University of Genoa

&

Italian Institute of Technology

PhD Program in Bioengineering and Robotics

Computational Sensing for ISM & OCT Guided Laser Microsurgery

by

Ajay Gunalan

Thesis submitted for the degree of *Doctor of Philosophy* (36° cycle)

27^{th} *March*, 2024

Thesis Supervisor

Dr. Leonardo De Mattos — Head of the Biomedical Robotics Lab, IIT

Head of the Ph.D. Program

Dr. Paolo Massobrio — Associate Professor, University of Genoa

Thesis Jury (External Examiners):

Dr. Edward Grant — Professor Emeritus, North Carolina State University

Dr. Brian Davies — Professor Emeritus, Imperial College London

Advanced Robotics,
Italian Institute of Technology

Department of Informatics, Bioengineering,
Robotics and Systems Engineering

This doctoral thesis and my life is dedicated to the unconditional love and tears of my mother ...

Love is God
If not for your love, I would have ceased to exist.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains fewer than 60,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 160 figures.

Ajay Gunalan
April 2024

Acknowledgements

A Ph.D. is an individual achievement, but this journey has definitely not been a solitary path. Throughout these three years, many people have directly or indirectly contributed to my endeavor to obtain a doctorate. I would like to thank each and every one of them from the bottom of my heart. Your support means a lot to me, and this achievement would not have been possible without it. Please forgive me if I inadvertently fail to mention someone in particular. I would like to start by thanking my Ph.D. thesis supervisor, Dr. Leonardo De Mattos, for providing me with this opportunity for doctoral research. He has been nothing but the best Ph.D. supervisor a student could hope for. He allowed me to work at my own pace, yet he was always available for guidance whenever I encountered a roadblock. Thank you for all the advice, ideas, moral support, and constructive criticism. It has truly been an honor to conduct my research under your supervision. I thank Prof. Edward Grant and Prof. Brain Davies for the time and energy they spent revising my thesis and providing constructive feedback. I also extend my gratitude to Dr. Darwin Caldwell for his support as ADVR director. I thank Nikhil Deshpande, who trusted me and gave me the opportunity to work at IIT, and for his patience throughout the long visa process. I am grateful to Andres Geraldes for encouraging me to pursue a Ph.D. and guiding me during my initial days of the program. Also, Veronica Penza for being exceptionally kind and supportive on both a personal and professional level. I am truly grateful for it. Next, I thank Yonas and Dario, who have created a lasting impact on my life. Yonas, with his hard work and attention to detail, and Dario, with his expertise, knowledge, and humility, have been inspiring to me. Now, I thank all current and past members of BRL & Vicarios Lab: Alberto, Federico, Chiara, Carola, Enrico, Valreo, Nabeel, Maria, Yaesol, et al, for creating a positive atmosphere at IIT. I would like to thank my friends Kapil, Shunlei, Adeel, Dr. Charanraj, Dr. Karthikeyan, Divya, Siddharth, Prashanth, and Vishal for numerous dinners and for making my stay here in Genoa homely and enjoyable. Rohith J.K., for being a true friend who supports me throughout my ups and downs. Finally, I thank my family, who has been an unwavering support to me.

Abstract

Lasers surgery offers precision and fewer complications, but challenges include controlling ablation depth and identifying tumor edges. Intraoperative imaging, especially optical laser scanning microscopes (LSM), such as optical coherence tomography (OCT) and image scanning microscopes (ISM), may overcome these issues. They provide benefits such as, cost-effectiveness, and no radiation exposure. This work investigates the use of OCT and ISM to enhance surgeries providing better depth perception, clearer tumor margins, and surgical precision. Transitioning optical imaging from bench-top to clinical settings requires overcoming challenges like limited focus depth, motion blur, and slow data acquisition. Computational sensing may be a solution to tackle this problem by co-designing hardware and software through techniques like optimization and machine learning. In particular, compressive sensing is of special interest here, as it enables sampling data below the Nyquist rate and is able to reconstruct original signals using computational methods. Compressive sensing is typically used in single-pixel cameras. This work explores applying compressive sensing to LSM, specifically ISM and OCT. First, we investigate how to combine laser surgery with optical imaging modalities. Then, we apply compressive sensing ISM and leverage the micro-images from a single-photon avalanche diode (SPAD) to produce better images. We address the limitations of 1D and 2D OCT scans under the separate optical path (SOP) category by using a 3D-OCT scan to determine the ablation depth. OCT scanning time is reduced by applying compressive sensing without significant loss in the quality of the depth map. Finally, we use the depth map to control the exposure time (laser on/off time) to accurately ablate a given depth across different tissue types in a feedback controller. So far, ablation over a given point has been performed using a high-speed scanner, but the ablation area remains limited. Auto-CALM extends the capability of CALM to ablate over a larger area defined by the surgeon automatically. It employs target tracking, laser tracking, and an ablation algorithm. Tested on a porcine model simulating breathing, Auto-CALM showed high precision and promised significant surgical advancements. Integrating compressive sensing with LSM modalities such as OCT and ISM is a significant step in transferring this technology from bench-top to clinical settings. We have shown that it reduces the photo-bleach of ISM, reduces the scanning time of OCT, and enables the control of the laser ablation depth. Additional efforts, such as the development of endoscopic probes that incorporate these technologies and the extension of the depth of focus, are essential for the clinical application of these advancements.

Key-words: Optical Coherence Tomography; Laser Microsurgery; Confocal Microscopy; Compressive Sensing; CALM; Laser Spot Tracking; Target Tracking; Auto-CALM;

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Abstract

The discussion begins with laser microsurgery and the challenges it introduces. Following this, the focus shifts to imaging methods utilizing lasers, specifically Optical Coherence Tomography (OCT) and Image Scanning Microscope (ISM). The organization of the thesis is then outlined, and the study's main contributions are highlighted.

1.1 Laser Microsurgery

The history of laser surgery begins in 1951, with the first medical application reported by Goldman in 1962. Since then, laser technology has evolved significantly, leading to minimally invasive procedures, shorter recovery times, and less risk to patient health [19], [6]. Lasers are used in a wide range of surgical applications, including cardiovascular surgery, ophthalmology, lithotripsy, cancer treatment, and cosmetic procedures. In cardiovascular surgery, for instance, lasers have proven effective in angioplasty for coronary and peripheral vessels [19]. In ophthalmology, lasers have revolutionized procedures like LASIK for vision correction [66]. In the field of oncology, lasers are used for tumor ablation and other cancer treatments [6]. In cosmetic and dermatological procedures, lasers are used for skin resurfacing, scar revision, hair removal, and tattoo removal [75], [45]. Laser surgery in Ear, Nose, and Throat (ENT) or Otorhinolaryngology has become a common practice due to its precision, minimally invasive nature, and wide range of applications. The most commonly used lasers in ENT are the CO2 (carbon dioxide) laser, the diode laser, and the neodymium YAG (Nd:YAG) laser[69]. Laser surgery offers several benefits, including non-contact tissue ablation, precise cutting, hemostasis, low cicatrization, reduced postoperative pain, and reduced tissue swelling. In addition, they can be combined with endoscopes, microscopes, and surgical robotic systems. The development of the CO2 laser in 1964 [62] brought an alternative approach to perform incisions on soft tissues. In 1972, Jako et al. conducted animal studies to demonstrate the viability of performing surgical interventions in the larynx using a CO2 laser as a tissue cutting tool [40]. Three years later, Strong [77] reported on the use of a CO2 laser in conjuction to a microscope for the excision of a laryngeal carcinoma from a human patient. Laser surgery presents numerous advantages over traditional cold-steel surgery. These advantages include better control over bleeding, enhanced sterility of the surgical site, and the absence of post-operative edema [82], [85]. Laser surgery is now an established alternative to radiation therapy for the treatment of laryngeal cancers [74], [68]. A review by Milford and O'Flynn from 1991 provides a detailed comparitive study on radiotherapy and laser excision for the treatment of laryngeal carcinomas taking into account recurrence and total treatment period. The conclusion was a positive assessment of laser surgery as a valid alternative treatment [57]. Soon after, it was reported that laser surgery enables effective treatment of T1 and T2 type tumors in the larynx, and constitutes a convenient and acceptable alternative for the patient [26].

1.2 Transoral Laser Microsurgery (TLM)

Technological progress has transformed lasers in versatile tools for the treatment of diverse pathologies affecting delicate human organs. Together with other important applications such as eye surgery, transoral surgery is a major application field for lasers. Transoral Laser Microsurgery (TLM) is an evolution of the early techniques for laser surgery of the larynx. It includes a suite of minimally invasive endoscopic techniques for the excision of minuscule laryngeal diseases [68] as shown in Fig 1.1. In degenerative diseases associated with cancer, the primary objective of TLM is to ensure a complete eradication of the malignant tissue. At the same time, surgeons try to minimize the removal of healthy tissues, to preserve as much organ function as possible [76]. The execution of such accurate resections requires the use of a microscope and precise control of the laser incisions. In today's surgical practice, laser incisions are performed manually, i.e. moving a free-beam laser by means of a joystick-like device, called laser micromanipulator [68]. The micromanipulator is coupled to the operating microscope, forming an effective surgical device, yet, it is difficult to master, especially because it breaks the hand-eye coordination of the surgeon [22].

