC) technique developed by Torr et al. [108] is used to obtain the optimal consensus set with minimization of the cost function. With the output of the parameters for needle model N, the needle location in all B-scan images can be predicted as shown in Fig. 6.3. Thus, the ROI bounding box $L_i = [Lx_i, Ly_i, w, h]$ (Lx_i and Ly_i are coordinated for the top left corner of the bounding box, w and h are the width and height of the bounding box) in i-th B-scan image is represented as,

$$L_i = \begin{cases} NaN, \text{if } y(Mz_i) < \frac{h}{2}. \\ [x(Mz_i) - \frac{w}{2}, y_{max} - h, w, y_{max}], \\ \text{if } y(Mz_i) > y_{max} - \frac{h}{2}. \\ [x(Mz_i) - \frac{w}{2}, y(Mz_i) - \frac{h}{2}, w, h], \\ \text{if } \frac{h}{2} < y(Mz_i) < y_{max} - \frac{h}{2}. \end{cases} \tag{6.5}$$

where NaN means the bounding box does not exist, y_{max} is maximum imaging depth. In this way, we could crop bounding box with lots of candidates where the needle may appear under retina. The needle reaching out of image will be ignored since needle segment in these candidates are always too far away above the retina which are not important information. w and h are decided by the needle diameter which ensures the needle to be covered by the cropped bounding box. Here with the needle of 40 gauge (Ø=0.16 mm), we set $w = 120$ and $h = 240$ where the larger h can help to include the shadow information.

6.3.2 Automatic Needle Localization Under Retina

For needle point classification and detection, RetinaNet [107] model is used. This network is a single unified network composed of backbone network and task specific subnetwork as shown in Fig. 6.1. Considering that we feed the image with the cropped size, the backbone network used here is Resnet18 for providing convolutional feature map and subnetwork perform the classification and bounding box regression task. The RetinaNet is one detector which can achieve comparable accuracy as two-stage detector while have the real time processing speed. The detectors normally have the problem of class imbalance which is even more serious in our needle detection scenario. Based on the subretinal injection tests on the ex-vivo pig eyes, the bounding box area of the needle segment is ranging from 352 to 1131 pixels which is a small region compared to the cropped image with 28,800 pixels and the original image with 524,288 pixels. This means that even though we have already cropped the image into a significantly smaller size compared to original image, the imbalance of the foreground and background problem still exits. Thus the focal loss concept is introduced to evaluate the unbalance of the foreground and background [107] as follows:

$$FL(p_t) = -\alpha(1 - p_t)^{\gamma} log(p_t) \tag{6.6}$$

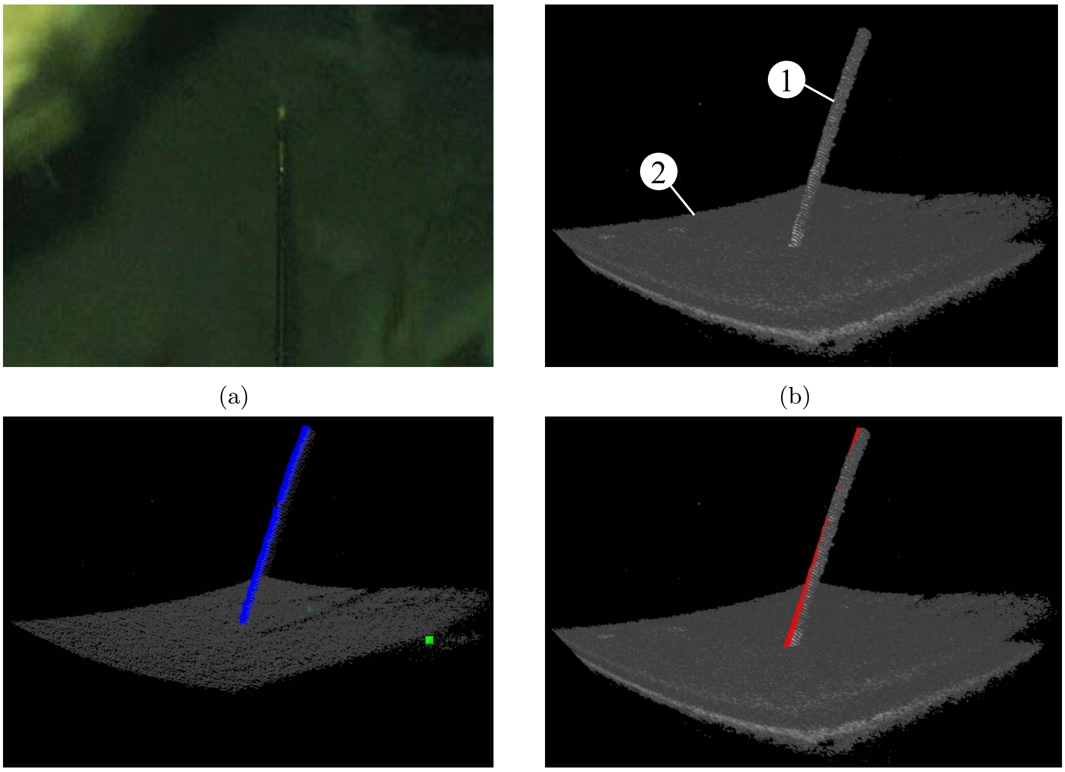
where p is the model's estimated probability for the specified class; $\alpha \in [0, 1]$ is the weighting factor for the specified class; and $\gamma \in [0, +\infty]$ is the tuneable focusing parameter. As previously introduced, the needle above the retinal surface can be detected and automatic annotated with the ellipse fitting algorithm. Therefore, the main difficulty comes when the needle detection under the surface. Thus we firstly manually annotate the needle which has the interaction with the retina (adhered to the retinal surface or under retinal surface), and then train the RetinaNet model with the annotated images. Moreover, we evaluate the trained model to localize the needle segment with our test dataset. The evaluation performance of the network will be introduced in the next section.

6.4 Experiments and Results

In this section, we present dataset preparation with our robot-assisted subretinal injection setup as well as the results related to the network performance with different parameters.

6.4.1 Dataset Preparation

The dataset is collected on the robot-assisted subretinal injection platform with ex-vivo fresh pig eyes as shown in Fig. 6.4. The fresh ex-vivo pig eye (prepared for experiments within 2 hours after removal) has very similar structure with the human eye. The iRAM!S robot is mounted on an adjustment bracket. The OPMI LUMERA 700 with integrated RESCAN 700 intraoperative OCT engine is fixed on the optical table to reduce the influence of ambient vibration. OCT scan



(a) (b)

(c) (d)

Figure 6.3: (a) The microscope image with needle inserted inside retina. (b) The rendered OCT volume in oblique view. ① represents needle and ② represents retina. (c) M_i points in OCT volume. The blue point are inliers and green points are outliers. (d) Needle model $N(x, y, z)$ in OCT volume. The red line is the RANSAC fitted needle model. The image crop operation are carried out based on $N(x, y, z)$ and Eqn. 6.5.

area can be easily relocated by the control panel. The iRAM!S eye surgical robot with 5DoF, with Pirezo motor technology, is utilized to perform the injection experiments. Piezo motor (SmarACT GmbH) provides 1 µm accuracy by using PID controller with integrated incremental optical encoder. The robot is adjusted manually by the adjustment bracket to make the needle tip approach to the retina. During the injection, only the motor that is holding the syringe is enabled to control the needle position. Thereafter, the motor is controlled to advance an exact 20 µm for each step and the OCT volume is captured after each movement. We stop the movement with enough insertion depth and the same procedure is repeated on the several ex-vivo pig eyes. Finally, we get 150 cubes with 19,200 images.

All images are processed by the ROI crop method and are divided into three categories, no needle (NN), needle above retina (NAR) which means needle has no interaction with retinal tissue, needle under retina (NUR) which means needle has interaction with retinal tissue. Since most of the images present the needle above the retina, we randomly select 1000 images NN, 1000 images NAR, and 1000 images NUR from 140 cubes to balance the different categories. The rest 10 cubes with 811 images (287 images NN, 440 images NAR, and 84 images NUR) are selected as test dataset. All the selected and processed images are put into two datasets, one is original image dataset where the image size is 512×1024 and the other one is the cropped image dataset where the image size is 120×240.