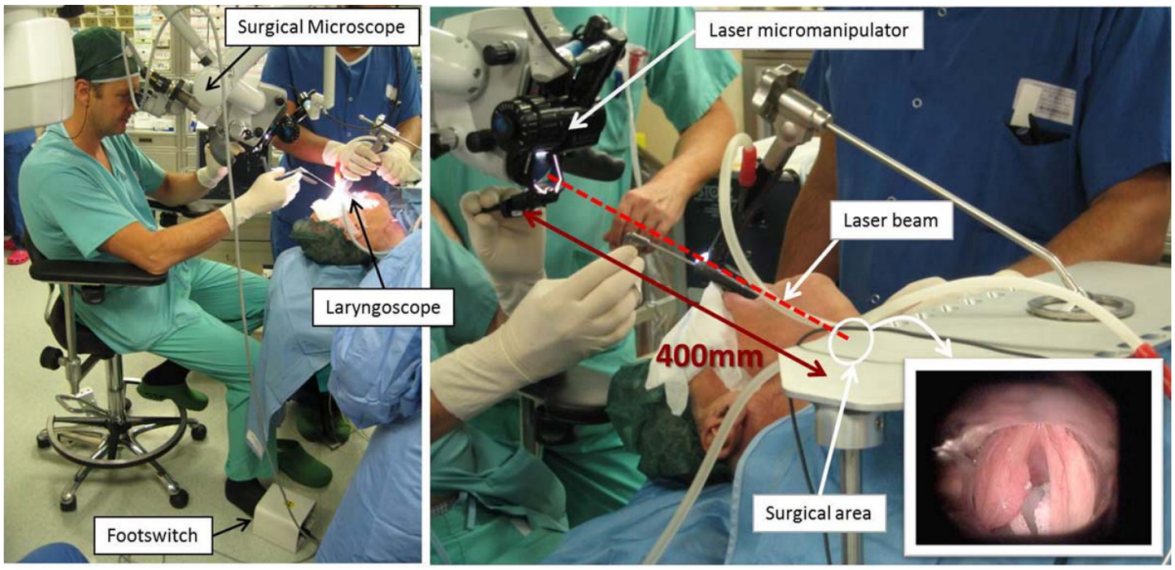

Figure 1.1 The state-of-the-art surgical setup in laser phonomicrosurgery [54]

1.3 Challenges in TLM

There are two main issues in TLM. First, the presently available technologies for TLM do not include any support for sensing the depth of laser cuts. The quality of incisions relies entirely on the dexterity and visual perception of clinicians. Extensive training is required to develop an effective laser cutting technique,which includes both (i) the acquisition of basic knowledge of the physical principles behind laser ablation of tissue; and (ii) the ability to manipulate the laser parameters and its exposure time in order to provide accurate cutting [68], [76]. Parameters traditionally used in TLM include power, energy delivery mode, pulse duration and incision scanning frequency. The resulting incision depends on the combination of these parameters plus the total time of laser exposure. However, it is not evident how to regulate these parameters in order to achieve the desired cutting level. While experienced clinicians normally have sufficient knowledge and understanding of the laser ablation processes, lack of experience represents a practical problem for many others. Secondly, there is the difficulty in determining tumor margins. laser surgery is challenging because of the complex laser-tissue interaction process. This leads to poor ablation depth control and tissue specificity, increasing the risk of accidental damage to nerves or blood vessels [59]. In our lab, depth control of laser surgery was previously achieved based on a tissue model using feedforward controllers [30]. Nevertheless, the efficacy of these controllers is contingent upon the fidelity of the tissue model, which presents significant challenges due to the heterogeneity of tissue types. Further complexity arises within individual tissue classifications, where variability is influenced by factors such as moisture content and the presence of neoplastic tissues, both of which substantially complicate the modeling process.

1.4 Intraoperative imaging

Intraoperative imaging can mitigate these challenges by providing accurate depth of ablation which can be used in a feedback controller. The current lack of intraoperative imaging capabilities has been associated with increased chances of over or under-treatment. Under-treatment requires repeated surgery with the associated cost, while over-treatment can lead to complications and reduce surgical quality. Various medical imaging technologies, such as MRI and CT, are commonly used in preoperative or postoperative settings. However, their use during intraoperative scenarios is generally limited. In contrast, ultrasound can be utilized intra-operatively, offering good imaging depth but

has lower resolution (0.2–2 mm) [34]. Optical coherence tomography (OCT) addresses this niche challenge of high-resolution (5–10 μm) intraoperative imaging. It has low-cost, radiation-free, and fiber-based advantages. In addition, OCT can be integrated with hand-held probes, laparoscopes, catheters, and endoscopes [33]. Therefore, OCT and ISM (Image Scanning Microscope) imaging provide microns and sub/micron-level resolution, respectively. Secondly, they can easily integrated with laser-based microsurgery.

1.5 Image Scanning Microscope

Laser scanning microscopy (LSM) is the optical architecture at the base of many imaging and spectroscopy techniques widely used in the fields of material science [80], biology [7], and medicine [36]. The reason for the success of LSM-based imaging techniques is their high spatiotemporal resolution and their capability to provide quantitative information. The LSM concept consists in focusing and scanning a laser beam on a sample to record either the scattered or fluorescent light with a single-element detector. Therefore, a complete image is built pixel-bypixel by arranging the recorded intensity values along the scanning pattern [70]. Among the LSM techniques, confocal laser scanning microscopy (CLSM) became especially popular, thanks to its ability of rejecting out-of-focus light and its superior spatial resolution. In detail, CLSM setups are designed to image the focal plane of the objective lens onto a circular pinhole, placed before the detector. Thus, the pinhole acts as a spatial filter allowing most of the light coming from the focal plane to reach the detector, while blocking most of the out-of-focus light. The closer the pinhole, the higher the spatial resolution and the signalto-background (SBR) ratio of the images [20]. Notably, CLSM images achieved a lateral resolution twice better than the optical diffraction limit in the extreme case of a point-like pinhole aperture [21], [10], [9]. Thus, theoretically, CLSM should be a super-resolution technique, but in a realistic scenario, the pinhole cannot be fully closed. Indeed, closing the pinhole reduces the overall amount of recorded photons, compromising the signal-to-noise ratio (SNR) of the acquired images and hindering the capability of CLSM to achieve super-resolution. Recently, image scanning microscopy (ISM) transformed confocal microscopy into a practical super-resolution technique. Theoretically developed in the 1980s [10], [71], [9], it has been experimentally realized for the first time in 2010 [58]. The core idea behind ISM is to replace the single-element detector with an array of detectors as shown in Fig 1.2, each detector acting as in a closed pinhole configuration. Thus, after a complete scan, each detector element generates a confocal image which represents

the same sample from a slightly different point-of-view. Furthermore, each detector contributes to the photon detection efficiency, maximizing the SNR of the acquisition. As a result, the ISM microscope provides a multi-dimensional dataset, which can be intuitively seen as a collection of scanned images, as many as the elements of the detector array, almost identical in content, but different in SNR and each shifted from each other by a quantity named shift-vector. Thus, estimating and compensating for the shifts enables the summation of all the collected images, constructing an ISM image with sub-diffraction resolution and enhanced SNR.

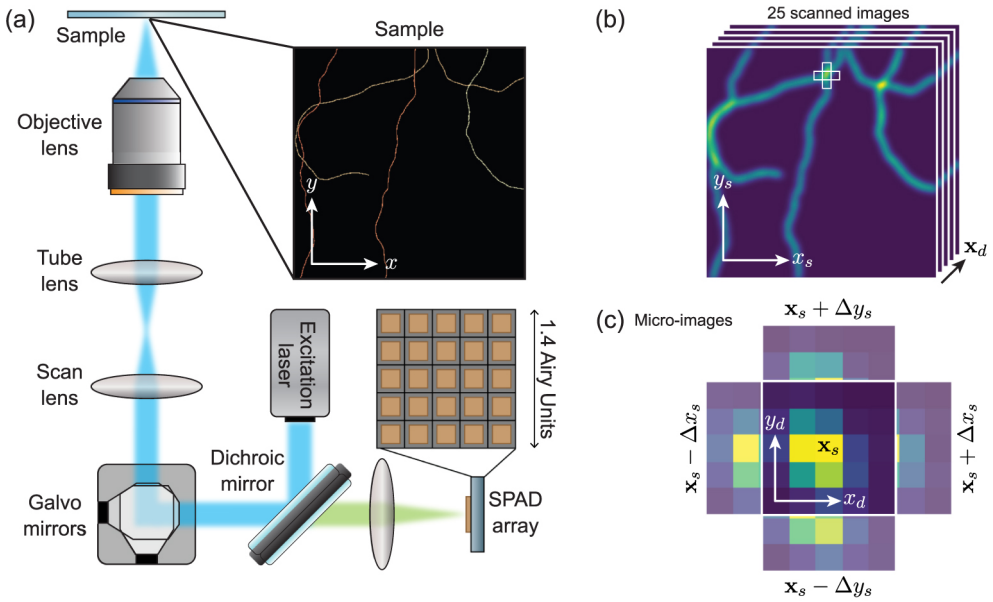

Figure 1.2 Image scanning microscopy. (a) A sketch of the laser scanning microscope equipped with a SPAD array detector. (b) The ISM dataset, seen as a set of scanned images as many as the number of elements of the detector array. (c) The ISM dataset, seen as a collection of micro-images, as many as the scan points. The depicted microimages correspond to the scan points highlighted in (b) as white boxes [94]. Figure licensed under CC BY 4.0.

1.6 Optical Coherence Tomography

OCT has a good potential to address these challenges given the following: (1) its imaging resolution ranges between 5 and 10 μm, which is enough to resolve precancerous lesions; (2) its penetration depth is normally between 1 and 3 mm, which is sufficient to image

the entire depth of epithelial tissues; and (3) it is an optical fiber-based imaging system and thus has potential to be compact, free of radiation, and affordable [78].

Optical Coherence Tomography (OCT) is a non-invasive imaging technique that resembles ultrasound imaging. However, unlike ultrasound, it uses an infrared laser to image both surface and subsurface details. OCT images are fundamentally made up of a one-dimensional point scan called an A-scan. A point on the sample is scanned to acquire its raw spectral data, consisting of spectrum intensity along various depths. These data are processed to obtain the A-scan data, which represents the reflectivity signal strength at various depths in the direction of that point.When the A-scan is repeated along a line in the sample, a B-scan (2D images) is obtained. Repeating the B-scan along the surface of the sample yields a C-scan (3D images). Laser scanners, which may be based on various technologies such as mirror galvanometers, MEMS-mirrors, MEMS-lenses, or fiber scanners, perform both B-scans and C-scans.

The most straightforward OCT configuration is based upon the Michelson interferometer, which splits the light 50/50 into a sample arm and a reference arm. In this simple configuration, nearly 75% of the optical power is wasted [67]. The SNR of OCT is approximately proportional to the optical power and is inversely proportional to the bandwidth of the detection electronics. Hence, alternative power-conserving configurations have been developed. As described in [67], such alternative configurations may use optical circulators, unbalanced couplers, and/or balanced heterodyne detection.

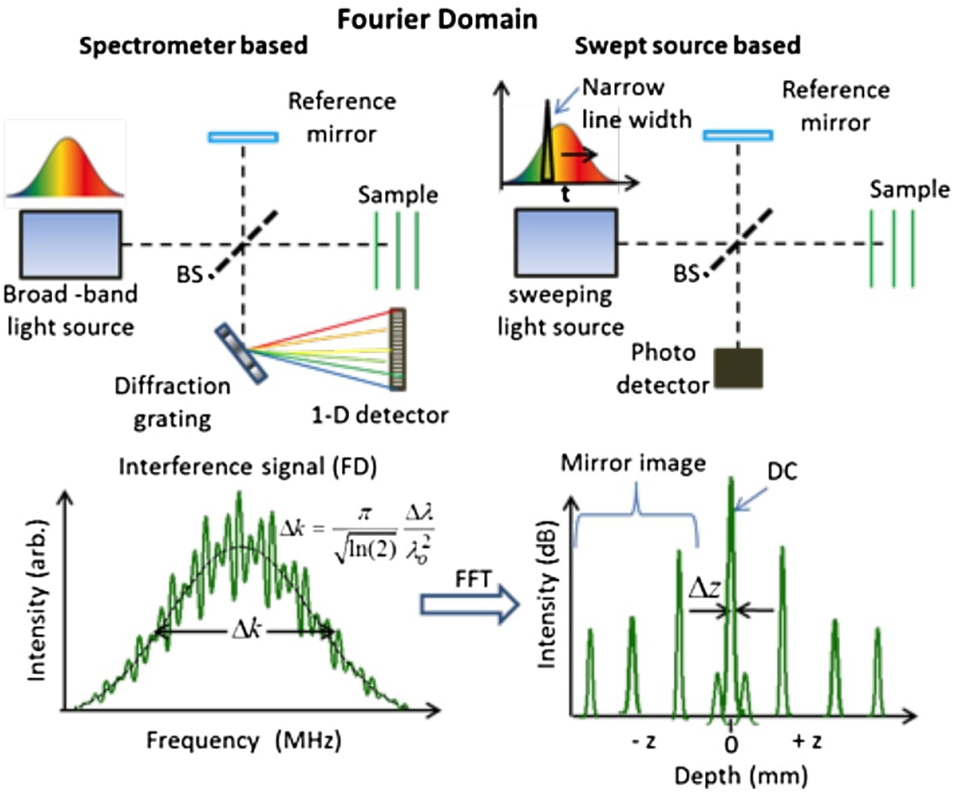

Figure 1.3 Comparison between Spectral Domain OCT (SD-OCT) and Swept Source OCT (SS-OCT) by [24]. Figure licensed under CC BY 4.0.

The first-generation OCT is known as Time-Domain OCT (TD-OCT). In these systems, the reference arm was mechanically scanned, limiting the image acquisition rate. Evolution toward Fourier Domain OCT (FD-OCT) gave rise to the second-generation OCT, which eliminated the need to mechanically scan the reference arm. These newer devices can be classified into either Swept Source (SS-OCT) or Spectral Domain (SD-OCT), as illustrated in 1.3. Both methods are equivalent from a theoretical point of view. However, SD-OCT uses a broadband light source with a spectrometer at the interferometer exit. SS-OCT uses a laser with a narrow instantaneous line width that is rapidly tunable over a large wavenumber range, which is combined with a single detector [4]. Both SD-OCT and SS-OCT are capable of real-time imaging with resolution (5–10 μm), which is about two orders of magnitude finer than ultrasound imaging (0.2–2 mm) [35]. Detailed mathematical descriptions of OCT fundamentals can be found at [13, 39].

Another modern OCT technology is the polarization-sensitive optical coherence tomography (PS-OCT). This is a set of hardware and software extensions to OCT

that allows measuring the birefringence of local regions of the tissue [1]. A tissue is said to be birefringent if the real part of its refractive index is polarization state dependent. Tissues such as muscle, cartilage and tendons exhibit birefringence due to their internal arrangement of sub-cellular fibrous structures. When the tissue is damaged or undergoes necrosis, this structure degenerates with a corresponding reduction in birefringence. The degree of birefringence, thus, gives an indication of the degree of tissue damage [39, 1].

Despite its capabilities and different imaging modalities, standard OCT systems cannot image internal organs due to their limited imaging depth.

1.7 Complementary Nature of ISM and OCT

OCT is capable of visualizing micron-scale tissue structural morphologies with the imaging contrast predominantly sensitive to the scattering property of tissues. Fluorescence imaging techniques such as confocal, ISM, two-photon excitation fluorescence (TPEF) can provide depth-resolved submicron-scale images with the imaging contrast coming from exogenous or endogenous fluorophores, thus providing molecular information which cannot be obtained by OCT. The two complementary imaging modalities provide important, yet different optical information based on unique contrast mechanisms. There is, hence, a strong motivation for developing integrated platforms for performing both OCT and fluorescence imaging. Furthermore, this combination has proven effective in visualizing morphology and fluorophore distribution in several benchtop systems [55].

1.8 Thesis Organisation & Contributions

2. **Literature Review**

This chapter delves into the current advancements in OCT-guided laser surgery in detail. Subsequently, the importance of separate optical paths (SOP) and computational sensing is briefly highlighted.

- *Publications:* ***A. Gunalan*** *and L. S. Mattos, "Towards OCT-Guided Endoscopic Laser Surgery—A Review," Diagnostics, vol. 13, no. 4, pp. 677, Feb. 2023. [Online]. Available: link*

3. Compressive Image Scanning Microscope

In the third chapter, the limitations of Image Scanning Microscopy (ISM) are overcome by reducing the number of scan points fourfold through a fixed sampling strategy and leveraging computational sensing to reconstruct the full image. To the best of our knowledge, this is the first application of compressive sensing to ISM.