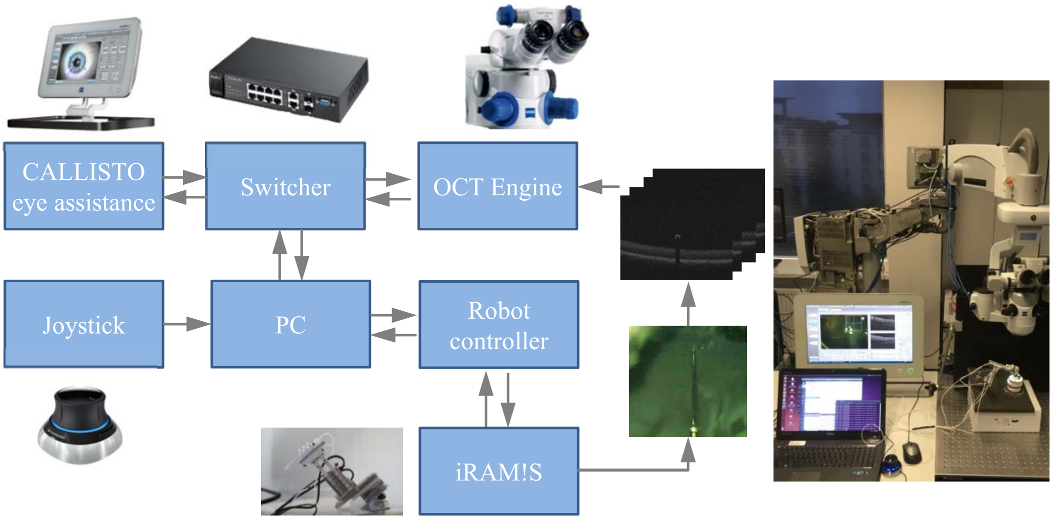


Figure 6.4: The robot-assisted subretinal injection setup.

6.4.2 Training and Evaluation

6.4.2.1 Metrics

Since the needle position is critical to be known in the future application, for example, detecting the needle insertion depth, intersection over union (IoU) for detected bounding box which represents the accuracy of needle position is used. The IoU is calculated as,

$$\mathrm{IoU} = A_o/A_u \tag{6.7}$$

where A_o is the area of the overlap for the detected bounding box and ground truth bounding box, A_u is the area of the union for the detected bounding box and ground truth bounding box. To determine the performance of the needle with respect to the overall performance, we use the mean average precision (MAP) to calculate the mean value of average precision (AP) for each individual query Q_i. MAP can be calculated as,

$$\mathrm{MAP} = \frac{1}{|Q|}\sum_{i=1}^{|Q|} \mathrm{AP}(Q_i) \tag{6.8}$$

where $|Q|$ denotes the number of the categories. By introducing both metrics, we could analyze the performance of category accuracy and also the needle localization precision.

6.4.2.2 Training

Both datasets are trained with the same RetinaNet network for 100 epochs from scratch with around 5 hours in NVidia Titan-X GPU with i7-7700K CPU and 16 GB RAM. In order to determine the optimized γ and α, different combinations of γ and α are tested and tuned with the objective function of AP and MAP. The output of the network contains many bounding boxes with the probability of which category it is. We only output the one with the maximum probability. For tuning the parameters, the IoU threshold is set to 0.5 which is considered as a good location predictor for filtering out the inaccurate predicted needle position. The trained model evaluation results are listed in Table 6.1 and Table 6.2 for the original dataset and cropped dataset, respectively. From these results, we can find that the cropped dataset trained model is significantly better than the original dataset trained model overall with the best-tuned result mAP of 0.97 and 0.81, respectively. Moreover, the interface time of cropped dataset trained model is furthermore reduced by 37% compared to the original dataset trained model. From these results, it can be seen that γ and α have the ability to tune the performance to a certain degree. However, the original data trained model introduces not only larger foreground-background unbalance, but also more noise and reflection reducing the information entropy.

Table 6.1: The original dataset trained model.

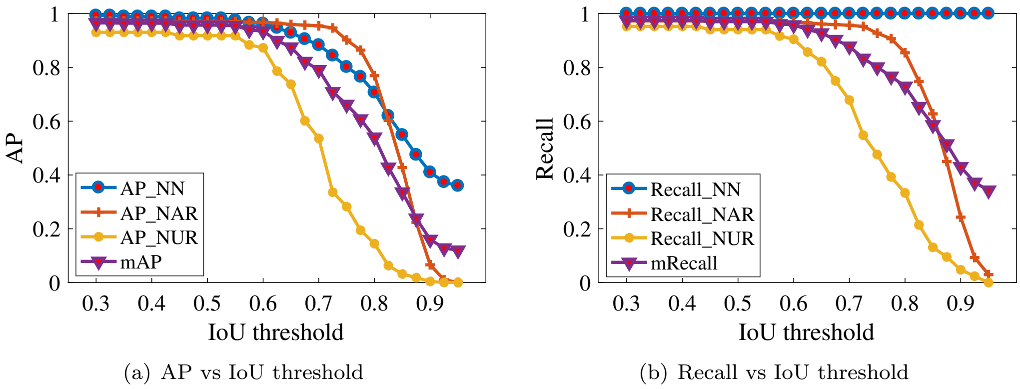
γ	α	AP_NN	AP_NAR	AP_NUR	MAP	time(ms)
1.0	0.25	0.3539	0.0	0.0	0.1769	50.6
1.5	0.25	0.9663	0.8490	0.2639	0.7614	54.7
2.5	0.25	0.9503	0.8823	0.4367	0.8049	52.3
5.0	0.25	0.8750	0.8272	0.2560	0.6959	54.8
1.0	0.5	0.3539	0.8945	0.2938	0.4740	54.6
1.5	0.5	0.9535	0.8579	0.2595	0.7561	54.5
2.5	0.5	0.9696	0.8544	0.4194	0.8033	53.4
5.0	0.5	0.9025	0.8715	0.2671	0.7359	54.0
1.0	0.75	0.9631	0.8821	0.3148	0.7808	53.3
1.5	0.75	0.9795	0.8890	0.2828	0.7827	52.6
2.5	**0.75**	**0.9729**	**0.8602**	**0.4552**	**0.8153**	**53.7**
5.0	0.75	0.9535	0.8706	0.3184	0.7740	128.4

Table 6.2: The cropped dataset trained model.

γ	α	AP_NN	AP_NAR	AP_NUR	MAP	time(ms)
1.0	0.25	0.9897	0.9702	0.9011	0.9627	33.6
1.5	0.25	0.9931	0.9701	0.9180	0.9686	33.5
2.5	0.25	0.9863	0.9701	0.8991	0.9604	33.6
5.0	0.25	0.9729	0.9680	0.8470	0.9402	36.5
1.0	0.5	0.9897	0.9740	0.9057	0.9648	33.3
1.5	0.5	0.9897	0.9655	0.9221	0.9667	35.9
2.5	0.5	0.9931	0.9725	0.9125	0.9678	34.4
5.0	0.5	0.9897	0.9680	0.8732	0.9551	33.5
1.0	0.75	0.9931	0.9772	0.9080	0.9678	36.1
1.5	0.75	0.9931	0.9698	0.9006	0.9641	36.1
2.5	**0.75**	**0.9931**	**0.9770**	**0.9201**	**0.9708**	**33.8**
5.0	0.75	0.9897	0.9627	0.9043	0.9616	36.1

6.4.2.3 Evaluation

In order to give the sensitivity analysis of the IoU threshold value, we evaluate the AP and recall performance with different IoU threshold on the cropped dataset trained model(γ=2.5, α=0.75) shown in Fig. 6.5. The AP and recall value show steady until IoU threshold is more than 0.55 and then the AP and recall for needle detection (NAR and NUR) decrease to 0 when the IoU threshold is more than 0.9. Table 6.3 shows the evaluation result of annotation and detection with the IoU threshold of 0.55 in the best model. By integrating NAR and NUR into one category, meaning that the needle existing in the image regardless of the interaction situation of needle and retina, only 5 images are misclassified to NN. This indication shows that the method has the ability to localize the needle with the accuracy of 0.55 (IoU value) in the confidence of 99.2%. Fig. 6.6 shows some of the detected examples providing the method is working properly in most of the cases. It only fails to detect when needle tip is very small (usually the first needle tip in the B-scan). This will not reduce the operational quality since missing one B-scan needle tip detection will only cause 23 µm (the scan resolution in Z direction) at most for needle localization error in one direction. This error is within the tolerance of needle localization requirement in subretinal injection which is 25 µm.