- *Publications:* ***A. Gunalan****, M. Castello, S. Piazza, S. Li, A. Diaspro, L.S. Mattos, P. Bianchini, "Compressive Image Scanning Microscope," Presented at International Symposium on Computational Sensing, Luxembourg, 2023. link*

4. Compressive 3D OCT-Guided Depth Control

To overcome the limitations of 1D and 2D OCT scans under SOP categories, a 3D-OCT scan was employed to determine ablation depth. Additionally, compressive sensing was applied to OCT hardware to decrease scanning time without a significant loss in depth map quality. Lastly, the depth map was utilized to control the exposure time (laser on/off time) to accurately ablate a given depth across various tissue types using a feedback controller.

- *Publications:* (in-preparation)

5. Auto-CALM: Automatic Computer-Assisted Laser Microsurgery

In the previous chapters, ablation over a given point was performed using a high-speed scanner, but the area of ablation remained limited. This chapter extends the capability of CALM to automatically ablate over larger areas defined by the surgeon. Auto-CALM includes laser spot tracking, target tracking, and an automatic ablation algorithm.

- *Publications: S. Li,* ***A. Gunalan****, M.A. Azam, V. Penza, D. G. Cladwell, L.S. Mattos, "Auto-CALM: Automatic Computer-Assisted Laser Microsurgery," IEEE Transactions on Medical Robotics and Bionics.* (accepted)

6. Conclusion

- Summarize the entire work and discuss further future research direction.

Chapter 2

Literature Review

Abstract

This chapter reviews various state-of-the-art approaches for integrating OCT and laser surgical systems. These methods are categorized into three types: a double-clad fiber (DCF), a dichroic mirror, and separate optical paths. Additionally, the critical factors to consider when designing these systems are summarized, with a special emphasis on the use of Separate Optical Paths (SOP). The chapter concludes by introducing computational sensing and its relevance in this context.

2.1 Introduction

In this chapter, the state of the art in OCT-guided laser surgery is reviewed. The references are organized based on the methods utilized to integrate a high-power surgical laser with OCT. To date, these methods include: (1) double-clad fiber [41]; (2) dichroic mirror [43]; and (3) separate optical paths [28].

2.2 Double-Clad Fiber (DCF)

When light travels in an optical fiber, only waves at certain discrete angles greater than or equal to the fiber's critical angle will propagate, resulting in discrete modes or eigenfunctions. In a single-mode fiber (SMF), due to the small core diameter, only one mode of light propagates through it. Whereas in the multi-mode fiber (MMF), multiple modes propagate. Double-clad fibers (DCF) have a single-mode core surrounded by a multi-mode inner cladding, as shown in Figure 2.1a. The inner core is used for OCT imaging, and the outer core is used for laser therapy. DCF can support wavelengths up to 1750 nm [52]. However, the DCF is known to have cross-talk between the single-mode core and inner cladding, which results in ghost images and/or increased noise floor [8].

Maltais-Tariant et al. synchronized the therapy galvanometer and imaging galvanometer of a double-clad fiber system to deliver the treatment laser only in targeted areas [52]. A shutter was used to stop the therapy laser once the targeted therapy duration or coagulation was reached. A diagram of the system setup is presented in Figure 2.1b. It included a reflective collimator based on a 90° off-axis parabolic mirror, which contributed to avoid damaging the OCT system with the therapy laser's high power. Ghost artifacts induced by the DCF were reduced by splicing an extra 1 m of Single Mode Fiber (SMF) between OCT and DCF. The process was demonstrated on an ex vivo rat tongue and abdominal muscles for ablation depths ranging from 500 to 1000 µm.

Jivraj et al. coupled an OCT signal into the signal core of a fiber laser using a commercially available pump combiner [42]. The tilted fiber gratings were used to protect the OCT system from the high-power laser. The saturation of the OCT's balance detector by the reflected leakage light from the combiner was avoided by firing ablation pulses during the latter half of the A-line period, when no OCT data are being gathered. The system demonstrated the feasibility of the method, but the following drawbacks were observed: (1) OCT sample beam dispersion due to the fiber laser; (2) attenuation of

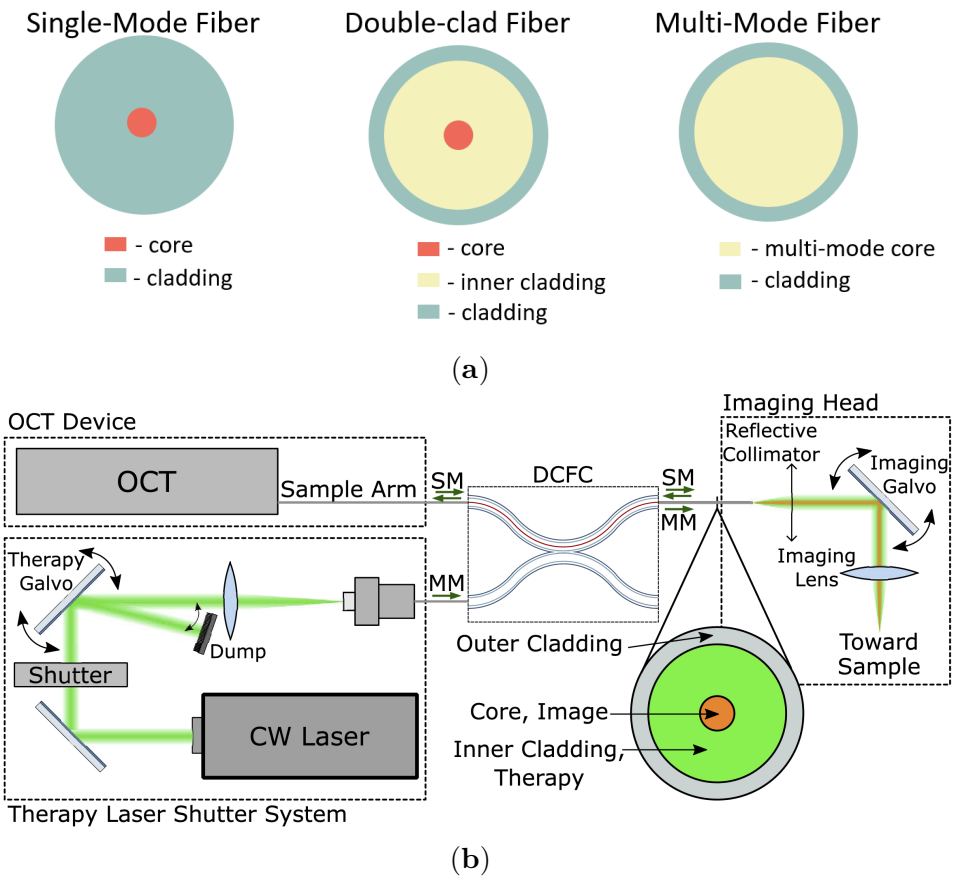

Figure 2.1 OCT-guided laser surgery systems based on double-clad fibers. (**a**) Different types of optical fibers: single-mode fiber (SMF), double-clad fiber (DCF) and multi-mode fiber (MMF). (**b**) Experimental setup proposed for co-localized OCT imaging and laser therapy, used with permission from [52] © The Optical Society.

OCT due to the large mode area of the fiber laser; and (3) bending loss. In addition, the system did not incorporated any steering control of the end of the fiber.

Chang et al. described the use of en-face OCT maps to classify tumors and non-tumor tissues based on Otsu's method [18]. Then, by combining this information with laser ablation units, they generated driving signals for the therapy laser. In the system, the use of a DCF avoided the need for image registration. However, the filters used to avoid high-power backscattered radiation attenuated the OCT signal, affecting image quality. An additional limitation included the fact that the system did not consider potential patient motions between the acquisition of OCT images and the subsequent laser therapy.

2.3 Dichroic Mirror (DM)

A dichroic mirror allows light of a certain wavelength to pass through, while light of other wavelengths is reflected. They are made by alternating layers of optical coatings with different refractive indices. The interfaces between these layers produce phased reflections, selectively reinforcing certain wavelengths of light. Reflective collimators are typically based on a 90° off-axis parabolic mirror, which has a focal length that remains constant over a broad wavelength range. In the studies reviewed in this section, dichroic mirrors and reflective collimators are used to align OCT and high-power lasers. This combination of optics is useful to avoid damaging the OCT setup with the intense backscattered high power laser [83].

Zhang et al. investigated the potential of OCT-guided laser surgery in cochleostomy (inner ear surgery) [89]. This procedure involves drilling the cochlea while avoiding damage to the cochlear endosteum. However, due to the small diameter and thickness of the endosteal layer, the drilling process reaches the limits of human capabilities. Thus, the researchers used OCT images to control the pulse position and duration of a high-power laser in closed-loop, as illustrated in Figure 2.2. This enhanced the precision of the operation, preventing laser exposure to the critical structure. Nonetheless, detection of the critical structure was difficult due to the presence of the highly scattering bone. Therefore, they developed a speckle averaging technique called History Compounding. Finally, they also used OCT as an accurate optical tracking system by locating small laser-ablated landmarks surrounding the cochleostomy. However, open challenges for clinical translation were still identified. For example, the experiments demonstrated that 100% protection of the endosteum was not guaranteed, as the channel bottom penetrated the “stop surface” in some points. In addition, the critical structure segmentation was semi-automatic and OCT imaging, processing, and CO_2 laser control were conducted by three independent software packages. This resulted in high operation complexity, which may lead to increased surgical time.

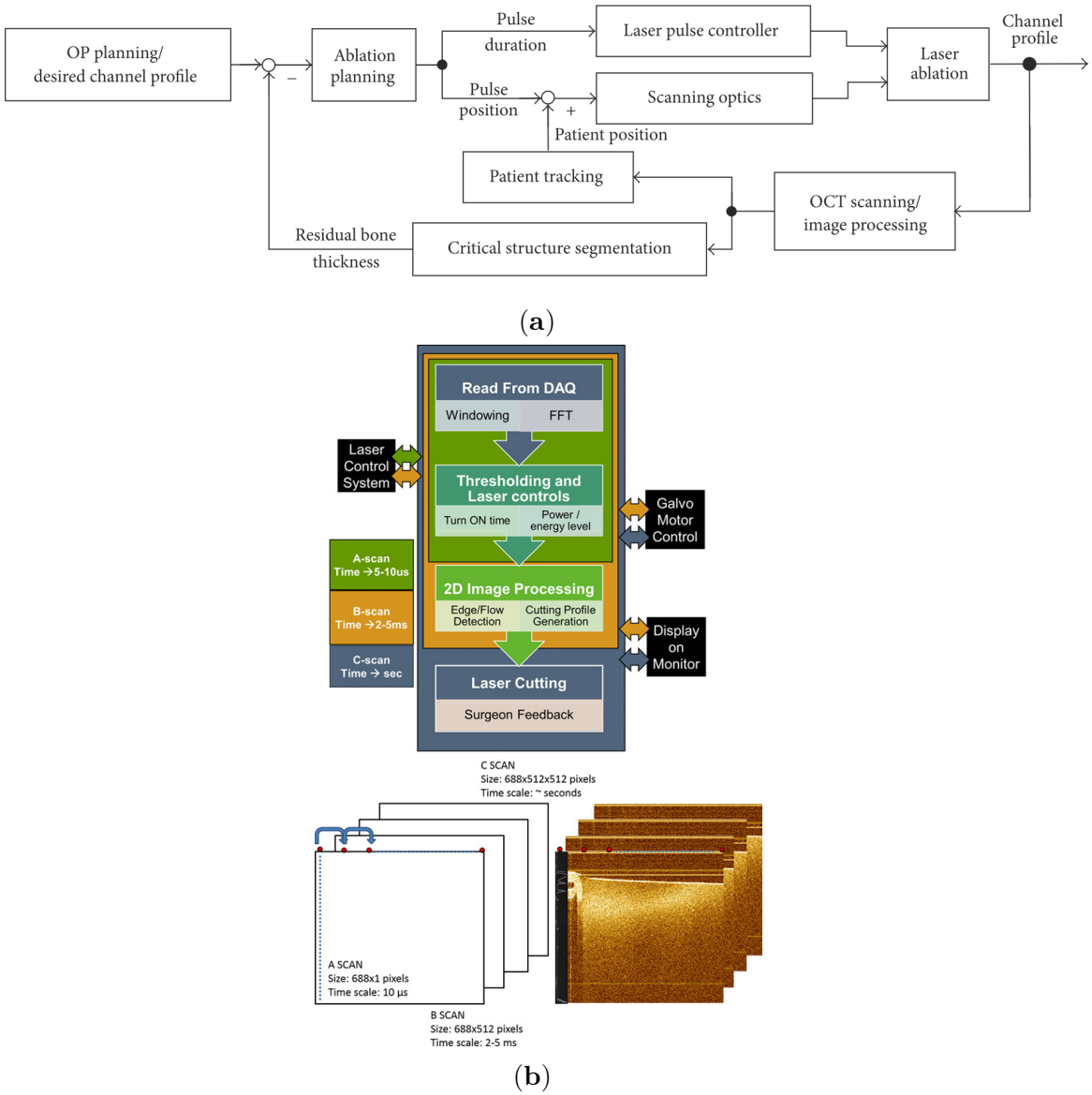

Figure 2.2 OCT-guided laser surgery systems based on dichroic mirrors. (**a**) Control loop scheme of the OCT-guided laser cochleostomy by [89] is licensed under CC BY 3.0. (**b**) Data flow and representation in the smart laser surgical system, used with permission from [43] © The Wiley .

Katta et al. used prior OCT images to obtain a laser ablation pattern avoiding blood vessels in a tissue phantom [43]. As shown in Figure 2.2, the laser power and its ON/OFF times were dynamically controlled based on a "blow-off" model during A-scan OCT. However, the field curvature in the scanning element and the laser's finite ON/OFF time resulted in an uneven cut. Recently, they used prior OCT images to compute the

three-dimensional tumor margin and angiography images to guide the coagulation and ablation steps in a brain cancer surgery of a xenograft model [44].

In 1999, Boppart et al. used OCT to image 65 sites on five ex vivo rat organ tissues before, during, and after laser ablation [11]. Following imaging, tissue registration was achieved by histologic processing to confirm the ablation site's morphology. They identified that the carbonized tissue layer rapidly absorbs and scatters both the incident argon laser and OCT imaging beams. Because of this, OCT imaging penetration is reduced and shows shadowing artifacts. The same group performed another study on prostate ablation [12]. In this case, it was noted that the loss of cell viability in the in vitro specimens reduced the contrast when compared to in vivo studies. It was also suggested that the presence of blood will further reduce OCT imaging penetration.

Leung et al. ablated ex vivo cortical (outer bone) using pulsed lasers while imaging in real time [48]. The intense backscattered light from the ablation process was avoided using inline coherence imaging (ICI). The ICI is analogous to M-mode ("motion-mode") ultrasonography. This provided the depth information in real time, avoiding tomography. They studied both the thermal and the ultrashort pulsed regimes of the ablation laser, finding that carbonization from the thermal regime affected image quality.

Ohmi et al. studied tissue laser ablation using OCT [60]. For this, they controlled the ablation lasers using an electronic shutter and used OCT to image the ablation crater. Laser ablation and OCT imaging were repeated following a sequential automatic procedure. Using this method, they performed experiments on human teeth and chicken bones. The depth of the crater and the ablation rate was determined from the OCT images. The ablation rate was estimated as 0.21 μm per pulse on the human tooth. However, the resolution of the OCT imaging system was only 10 μm, leading to the imprecise measurement. Real-time feedback was not possible due to the slower data acquisition of TD-OCT used.

2.4 Separate Optical Paths (SOP)

Fan et al. integrated a bench-top OCT setup and an endoscope systems with fiber laser to create a proof-of-concept system for OCT diagnostics and guidance during laser surgery [28]. In the system, the OCT beam and the ablation laser have separate optical paths, but they are manually aligned using the endoscopic images. In this setup, 2D-OCT is initially used to generate a tumor map and then to monitor the laser ablation. The results demonstrated the potential of OCT-guided laser surgery for accurate tumor

resections. Nonetheless, significant challenges remained open to achieve full integration of OCT and laser ablation into an endoscopic system.

Li et al. fastened an electromagnetically actuated forward scanning OCT probe of 0.5 mm diameter with a hollow waveguide to permit co-planar ablation and imaging [51]. The waveguide's tip had a CaF_2 window to align the laser with the OCT signal, as shown in Figure 2.3. The system was tested in gelatin and ocular tissues. Real-time 1-D OCT images were used to monitor the lasing operation to avoid ablating critical structures. The combined probe functioned as long as the surface being ablated was within the scanned volume, approximately 3 to 5 mm from the probe tip.