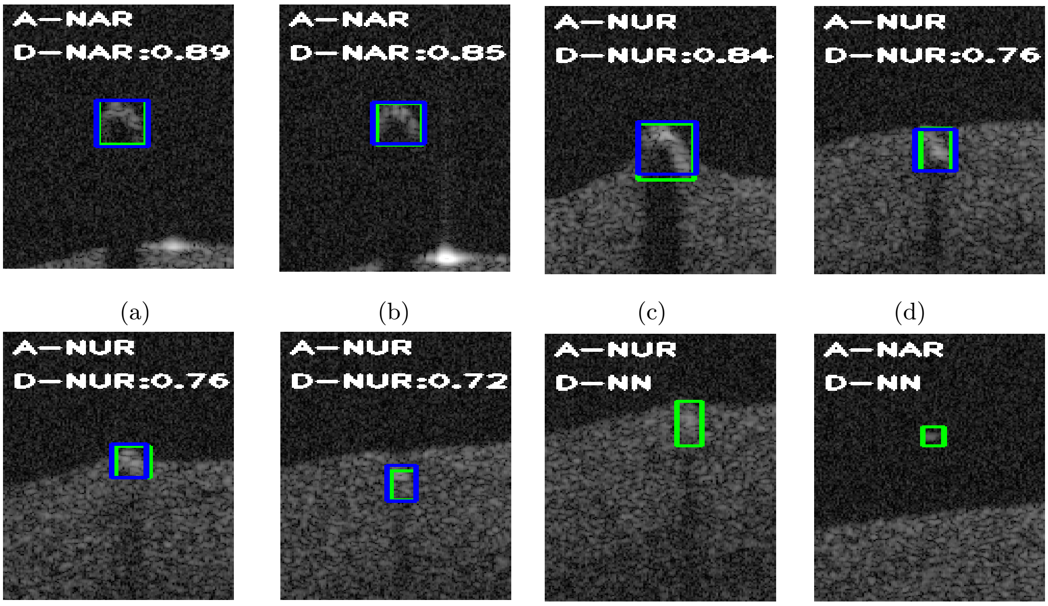
(a) AP vs IoU threshold

(b) Recall vs IoU threshold

Figure 6.5: (a) The function of IoU threshold with AP in three categories (AP_NN, AP_NAR, and AP_NUR). (b) The function of IoU threshold with AP in three categories (Recall_NN, Recall_NAR, and Recall_NUR), where mRecall denotes the mean recall value of three category.

6.5 Conclusion

In this chapter, we presented a flexible and robust framework for needle detection and localization in subretinal injection using MI-OCT images based on deep learning. The aim of the



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

(e) (f) (g) (h)

Figure 6.6: (a), (b), (c), (d), (e) and (f) are examples with correct detection. (g) and (f) are examples of incorrect detection when needle tip is very small. The yellow bounding box is the annotation-category (A-NAR, A-NUR, and A-NN) and blue bounding box is the detection-category (D-NAR, D-NUR, and D-NN). The number following is the IoU value.

Table 6.3: The evaluation of the three categories.

	NN	NAR	NUR	Detection
NN	287	2	3	292
NAR	0	427	2	429
NUR	0	11	79	90
Annotation	287	440	84	

proposed method is to overcome the shortcoming of the proposed method in Chapter 5 which has the assumption of needle has no significant deformation. Different from feeding data directly into the network, we take advantage of the needle geometrical features in volumetric OCT images to design a robust ROI crop method, thus the image size is significantly reduced. Afterward, the state of art one-stage detector named RetinaNet is applied to train the cropped images for the needle detection and localization. The evaluation results on ex-vivo pig eyes show that the performance of cropped image trained model is significant better than the original image trained model.The cropped image trained model can localize the needle accurately with a confidence of 99.2%. In this research, we have not investigated how much the needle will be deformed for 40-41 G needle. This could be an interesting problem for further direction. What is more, using the proposed method in this chapter to tracking the deformation of the needle during the insertion would be also an interesting point for the future.

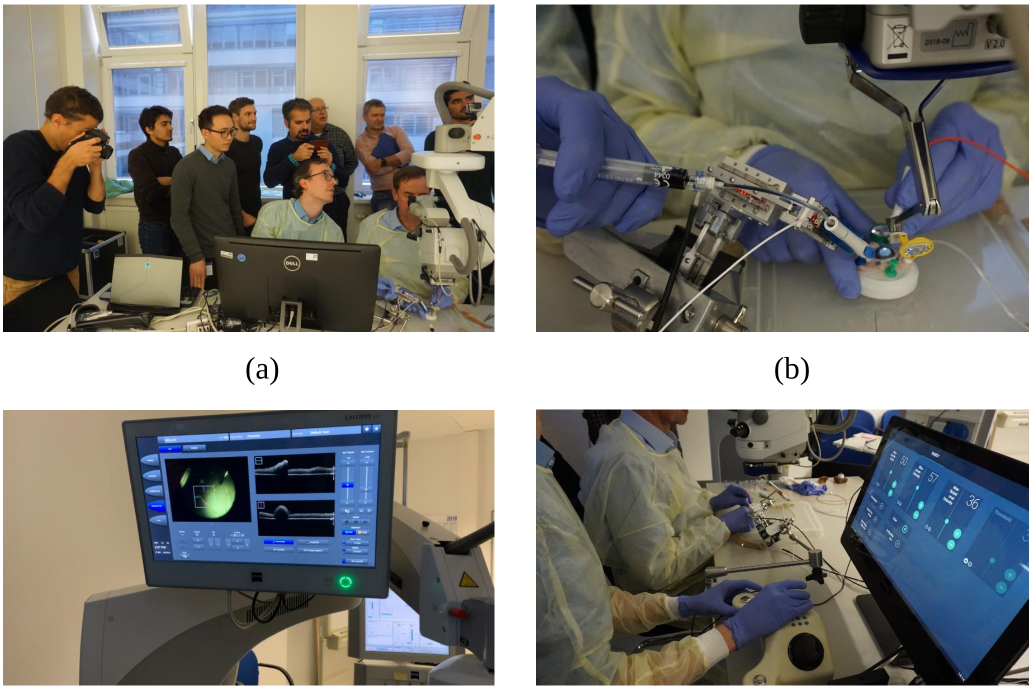
6.6 Related Publication

1. **Mingchuan Zhou**, Xijia Wang, Jakob Weiss, Abouzar Eslami, Kai Huang, Mathias Maier, Chris P. Lohmann, Nassir Navab, Alois Knoll, M. Ali Nasseri. Needle Localization for Robot-assisted Subretinal Injection based on Deep Learning.(accepted) In 2019 IEEE International Conference on Robotics and Automation(ICRA), May 2019.

Chapter 7

Conclusions

This chapter presents closing arguments for this thesis. Specially, a summary of contributions as well as directions for future work.


(a) (b)

(c) (d)

Figure 7.1: The wetlab test of robotic-assisted subretinal injection manually with OCT navigation. (a) Surgeons test the robot with OCT navigation on ex vivo pig eye. (b) The robot setup. (c) The interface for microscope and OCT image. (c) The robot control interface and control joystick.