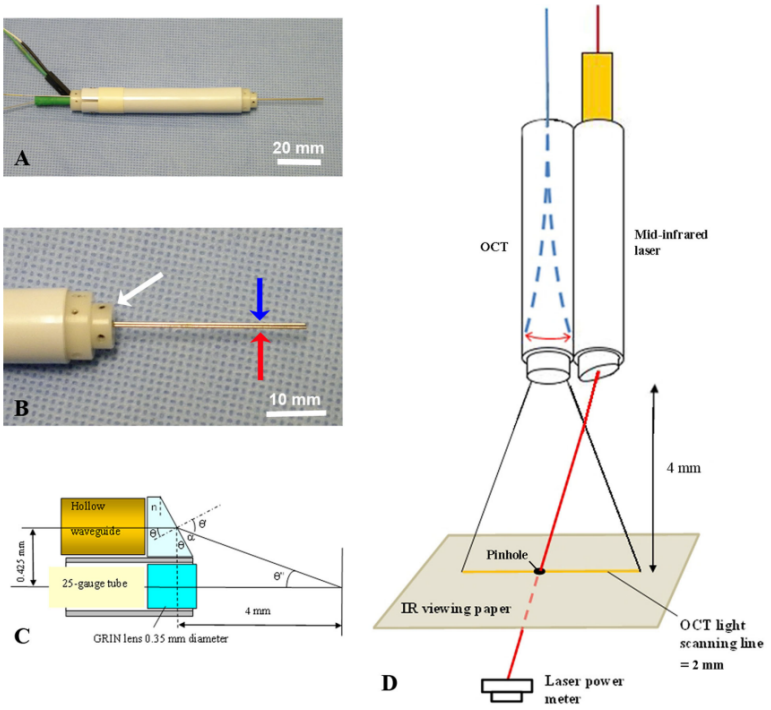

Figure 2.3 Combined miniature B-scan OCT and surgical laser in an intraocular probe, used with permission from [51] The Wiley. (A) Picture of the device. (B) Magnified picture showing the support strucutre (white arrow) and the combined OCT (red arrow) and laser (blue arrow) probe tip. (C) Details of the optical components that allow co-planar ablation and imaging. (D) Diagram of the combined OCT and laser probe tip showing the scanning optical fiber.

2.5 Why SOP?

Regarding OCT-guided laser surgery, it is clear from this review that three major factors should be considered when designing such systems. First, the OCT detector should not be saturated with the backscatter reflected light from the ablation laser. This can be achieved by using a filter or by synchronizing the timing of the OCT imaging and the therapy laser pulse, as described in the papers reviewed herein.Second, the focusing optics have to be carefully designed considering both lasers. Mirrors are a good choice as they provide broadband and high thermal load. However, they make the integration into an endoscopic system more challenging. Nonetheless, the use of MEMS technology may be a viable solution to these miniaturization and integration challenges. Finally, the three major methods to co-align the OCT laser and a therapy laser include the use of a double-clad fiber (DCF), a dichroic mirror, or separate optical paths. Although the DCF allows for easier endoscopy integration, the laser power is limited. It is also known to have crosstalk between the single-mode core and inner cladding, which results in ghost images and/or increased noise floor [42]. The Dichroic mirror on the hand can easily handle power laser, but it is difficult to miniaturize. It will also require redesigning the existing high-power surgical setup, increasing the cost and slowing the clinical adaptation. On the contrary, a Separate Optical Path (SOP), in which two separate probes are used, provides easier clinical adoption with the existing surgical workflow. Hence, this work will focus on the separate optical path.

2.6 Computational Sensing

Despite the numerous advantages offered by these optical imaging modalities, adapting them from a traditional bench-top setup to a clinical workflow presents numerous challenges, such as shallow depth of focus, short working distance, motion blurring and slow acquisition, etc. Computational sensing is a rapidly evolving area within the signal processing field that aims to tackle the above-mentioned problems by co-designing hardware and software. Unlike the traditional methodology of designing hardware and software separately, here, the design is approached synergistically by leveraging computational techniques such as optimization, machine learning, and hardware design to achieve performance previously considered impossible. One such example is compressive sensing which samples below the Nyquist sampling frequency and reconstructs the original signal by leveraging computational resources. Typically, compressive sensing is

implemented by Micromirror Device (DMD) or coded aperture in single-pixel cameras. The application of CS to Laser Scanning Imaging is limited.

Chapter 3

Compressive Image Scanning Microscope

Abstract

A novel approach to implementing compressive sensing in laser scanning microscopes (LSM), specifically in image scanning microscopy (ISM), is presented, utilizing a single-photon avalanche diode (SPAD) array detector. This method overcomes two significant challenges in applying compressive sensing to LSM: the computational time for the sampling matrix and the quality of reconstructed images. A fixed sampling strategy is employed, which skips alternate rows and columns during data acquisition, reducing the scan points by a factor of four and eliminating the need for computing different sampling matrices. The parallel images generated by the SPAD array are exploited to enhance the quality of the reconstructed compressive-ISM images compared to standard compressive confocal LSM images. The results showcase the approach's effectiveness in producing higher-quality images with reduced data acquisition time and its potential benefits in minimizing photobleaching.

3.1 Introduction

Compressive sensing allows the reconstruction of high-dimensional (N) signals x from low-dimensional (M) measurements y, as long as the signal is sparse on a particular basis, such as wavelet or shearlet ($N > M$). Generally, solving for x in $y = Ax$ is an ill-posed problem, meaning there is no unique solution or the solution is not robust to small data perturbations [27]. Therefore, a regularizer term (ϕ) is introduced to solve it.

$$\arg\min_{x} \ \phi(x) \quad s.t. \quad Ax = y \tag{3.1}$$

Compressive sensing is commonly implemented using a Digital Micromirror Device (DMD) or coded aperture in single-pixel cameras. Each row of A corresponds to a unique binary mask, as illustrated in Fig. 3.1. M unique binary masks sample the object of interest sequentially to obtain each element of the measurement matrix Y [86], [25].

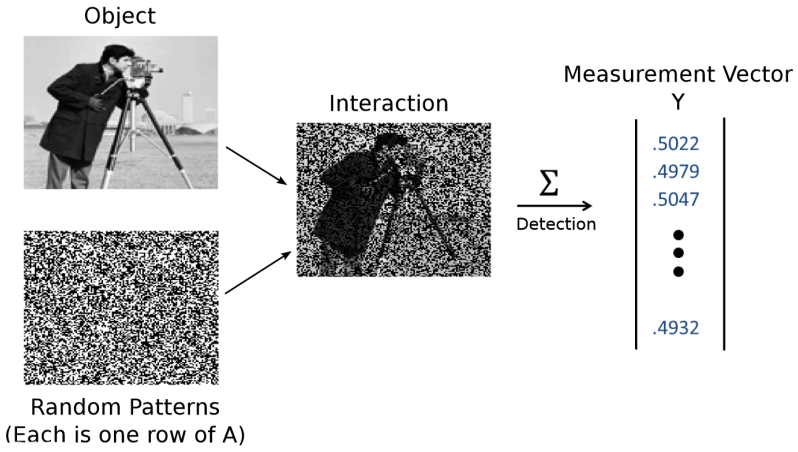

Figure 3.1 Standard optical architecture for compressive sensing based on the single-pixel camera.

In Laser Scanning Microscopes (LSM), images are formed by scanning point-by-point, which is the basis for multiple imaging modalities like confocal microscopy, image scanning microscopy (ISM), and optical coherence tomography. Compressive sensing could enable higher temporal resolution and reduced photobleaching through efficient sampling. However, most existing laser scanning hardware doesn't use encoding devices like DMD or coded aperture, limiting the application of compressive sensing in such configurations.

Pavilion (2016) was the first to implement compressive sensing on such a configuration [63], using the point spread function (PSF) of the optical setup as a smoothing function,

which led to a reduction of confocal fluorescence measurements by 10-15 times. In 2018, Francis *et al.* [31] used a multi-resolution approach for better reconstruction in confocal images, but both the quality, and time, of the reconstructed image is lower than standard TVAL3 solver [72]. Further details about TVAL3 are explained in section 3.2.2. In 2021, Hu *et al.* improved the speed in Raman imaging by sampling only the region of interest and avoiding scanning the background substrate, but the sampling matrix A needed to be recomputed for every new image [38]. The practical application of compressive sensing to laser scanning imaging modalities is limited by three main factors: (1) the solver's reconstruction time, (2) the time to compute the sampling matrix (A), and (3) the quality of the reconstructed images.

In this work, the focus is on addressing the latter two issues. Initially, alternate rows and columns are skipped during data acquisition, which reduces the number of scanned points by a quarter, as detailed in section 3.2.3. This process establishes a fixed sampling matrix A for different images, removing the necessity to compute various A matrices. Secondly, the parallel images produced by the Image Scanning Microscope (ISM) on the single-photon avalanche diode (SPAD) array detector [15] are utilized to enhance the quality of reconstructed images. To the best of the author's knowledge, this represents the first instance of implementing compressive sensing on an image-scanning microscope (ISM).

3.2 Materials and Methods

3.2.1 Simulation Setup and Ground Truth

The open-source ISM simulation software, BrightEyes-ISM, is utilized to create ISM images. Initially, a 2D point spread function (PSF) is generated for each element in the SPAD array detector as shown in Fig. 3.2. The simulation space is configured with a pixel size of 25 nm, detector element size of 50 nm, detector element pitch of 75 nm, and the total magnification of the optical system is set to 500. The PSF is simulated for excitation and emission wavelengths of 640 nm and 660 nm, respectively. Subsequently, synthetic tubulin phantoms are generated as depicted in Fig. 3.3 using the BrightEyes-ISM package [92]. These phantoms mimic various tubular microscopic structures frequently imaged by ISM, with variations in the number of tubular filaments and the thickness of each filament to produce a diverse set of samples. These phantoms are then convolved with

the PSF for each element of the SPAD array detector to produce several parallel images. Finally, Poisson noise is added to the resulting images as illustrated in Fig. 3.4.