7.1 Summary

The main target of this thesis is to figure out the possibility of autonomy for vitreoretinal surgery. The main contributions are to address the critical challenges with image guided robot-assisted surgery. The main motivation is introduced in the introduction that we intend to provide the surgical tools for reducing the stress of physical tremors and limited perception which challenges the surgeon during the surgery, thus they can focus more on the surgical plan and treatment decision. The current robotic system is tested intensively with the wetlab experiment, which shows that the system has the capability of performing the subretinal injection with OCT images navigation. We benefit a lot from the interaction with surgeons and the design of the system is developed based on the surgeons' suggestions with maximum technical possibility. All demonstrations and findings in this thesis were published in peer-reviewed conferences and journals, thus proving its credibility and novelty.

As Yang et al. propose in [109] regarding the autonomy level definitions of medical robotics, unlike autonomous cars, the technology challenges of the medical robot vary differently based on a specific task, environment, technology, and risk. The content in this thesis is intended to provide some fundamental clues to support the Level 3 (L3) of the autonomy, which is Conditional Autonomy, a system generates task strategies but relies on the human to select from among different strategies or to approve an autonomously selected strategy. For the subretinal injection application, this L3 autonomy can be automatic injecting the needle inside the retina and control the depth for around one minute which is an almost impossible task for the human hand. In this thesis, we explore this autonomy task by the precision robot and high-resolution OCT. Based on our current research results, we could conclude the autonomy of the subretinal injection is feasible. However, all current tests are performed on the ex-vivo well-fixed pig eyes. There would be more challenges for the in-vivo eye but somehow could be further tackled with faster OCT. This would be also very attractive to find solutions for these challenges.

7.2 Contributions

This thesis tries to tackle the problems in the enabling of the automatic subretinal injections, which we can summarize the main contributions as follows,

First of all, the workflow of automatic subretinal injection is designed with a proposed system. The proposed system has mainly two parts, the precision robot and high-resolution

OCT. The surgical workflow using this system consists of two main parts: the first part is the manual robot control part which is to aim the target before approaching the retinal surface while considering remote center of motion (RCM) constraint. When the injection area is precisely located, the second part which is needle injection mode will be activated. We present a robot system for subretinal injection integrated with intraoperative OCT. To ensure surgical safety, needle insertion depth will be estimated continuously from OCT images. A soft RCM control method is designed and integrated for the controller of our hybrid parallel-serial surgical robot. Safety and accuracy performance evaluation with a 15 ms control loop shows RCM deviation error is within 1 mm. Experimental results demonstrated that the proposed system can improve surgical outcomes by helping the surgeon to overcome their physical limitation for enabling a better dexterous motion.

Second, based on the proposed system and workflow, the needle should be controlled and located when it is above the retina. To archive this, this thesis presents a method to estimate the 6DOF needle pose specifically for the application of robotic intraocular needle navigation using OCT volumes. The key ingredients of the proposed method are (a) 3D needle point cloud segmentation in OCT volume and (b) needle point cloud 6DOF pose estimation using a modified iterative closest point (ICP) algorithm. To address the former, a voting mechanism with geometric features of the needle is utilized to robustly segment the needle in OCT volume. Afterward, the CAD model of the needle point cloud is matched with the segmented needle point cloud to estimate the 6DOF needle pose with a proposed shift-rotate ICP (SR-ICP). This method is evaluated by the existing ophthalmic robot on ex-vivo pig eyes. By using this needle segmentation method, we presents a flexible framework for hand-eye calibration of an ophthalmic robot with an OCT without any markers. The proposed method consists of three main steps: a) we estimate the OCT calibration parameters; b) with micro-scale displacements controlled by the robot, we detect and segment the needle tip in 3D-OCT volume; c) we find the transformation between the coordinate system of the OCT camera and the coordinate system of the robot. We verified the capability of our framework in ex-vivo pig eye experiments and compared the results with a reference method (marker-based). In all experiments, our method showed a small difference from the marker based method, with a mean calibration error of 9.2 μm and 7.0 μm, respectively. Additionally, the noise test shows the robustness of the proposed method.

Last but not least, the needle tracking method is developed for the needle insertion process. First, with the assumption of having no significant deformation for the needle, The geometrical based method is proposed to estimate the needle tip position under the retina. The basic idea is to calibrate the needle when it lies above the retina, and predict the tip position when the needle is inserted inside. We evaluated our approach with 32 G needle on both phantom tissues and ex-vivo porcine eyes. Evaluation results show that the average error in distance measurement is 4.7 µm (maximum of 16.5 µm). We furthermore, verified the feasibility of the proposed method to track the insertion depth of needle in robot-assisted subretinal injection. However, needle can have large deformation during the insertion when using small diameter needles. To consider the potential needle deformation during insertion, we furthermore demonstrate a robust framework for needle detection and localization in subretinal injection based on deep learning. The proposed method consists of two main steps: a) the preprocessing of OCT volumetric images; b) needle localization in the processed images. The first step is to coarsely localize the needle position based on the needle information above the retinal surface and crop the original image into a small region of interest (ROI). Afterward, the cropped small image is fed into a well-trained network for detection and localization of the needle segment. The entire framework is extensively validated in ex-vivo pig eye experiments with robotic subretinal injection. The results show that the proposed method can localize the needle accurately with the confidence of 99.2%.

7.3 Further Work

To estimate the instrument pose for vitreoretinal surgery using OCT volumetric images is relatively new and there are not so many references which we can refer to during my Ph.D. project from the year of 2015. This thesis tries to disclose the tip of the iceberg from the point of needle tip position & needle pose estimation with the background for the autonomy of subretinal injection. There are plenty of points still open for research in further work, which is summarized in the following,

1. The robot system used the piezo motor with a hybrid parallel-serial mechanism. The benefit of the robot is its light-weight and compact structure. However, due to the high cost of the motor in which the price is positively correlated to the travel distance of the motor, the workspace of the robot is limit to small space which can cover the eyeball area. Before the eye surgical, the surgeon needs to relocate the needle in an appropriate pose and position for entering the trocar. The robot also needs to remove the instrument away which far from the eye for some other operations. Therefore, the ability to move

the instrument far from the eyeball should be also enabled for the whole robotic system. For this purpose, a regular robot with 6 DoF can be integrated into the system.

2. We tried our robotic system on the plenty of ex-vivo pig eyes. We find out that the hand-eye calibration may be easily broken when there is a movement for the eye from the outside environment. This movement can be the heart beating and breath from the patient, the potential movement from eye muscle(even though with anesthesia), the friction force generated by needle and trocar, the drag force generated by the trocar and tissue. Even though there are some solutions for the fix the eye with a vacuum sucker, the performance still needs to be validated and tested in the in vivo trials. A dynamic compensation method for the maintaining or updating of the hand-eye calibration in the future.

3. The retinal layers segmentation needs to be addressed in further work. In this thesis, we mostly focus on the needle segmentation and ILM layer segmentation, as we present in the introduction that, ideally, the drug needs to be delivered between the layer of outer nuclear layer and RPE layer. Even though the thickness of the retina is around 250 μm, the thickness varies in a different area of the retina. In the future, we can design an algorithm to detect not only the ILM layer but also the RPE layer, so that we can make sure that the needle is above the RPE layer.

4. The quality of the OCT signal varies during the experiment. The light path through the cornea, lens, vitreous body is complicated and may reduce the reflection light back to the OCT engine, making the needle reflection signal weak sometimes. With the deep learning approach based needle detection method, we have the confidence of 99.2% to detect the needle, however, this confidence could still be improved for medical application. One potential method is to take advantage of which a needle is a continuous object in the OCT volumes. Therefore the sequence of B-scan images could be further used to increase the accuracy of the needle detection. Expect for the information fusion method, some coating technology may need to be considered for the needle to enhance the needle imaging signal [110] which may be a more fundamental way to increase the detection accuracy of the needle.

5. Due to the image mechanism, the OCT image modality has an image depth of 2 mm in tissue. The retina tissue has a thickness of 0.2 mm, while the eyeball has a diameter of 25

> mm. It shows that the OCT is suitable for needle localization in a small range. Therefore, in this thesis, the OCT is used only for the fine localization in which the instrument is already controlled by a human via robot manually localize to the area of interest. However, to achieve a fully autonomous operation inside the eye, an additional image modality with a large imaging range should be compensatively adapted with OCT. The naturally good option would be the microscope image since it has already been intensively used during the operation.