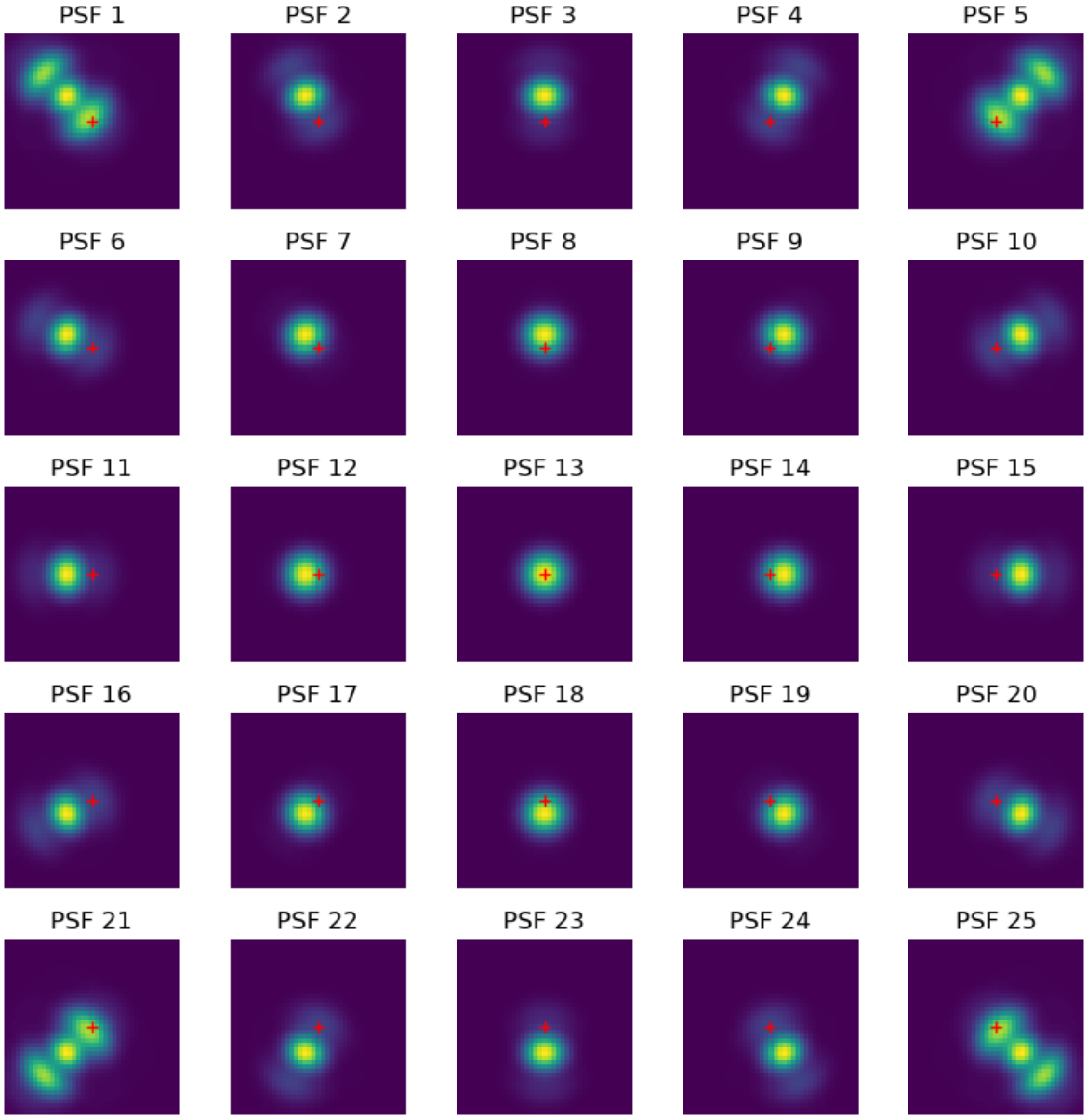

Figure 3.2 Point Spread Function (PSF) for each element in the SPAD array detector