Although the difficulties for the autonomy of vitreoretinal surgery is challenging from both technical and ethical aspects, it has a huge benefit for human society and releases the hard workload from the surgeons, allowing surgeons to focus more on the intraoperative surgical planning and decision making. It would be also interesting to see the overall development of autonomy in a surgical task in the next decades and how the new solutions are proposed and interacted with the challenges of technology and ethic. The previous chapter proposed an algorithm to predict the needle tip under retina using the geometrical information of the needle. The main premise is that the needle does not have large deformation during the insertion. However, the thin needle can bring better clinic performance for reducing injection trauma to the retina.

References

[1] Taiga Nakano, Naohiko Sugita, Takashi Ueta, Yasuhiro Tamaki, and Mamoru Mitsuishi. **A parallel robot to assist vitreoretinal surgery**. *Int. J. Comput. Assist. Radiol. Surg.*, **4**(6):517–526, 2009. x, 3, 5, 6, 13, 58

[2] *The History of Ophthalmology.* 2013. 1

[3] Daniel M Albert and Mark J Lucarelli. *Clinical Atlas of Procedures in Ophthalmic Surgery.* Amer Medical Assn, 2004. 1, 5

[4] Allen Foster and Serge Resnikoff. **The impact of Vision 2020 on global blindness**. *Eye*, **19**(10):1133, 2005. 1

[5] Donatella Pascolini and Silvio Paolo Mariotti. **Global estimates of visual impairment: 2010**. *British Journal of Ophthalmology*, **96**(5):614–618, 2012. 1

[6] Robin Casten, Barry W Rovner, and Joseph L Fontenot. **Targeted Vision Function Goals and Use of Vision Resources in Ophthalmology Patients with Age-Related Macular Degeneration and Comorbid Depressive Symptoms.** *J. Vis. Impair. & Blind.*, **110**(6):413–424, 2016. 1

[7] Wan Ling Wong, Xinyi Su, Xiang Li, Chui Ming G Cheung, Ronald Klein, Ching-Yu Cheng, and Tien Yin Wong. **Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis**. *The Lancet Global Health*, **2**(2):e106–e116, 2014. 1

[8] Sonia Mehta. **Age-related macular degeneration**. *Primary Care: Clinics in Office Practice*, **42**(3):377–391, 2015. 1

[9] Sachin Parikh, Andrew Le, Julian Davenport, Michael B Gorin, Steven Nusinowitz, and Anna Matynia. **An Alternative and Validated Injection Method for Accessing the Subretinal Space via a Transcleral Posterior Approach**. *J. Vis. Exp.*, (118):e54808—-e54808, 2016. 2

[10] Kanmin Xue, Jasleen K Jolly, Alun R Barnard, Anna Rudenko, Anna P Salvetti, Maria I Patrício, Thomas L Edwards, Markus Groppe, Harry O Orlans, Tanya Tolmachova, et al. **Beneficial effects on vision in patients undergoing retinal gene therapy for choroideremia**. *Nature medicine*, **24**(10):1507, 2018. 2

[11] G Alex Ochakovski, K Ulrich Bartz-Schmidt, and M Dominik Fischer. **Retinal gene therapy: surgical vector delivery in the translation to clinical trials**. *Frontiers in neuroscience*, **11**:174, 2017. 2

[12] Iulian Iordachita, Zhenglong Sun, Marcin Balicki, Jin U Kang, Soo Jay Phee, James Handa, Peter Gehlbach, and Russell Taylor. **A sub-millimetric, 0.25 mN resolution fully integrated fiber-optic force-sensing tool for retinal microsurgery**. *International journal of computer assisted radiology and surgery*, **4**(4):383–390, 2009. 3

[13] Stanislao Rizzo, Fabio Patelli, and David Chow. *Vitreo-retinal surgery: progress III.* Springer Science Business Media, 2008. 3

[14] Wei Wei, Roger Goldman, Nabil Simaan, Howard Fine, and Stanley Chang. **Design and theoretical evaluation of micro-surgical manipulators for orbital manipulation and intraocular dexterity**. In *Robot. Autom. 2007 IEEE Int. Conf.*, pages 3389–3395. IEEE, 2007. 3, 6, 13

[15] V Narendran and Abhishek Kothari. *Principles and Practice of Vitreoretinal Surgery.* JP Medical Ltd, 2014. 3

[16] Russell H Taylor, Janez Funda, David D Grossman, John P Karidis, and David A LaRose. **Remote center-of-motion robot for surgery**, 1995. 3

[17] Dan Stoianovici, Louis L Whitcomb, James H Anderson, Russell H Taylor, and Louis R Kavoussi. **A modular surgical robotic system for image guided**

percutaneous procedures. In *Int. Conf. Med. Image Comput. Comput. Interv.*, pages 404–410. Springer, 1998. 3

[18] Ben Mitchell, John Koo, Iulian Iordachita, Peter Kazanzides, Ankur Kapoor, James Handa, Gregory Hager, and Russell Taylor. **Development and application of a new steady-hand manipulator for retinal surgery**. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 623–629. IEEE, 2007. 4

[19] SPN Singhy and CN Riviere. **Physiological tremor amplitude during retinal microsurgery**. In *Proceedings of the IEEE 28th Annual Northeast Bioengineering Conference (IEEE Cat. No. 02CH37342)*, pages 171–172. IEEE, 2002. 4

[20] Michael U Humayun, R Scott Rader, Dante J Pieramici, Carl C Awh, and Eugene de Juan. **Quantitative measurement of the effects of caffeine and propranolol on surgeon hand tremor**. *Archives of ophthalmology*, **115**(3):371–374, 1997. 4

[21] M Lakie, K Frymann, F Villagra, and P Jakeman. **The effect of alcohol on physiological tremor**. *Experimental Physiology: Translation and Integration*, **79**(2):273–276, 1994. 4

[22] Kyle M Fargen, Raymond D Turner, and Alejandro M Spiotta. **Factors that affect physiologic tremor and dexterity during surgery: a primer for neurosurgeons**. *World neurosurgery*, **86**:384–389, 2016. 4

[23] Michael J Elman, Joel Sugar, Richard Fiscella, Thomas A Deutsch, James Noth, Michael Nyberg, Kirk Packo, and Robert J Anderson. **The effect of propranolol versus placebo on resident surgical performance.** *Transactions of the American Ophthalmological Society*, **96**:283, 1998. 4

[24] Daniel R Evans. **Arm and hand rest device for microsurgery**, April 19 1977. US Patent 4,018,217. 4

[25] Cameron N Riviere and Patrick S Jensen. **A study of instrument motion in retinal microsurgery**. In *Eng. Med. Biol. Soc. (EMBC), 2000 22th Annu. Int. Conf. IEEE*, **1**, pages 59–60. IEEE, 2000. 4, 44, 57

[26] Wei Tech Ang, Cameron N Riviere, and Pradeep K Khosla. **Design and implementation of active error canceling in hand-held microsurgical instrument**. In *ntelligent Robot. Syst. 2001IEEE/RSJ Int. Conf.*, **2**, pages 1106–1111. IEEE, 2001. 4, 6, 13, 44

[27] E Rahimy, J Wilson, T C Tsao, S Schwartz, and J P Hubschman. **Robot-assisted intraocular surgery: development of the IRISS and feasibility studies in an animal model**. *Eye*, **27**(8):972–978, 2013. 4, 6, 13, 44, 58

[28] Andy Gijbels, EB Vander Poorten, Benjamin Gorissen, Alain Devreker, Peter Stalmans, and Dominiek Reynaerts. **Experimental validation of a robotic comanipulation and telemanipulation system for retinal surgery**. In *2014 5th IEEE RAS EMBS Int. Conf. Biomed. Robot. Biomechatronics*, pages 144–150. IEEE, 2014. 4, 6, 13, 44, 58

[29] Peter H Tomlins and RK Wang. **Theory, developments and applications of optical coherence tomography**. *Journal of Physics D: Applied Physics*, **38**(15):2519, 2005. 4, 9

[30] Justis P Ehlers, Peter K Kaiser, and Sunil K Srivastava. **Intraoperative optical coherence tomography using the RESCAN 700: preliminary results from the DISCOVER study**. *Br. J. Ophthalmol.*, pages bjophthalmol—-2014, 2014. 5, 9, 25, 42, 59, 63

[31] Maximilian Pfau, Stephan Michels, Susanne Binder, and Matthias D Becker. **Clinical experience with the first commercially available intraoperative optical coherence tomography system**. *Ophthalmic Surgery, Lasers and Imaging Retina*, **46**(10):1001–1008, 2015. 5

[32] Xingchi He, Marcin A Balicki, Jin U Kang, Peter L Gehlbach, James T Handa, Russell H Taylor, and Iulian I Iordachita. **Force sensing microforceps with integrated fiber bragg grating for vitreoretinal surgery**. In *Optical Fibers and Sensors for Medical Diagnostics and Treatment Applications XII*, **8218**, page 82180. International Society for Optics and Photonics, 2012. 5

[33] Eleanora P Westebring-van der Putten, Richard HM Goossens, Jack J Jakimowicz, and Jenny Dankelman. **Haptics in minimally invasive surgery, a review**. *Minimally Invasive Therapy Allied Technologies*, **17**(1):3–16, 2008. 5

[34] Puneet K Gupta, Pahick S Jensen, and Eugene de Juan Jr. **Surgical forces and tactile perception during retinal microsurgery**. In *Int. Conf. Med. Image Comput. Comput. Interv.*, pages 1218–1225. Springer, 1999. 5

[35] Maneesh Dewan, Panadda Marayong, Allison M Okamura, and Gregory D Hager. **Vision-based assistance for ophthalmic micro-surgery**. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 49–57. Springer, 2004. 5

[36] Seung Hyuk Baik. **Robotic colorectal surgery**. *Yonsei medical journal*, **49**(6):891–896, 2008. 6

[37] Shawn Tsuda, Dmitry Oleynikov, Jon Gould, Dan Azagury, Bryan Sandler, Matthew Hutter, Sharona Ross, Eric Haas, Fred Brody, and Richard Satava. **SAGES TAVAC safety and effectiveness analysis: da Vinci surgical system (Intuitive Surgical, Sunnyvale, CA)**. *Surgical endoscopy*, **29**(10):2873–2884, 2015. 6

[38] Angelo Tsirbas, C Mango, and E Dutson. **Robotic ocular surgery**. *British journal of ophthalmology*, **91**(1):18–21, 2007. 6

[39] JEAN PIERRE Hubschman, ANGELO Tsirbas, and STEVEN D Schwartz. **Robotic Surgery in Ophthalmology**. *Retina Today*, pages 81–84, 2008. 6

[40] Mohammad Beheshti, Saman Rahbar, Hossein Davarpanah, Somayeh Jowkar, and Farzin Piltan. **Design Auxiliary Sliding Variable Sliding Mode Controller for Robot-Assisted Ophthalmic Surgery**. *International Journal of Bio-Science and Bio-Technology*, **7**(5):187–202, 2015. 6

[41] Russell Taylor, Pat Jensen, Louis Whitcomb, Aaron Barnes, Rajesh Kumar, Dan Stoianovici, Puneet Gupta, ZhengXian Wang, Eugene Dejuan, and Louis Kavoussi. **A steady-hand robotic system for microsurgical augmentation**. *Int. J. Rob. Res.*, **18**(12):1201–1210, 1999. 6, 13, 58

[42] Franziska Ullrich, Christos Bergeles, Juho Pokki, Olgac Ergeneman, Sandro Erni, George Chatzipirpiridis, Salvador Pané, Carsten Framme, and Bradley J Nelson. **Mobility experiments with microrobots for minimally**

invasive intraocular SurgeryMicrorobot experiments for intraocular surgery. *Invest. Ophthalmol. Vis. Sci.*, **54**(4):2853–2863, 2013. 6, 13

[43] H C M Meenink, RHPCJN Rosielle, M Steinbuch, H Nijmeijer, and M C De Smet. **A master-μslave robot for vitreo-retinal eye surgery**. In *Euspen Int. Conf.*, pages 3–6, 2010. 6, 13

[44] Cameron N Riviere, Wei Tech Ang, and Pradeep K Khosla. **Toward active tremor canceling in handheld microsurgical instruments**. *IEEE Transactions on Robotics and Automation*, **19**(5):793–800, 2003. 6, 13

[45] Cheol Song, Peter L Gehlbach, and Jin U Kang. **Active tremor cancellation by a smart handheld vitreoretinal microsurgical tool using swept source optical coherence tomography**. *Opt. Express*, **20**(21):23414–23421, 2012. 6, 13, 27, 58, 61

[46] M A Nasseri, Martin Eder, Saurabh Nair, E C Dean, Martin Maier, D Zapp, C P Lohmann, and Aaron Knoll. **The introduction of a new robot for assistance in ophthalmic surgery**. In *Eng. Med. Biol. Soc. (EMBC), 2013 35th Annu. Int. Conf. IEEE*, pages 5682–5685. IEEE, 2013. 6, 14, 58, 60

[47] Amir Molaei, Ebrahim Abedloo, Marc D de Smet, Sare Safi, Milad Khorshidifar, Hamid Ahmadieh, Mohammad Azam Khosravi, and Narsis Daftarian. **Toward the art of robotic-assisted vitreoretinal surgery**. *Journal of ophthalmic & vision research*, **12**(2):212, 2017. 6, 13

[48] TL Edwards, K Xue, HCM Meenink, MJ Beelen, GJL Naus, MP Simunovic, M Latasiewicz, AD Farmery, MD de Smet, and RE MacLaren. **First-in-human study of the safety and viability of intraocular robotic surgery**. *Nature Biomedical Engineering*, page 1, 2018. 6, 7, 13, 57, 58, 61

[49] Riccardo Muradore and Paolo Fiorini. **A PLS-based statistical approach for fault detection and isolation of robotic manipulators**. *IEEE Transactions on Industrial Electronics*, **59**(8):3167–3175, 2012. 6, 13

[50] Hirohiko Arai, Kazuo Tanie, and Susumu Tachi. **Path tracking control of a manipulator considering torque saturation**. *IEEE Transactions on Industrial Electronics*, **41**(1):25–31, 1994. 6, 13

[51] Berk Gonenc, Alireza Chamani, James Handa, Peter Gehlbach, Russell H Taylor, and Iulian Iordachita. **3-DOF Force-Sensing Motorized Micro-Forceps for Robot-Assisted Vitreoretinal Surgery**. *Sensors*, **17**(11):3526–3541, 2017. 7

[52] Nicola Rieke, David Joseph, Chiara Amat, Federico Tombari, Mohamed Alsheakhali, Vasileios Belagiannis, Abouzar Eslami, and Nassir Navab. **Real-time localization of articulated surgical instruments in retinal microsurgery**. **34**:82–100, 2016. 7

[53] Michael Waine, Carlos Rossa, Ronald Sloboda, Nawaid Usmani, and Mahdi Tavakoli. **3d shape visualization of curved needles in tissue from 2d ultrasound images using ransac**. In *Robot. Autom. 2015 IEEE Int. Conf.*, pages 4723–4728. IEEE, 2015. 8

[54] Young-Joon Jo, Dong-Won Heo, Yong-IL Shin, and Jung-Yeul Kim. **Diurnal variation of retina thickness measured with time domain and spectral domain optical coherence tomography in healthy subjects**. *Investig. Ophthalmol. & Vis. Sci.*, **52**(9):6497–6500, 2011. 8

[55] Nathan McDannold, Greg Clement, Peter Black, Ferenc Jolesz, and Kullervo Hynynen. **Transcranial MRI-guided focused ultrasound surgery of brain tumors: Initial findings in three patients**. *Neurosurgery*, **66**(2):323, 2010. 8, 27

[56] Elliot R McVeigh, Michael A Guttman, Robert J Lederman, Ming Li, Ozgur Kocaturk, Timothy Hunt, Shawn Kozlov, and Keith A Horvath. **Real-time interactive MRI-guided cardiac surgery: Aortic valve replacement using a direct apical approach**. *Magn. Reson. Med.*, **56**(5):958–964, 2006. 8, 27

[57] Gustaaf J Vrooijink, Momen Abayazid, and Sarthak Misra. **Real-time three-dimensional flexible needle tracking using two-dimensional ultrasound**. In *Robot. Autom. 2013 IEEE Int. Conf.*, pages 1688–1693. IEEE, 2013. 8, 27

[58] Tim T Lam, Paul Miller, Susan Howard, and T Michael Nork. **Validation of a Rabbit Model of Choroidal Neovascularization Induced by a Subretinal Injection of FGF-LPS**. *Investig. Ophthalmol. & Vis. Sci.*, **55**(13):1204, 2014. 8, 27, 61

[59] Martin F Kraus, Benjamin Potsaid, Markus A Mayer, Ruediger Bock, Bernhard Baumann, Jonathan J Liu, Joachim Hornegger, and James G Fujimoto. **Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns**. *Biomedical optics express*, **3**(6):1182–1199, 2012. 9

[60] Marcel Pfister, Jaw-Chyng L Lue, Francisco R Stefanini, Paulo Falabella, Laurie Dustin, Michael J Koss, and Mark S Humayun. **Comparison of reaction response time between hand and foot controlled devices in simulated microsurgical testing**. *BioMed research international*, **2014**, 2014. 22

[61] Tomoya Sakai, Kanako Harada, Shinichi Tanaka, Takashi Ueta, Yasuo Noda, Naohiko Sugita, and Mamoru Mitsuishi. **Design and development of miniature parallel robot for eye surgery**. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 371–374. IEEE, 2014. 22

[62] Nicola Rieke, David Joseph Tan, Chiara Amat di San Filippo, Federico Tombari, Mohamed Alsheakhali, Vasileios Belagiannis, Abouzar Eslami, and Nassir Navab. **Real-time localization of articulated surgical instruments in retinal microsurgery**. *Medical image analysis*, **34**:82–100, 2016. 25

[63] Iro Laina, Nicola Rieke, Christian Rupprecht, Josué Page Vizcaíno, Abouzar Eslami, Federico Tombari, and Nassir Navab. **Concurrent segmentation and localization for tracking of surgical instruments**. In *International conference on medical image computing and computer-assisted intervention*, pages 664–672. Springer, 2017. 25

[64] Thomas Probst, Kevis-Kokitsi Maninis, Ajad Chhatkuli, Mouloud Ourak, Emmanuel Vander Poorten, and Luc Van Gool. **Automatic Tool Landmark Detection for Stereo Vision in Robot-Assisted Retinal Surgery**. *IEEE Robotics and Automation Letters*, **3**(1):612–619, 2018. 25, 41, 61

[65] Wadim Kehl, Federico Tombari, Slobodan Ilic, and Nassir Navab. **Real-time 3D model tracking in color and depth on a single CPU core**. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 745–753, 2017. 26

[66] WIKIPEDIA. `https://en.wikipedia.org/wiki/Birmingham_gauge`. Accessed: 2018-08-12. 26

[67] YEQING LI, CHEN CHEN, XIAOLEI HUANG, AND JUNZHOU HUANG. **Instrument tracking via online learning in retinal microsurgery**. In *Int. Conf. Med. Image Comput. Comput. Interv.*, pages 464–471. Springer, 2014. 26

[68] NICOLA RIEKE, DAVID JOSEPH TAN, MOHAMED ALSHEAKHALI, FEDERICO TOMBARI, CHIARA AMAT DI SAN FILIPPO, VASILEIOS BELAGIANNIS, ABOUZAR ESLAMI, AND NASSIR NAVAB. **Surgical tool tracking and pose estimation in retinal microsurgery**. In *Int. Conf. Med. Image Comput. Comput. Interv.*, pages 266–273. Springer, 2015. 26, 61

[69] RAPHAEL SZNITMAN, KARIM ALI, ROGERIO RICHA, RUSSELL H TAYLOR, GREGORY D HAGER, AND PASCAL FUA. **Data-driven visual tracking in retinal microsurgery**. In *Int. Conf. Med. Image Comput. Comput. Interv.*, pages 568–575. Springer, 2012. 26

[70] HAORAN YU, JIN-HUI SHEN, KAREN M JOOS, AND NABIL SIMAAN. **Design, calibration and preliminary testing of a robotic telemanipulator for OCT guided retinal surgery**. In *Robot. Autom. 2013 IEEE Int. Conf.*, pages 225–231. IEEE, 2013. 27, 61

[71] XUAN LIU, MARCIN BALICKI, RUSSELL H TAYLOR, AND JIN U KANG. **Towards automatic calibration of Fourier-Domain OCT for robot-assisted vitreoretinal surgery**. *Opt. Express*, **18**(23):24331–24343, 2010. 27, 61

[72] JAKOB WEISS, NICOLA RIEKE, MOHAMMAD ALI NASSERI, MATHIAS MAIER, ABOUZAR ESLAMI, AND NASSIR NAVAB. **Fast 5DOF needle tracking in iOCT**. *International journal of computer assisted radiology and surgery*, **13**(6):787–796, 2018. 27, 63

[73] NILS GESSERT, MATTHIAS SCHLÜTER, AND ALEXANDER SCHLAEFER. **A deep learning approach for pose estimation from volumetric OCT data**. *Medical image analysis*, **46**:162–179, 2018. 27, 63

[74] HESSAM ROODAKI, KONSTANTINOS FILIPPATOS, ABOUZAR ESLAMI, AND NASSIR NAVAB. **Introducing Augmented Reality to Optical Coherence Tomography in Ophthalmic Microsurgery**. In *Mix. Augment. Reality, 2015 IEEE Int. Symp.*, pages 1–6. IEEE, 2015. 28, 29, 47, 48, 59, 63

[75] Mingchuan Zhou, Mahdi Hamad, Jakob Weiss, Abouzar Eslami, Kai Huang, Mathias Maier, Chris P Lohmann, Nassir Navab, Alois Knoll, and M Ali Nasseri. **Towards Robotic Eye Surgery: Marker-Free, Online Hand-Eye Calibration Using Optical Coherence Tomography Images**. *IEEE Robotics and Automation Letters*, **3**(4):3944–3951, 2018. 29, 80

[76] MA Nasseri, M Maier, and CP Lohmann. **A targeted drug delivery platform for assisting retinal surgeons for treating Age-related Macular Degeneration (AMD)**. In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, pages 4333–4338. IEEE, 2017. 34, 42, 48

[77] Mohammad Ali Nasseri, P Gschirr, Martin Eder, Suraj Nair, K Kobuch, Mathias Maier, Daniel Zapp, C Lohmann, and Alois Knoll. **Virtual fixture control of a hybrid parallel-serial robot for assisting ophthalmic surgery: An experimental study**. In *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 732–738. IEEE, 2014. 36

[78] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, and James G Fujimoto. **Optical coherence tomography**. *Science (New York, NY)*, **254**(5035):1178, 1991. 41

[79] Jesús Díaz Díaz, Dennis Kundrat, Kim-Fat Goh, Omid Majdani, and Tobias Ortmaier. **Towards intra-operative oct guidance for automatic head surgery: first experimental results**. In *MICCAI*, pages 347–354. Springer, 2013. 42

[80] Gyeong-Woo Cheon, Yong Huang, and Jin U Kang. **Active depth-locking handheld micro-injector based on common-path swept source optical coherence tomography**. In *SPIE BiOS*, pages 93170U—-93170U. International Society for Optics and Photonics, 2015. 42, 59, 63

[81] Roger Y Tsai and Reimar K Lenz. **A new technique for fully autonomous and efficient 3D robotics hand/eye calibration**. *IEEE Transactions on robotics and automation*, **5**(3):345–358, 1989. 43

[82] Éric Marchand, Fabien Spindler, and François Chaumette. **ViSP for visual servoing: a generic software platform with a wide class of robot control skills**. *IEEE Robotics & Automation Magazine*, **12**(4):40–52, 2005. 43

[83] Yiu Cheung Shiu and Shaheen Ahmad. **Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX= XB**. *IEEE Transactions on robotics and automation*, **5**(1):16–29, 1989. 43

[84] Isabella Morgan, Uditha Jayarathne, Adam Rankin, Terry M Peters, and Elvis CS Chen. **Hand-eye calibration for surgical cameras: a Procrustean Perspective-n-Point solution**. *Int. J. Comput. Assist. Radiol. Surg.*, pages 1–9, 2017. 43

[85] Stephen Thompson, Danail Stoyanov, Crispin Schneider, Kurinchi Gurusamy, Sébastien Ourselin, Brian Davidson, David Hawkes, and Matthew J Clarkson. **Hand–eye calibration for rigid laparoscopes using an invariant point**. *Int. J. Comput. Assist. Radiol. Surg.*, **11**(6):1071–1080, 2016. 43

[86] Christoph Bergmeir, Mathias Seitel, Christian Frank, Raffaele De Simone, H-P Meinzer, and Ivo Wolf. **Comparing calibration approaches for 3D ultrasound probes**. *Int. J. Comput. Assist. Radiol. Surg.*, **4**(2):203, 2009. 43

[87] Johan Sarrazin, Emmanuel Promayon, Michael Baumann, and Jocelyne Troccaz. **Hand-eye calibration of a robot-UltraSound probe system without any 3D localizers**. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 21–24. IEEE, 2015. 43

[88] Krittin Pachtrachai, Max Allan, Vijay Pawar, Stephen Hailes, and Danail Stoyanov. **Hand-eye calibration for robotic assisted minimally invasive surgery without a calibration object**. In *Intelligent Robot. Syst. 2006 IEEE/RSJ Int. Conf.*, pages 2485–2491. IEEE, 2016. 43

[89] Francisco Vasconcelos, Donald Peebles, Sebastien Ourselin, and Danail Stoyanov. **Similarity registration problems for 2D/3D ultrasound calibration**. In *European Conference on Computer Vision*, pages 171–187. Springer, 2016. 43

[90] Francisco Vasconcelos, Donald Peebles, Sebastien Ourselin, and Danail Stoyanov. **Spatial calibration of a 2D/3D ultrasound using a tracked needle**. *Int. J. Comput. Assist. Radiol. Surg.*, **11**(6):1091–1099, 2016. 43, 55

[91] Sam Van der Jeught, Jan AN Buytaert, Adrian Bradu, Adrian Gh Podoleanu, and Joris JJ Dirckx. **Real-time correction of geometric distortion**

artefacts in large-volume optical coherence tomography. *Measurement Science and Technology*, **24**(5):057001, 2013. 45

[92] K.S Arun, T.S Huang, and S.D Blostein. **Least-squares fitting of two 3-D point sets**. *IEEE Trans. Pattern Anal. Mah. Intell.*, **9**(5):698–700, 1987. 50

[93] Berthold K. P. Horn. **Closed-form solution of absolute orientation using unit quaternions**. *Journal of the Optical Society of America A*, **4**(4):629–642, 1987. 50

[94] Pierre Chatelain, Alexandre Krupa, and Maud Marchal. **Real-time needle detection and tracking using a visually servoed 3D ultrasound probe**. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 1676–1681. IEEE, 2013. 50

[95] Homa Alemzadeh, Daniel Chen, Andrew Lewis, Zbigniew Kalbarczyk, Jaishankar Raman, Nancy Leveson, and Ravishankar Iyer. **Systems-Theoretic Safety Assessment of Robotic Telesurgical Systems**. In *Computer Safety, Reliability, and Security*, pages 213–227. Springer International Publishing, 2015. 55

[96] Alexander Barthel, Diego Trematerra, M Ali Nasseri, Daniel Zapp, Chris P Lohmann, Alois Knoll, and Mathias Maier. **Haptic interface for robot-assisted ophthalmic surgery**. In *Eng. Med. Biol. Soc. (EMBC), 2015 37th Annu. Int. Conf. IEEE*, pages 4906–4909. IEEE, 2015. 58

[97] Berk Gonenc, Russell H Taylor, Iulian Iordachita, Peter Gehlbach, and James Handa. **Force-sensing microneedle for assisted retinal vein cannulation**. In *IEEE Sensors Journal*, pages 698–701. IEEE, 2014. 60

[98] Mingchuan Zhou, Kai Huang, Abouzar Eslami, Hessam Roodaki, Daniel Zapp, Mathias Maier, Chris P. Lohmann, Alois Knoll, and M. Ali Nasseri. **Precision Needle Tip Localization using Optical Coherence Tomography Images for Subretinal Injection, to be appear.** In *IEEE Int. Conf. on Rob. and Auto.*, 2018. 61

[99] Mingchuan Zhou, Hessam Roodaki, Abouzar Eslami, Guang Chen, Kai Huang, Mathias Maier, Chris P Lohmann, Alois Knoll, and Mohammad Ali Nasseri. **Needle segmentation in volumetric optical coherence tomography images for ophthalmic microsurgery**. *Applied Sciences*, **7**(8):748, 2017. 63, 65

[100] Hessam Roodaki, Chiara Amat di San Filippo, Daniel Zapp, Nassir Navab, and Abouzar Eslami. **A Surgical Guidance System for Big-Bubble Deep Anterior Lamellar Keratoplasty**. In *Int. Conf. Med. Image Comput. Comput. Interv.*, pages 378–385. Springer, 2016. 64

[101] Mingchuan Zhou, Kai Huang, Abouzar Eslami, Daniel Zapp, Haotian Lin, Mathias Maier, Chris P. Lohmann, Alois Knoll, and M. Ali Nasseri. **Beveled Needle Position and Pose Estimation based on Optical Coherence Tomography in Ophthalmic Microsurgery**. In *2017 IEEE International Conference on Robotics and Biomimetics*, December 2017. 65

[102] Radu Bogdan Rusu and Steve Cousins. **3d is here: Point cloud library (pcl)**. In *Robot. Autom. 2011 IEEE Int. Conf.*, pages 1–4. IEEE, 2011. 67

[103] Radu Bogdan Rusu, Nico Blodow, Zoltan Marton, Alina Soos, and Michael Beetz. **Towards 3D object maps for autonomous household robots**. In *Intell. Robot. Syst. 2007 IEEE/RSJ Int. Conf.*, pages 3191–3198. IEEE, 2007. 68

[104] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. *Surface reconstruction from unorganized points*, **26**. ACM, 1992. 68

[105] Lis Custodio, Tiago Etiene, Sinesio Pesco, and Claudio Silva. **Practical considerations on Marching Cubes 33 topological correctness**. *Comput. & Graph.*, **37**(7):840–850, 2013. 70

[106] Jasper RR Uijlings, Koen EA Van De Sande, Theo Gevers, and Arnold WM Smeulders. **Selective search for object recognition**. *International journal of computer vision*, **104**(2):154–171, 2013. 78

[107] Tsung-Yi Lin, Priyal Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. **Focal loss for dense object detection**. *IEEE transactions on pattern analysis and machine intelligence*, 2018. 78, 81, 82

[108] Philip HS Torr and Andrew Zisserman. **MLESAC: A new robust estimator with application to estimating image geometry**. *Computer vision and image understanding*, **78**(1):138–156, 2000. 81

[109] Guang-Zhong Yang, James Cambias, Kevin Cleary, Eric Daimler, James Drake, Pierre E Dupont, Nobuhiko Hata, Peter Kazanzides, Sylvain Martel, Rajni V Patel, et al. **Medical robotics-Regulatory, ethical, and legal considerations for increasing levels of autonomy**. *Sci. Robot*, **2**(4):8638, 2017. 92

[110] Hungyen Lin, Yue Dong, Daniel Markl, Zijian Zhang, Yaochun Shen, and J Axel Zeitler. **Pharmaceutical film coating catalog for spectral domain optical coherence tomography**. *Journal of pharmaceutical sciences*, **106**(10):3171–3176, 2017. 95