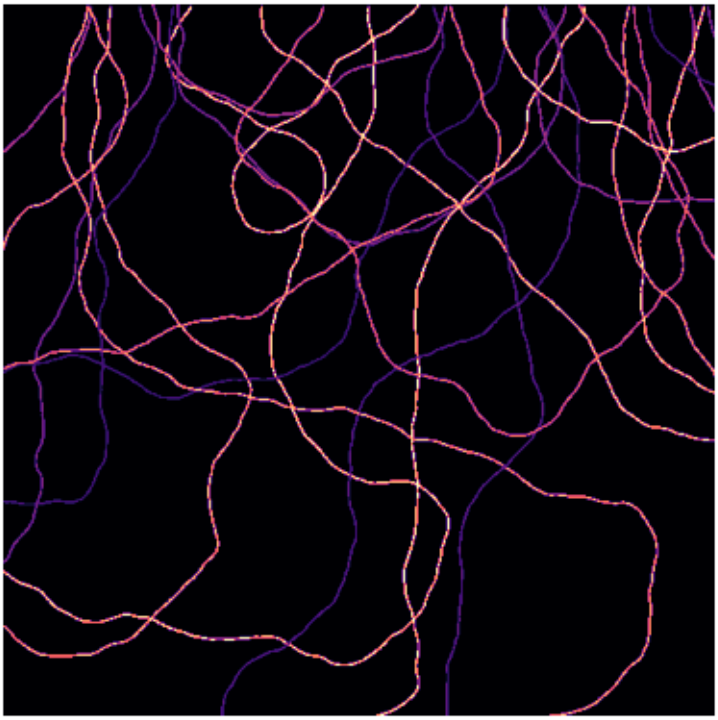

Figure 3.3 Phantom

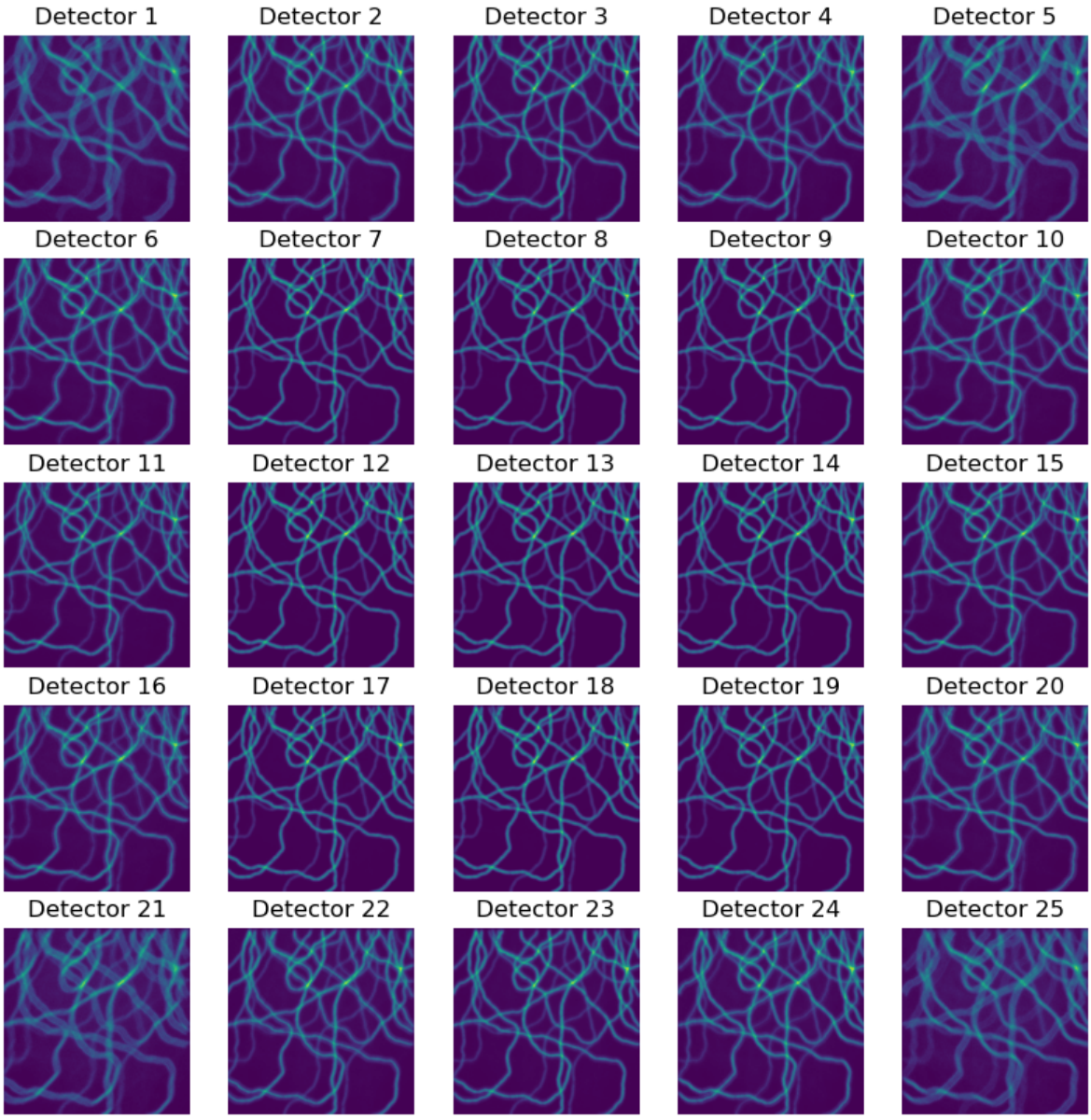

Figure 3.4 Parallel Micro-image on the Detector

3.2.2 Compressive Sensing Reconstruction

There are two common choices for the regularizer $\phi(x)$ in Eq. 3.1 [29]: (1) L_1 norm $||x||_1$ and (2) Total Variation (TV) norm $||x||_{TV}$:

$$||x||_{TV} := \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} |x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}| \tag{3.2}$$

TV regularization is more appropriate for image reconstruction because it preserves edges and boundaries [90]. A comprehensive review of various algorithms for solving L_1 norm and TV norm can be found in [72]. Based on the provided information, TVAL3 [49] was chosen due to its fast reconstruction time.

3.2.3 Scanning Pattern

Compressive sensing enables the reconstruction of the original signal from 10-15% of the sample through random scanning. However, implementing such an arbitrary scanning system on the galvo mirror poses significant challenges. As a solution, a fixed sampling strategy was adopted by scanning alternate rows and columns, as shown in Fig. 3.5, achieving an overall compression ratio of 25%. An additional benefit of this method is the consistency of the sampling matrix A across different images, unlike in random scanning, where different A matrices must be computed for different samples. The sampled data (y) and the sampling matrix (A) are then used to reconstruct the image using the TVAL3 solver.

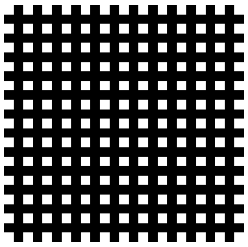

Figure 3.5 Sampling strategy: Alternate rows and columns are skipped. White indicates sampled location, and black shows the unsampled location.

3.2.4 ISM Reconstruction

The ISM image is generated by combining all parallel images from the SPAD array detector. Adaptive Pixel Reassignment (APR) [16], [17], available in the BrightEyes-

ISM package, is employed for this merging process. This procedure is applied to the corresponding reconstructed images from compressive sensing to obtain the Compressive-ISM images. For the Confocal LSM image, the central element of the SPAD array corresponds to the standard image, and in this case, the reconstructed image from compressive sensing is used directly.

3.3 Results and Discussion

The quality of reconstructed images for both Compressive ISM and Compressive Confocal LSM is evaluated by computing the relative error using Eq. 3.3:

$$\text{Relative Error} = \frac{\|I_{fs} - I_{cs}\|_{\text{F}}}{\|I_{fs}\|_{\text{F}}} \tag{3.3}$$

where $\|\|_{\text{F}}$ denotes the Frobenius norm, I_{fs} represents the fully sampled image corresponding to the top row in Fig.3.6, and I_{cs} corresponds to the compressive images corresponding to the bottom row of Fig.3.6. Twenty-five samples are used to calculate the mean and standard deviation, as shown in Table 3.1. Compressive ISM yields better results compared to Compressive Confocal LSM images, as illustrated in Fig.3.6. This improvement can be attributed to the utilization of parallel images generated by the SPAD array. Moreover, our proposed sampling strategy reduces the number of scanned points by a factor of four, leading to faster data acquisition, decreased photobleaching in the sample, and eliminating the need to compute the sampling matrix A for different samples.

Table 3.1 Quality of reconstructed Images

Imaging Technique	**Relative Error**
Compressive Confocal LSM	$14.65 \pm 2.67\,\%$
Compressive ISM	$12.50 \pm 0.71\,\%$

3.4 Conclusion

An efficient compressive sensing method for image scanning microscopy was devised using a SPAD array, incorporating a fixed sampling strategy to reduce acquisition time and

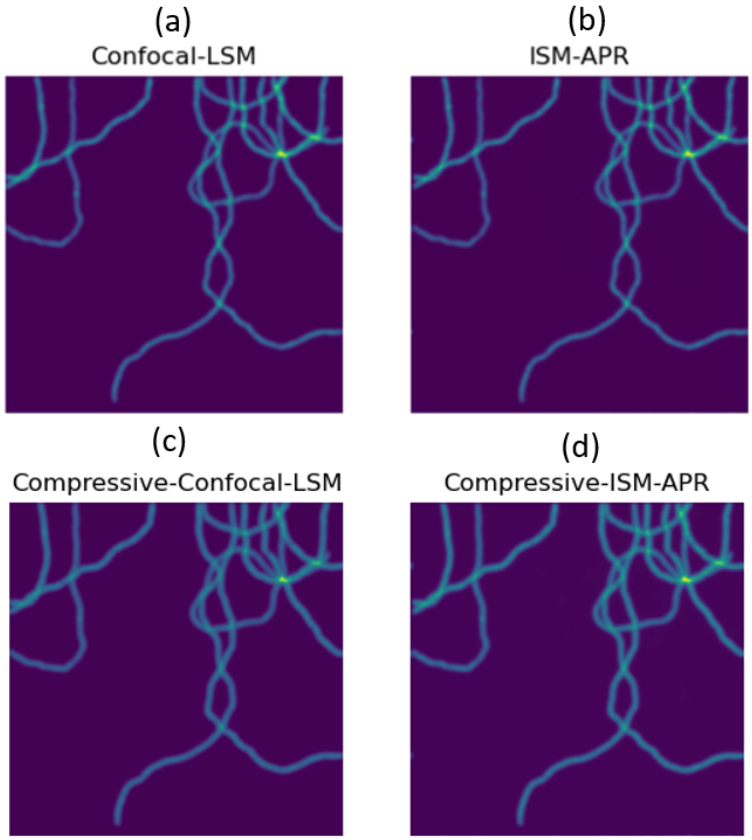

Figure 3.6 (a) Fully sampled CLSM. (b) Fully sampled ISM. (c) Compressive CLSM. (d) Compressive ISM.

negate the need for computing various sampling matrices. The Compressive-ISM reconstruction yielded images of superior quality compared to those obtained through standard Compressive Confocal LSM. This work paves the way for practical applications in laser scanning imaging modalities, improving temporal resolution, minimizing photobleaching, and enhancing image quality. Future research includes hardware implementation, block-based compressive sensing [88] for faster reconstruction, GPU-based solutions for parallel SPAD array reconstruction, and deconvolution-based ISM algorithms [93] to further improve image quality.

Chapter 4

Compressive 3D OCT-Guided Depth Control

Abstract

This study introduces a pioneering approach to determining ablation depths using 3D Optical Coherence Tomography (OCT) in laser microsurgery under the Separate Optical Path (SOP) category. This method overcomes the limitations of conventional 1D and 2D OCT scans by providing a more comprehensive and accurate depth assessment. A significant advancement is the implementation of compressive sensing in OCT hardware, which effectively reduces scanning time without compromising the quality of the depth map. Additionally, a feedback controller was introduced in order to utilize the depth map to precisely adjust the laser exposure time across different tissue types, ensuring accurate ablation.

4.1 Introduction

The SOP works mentioned previously in section 2.4 utilize 1-D and 2-D OCT scans for depth information. When ablating a point a 3D crater hole is produced. A 1-D OCT scan can give drastically different values depending on where the OCT is focused within the ablated crater, leading to incorrect estimates about the depth. In the case of 2D-OCT scans, it also depends on the plane of the scan. Hence, both 1D and 2D OCT data are insufficient since they depend on the scan location. Additionally, 2D image processing must be calibrated to a specific OCT data set. Alternatively, 3D scans offer a higher-fidelity map that is easier to interpret; the large data size ($\approx 2GB$) complicates data analysis; and the acquisition time ($\approx 15s$) results in motion blur. Compressive sensing allows reconstruction of high-dimensional (N) signals x from low-dimensional (M) measurements y, as long as the signal is sparse on a certain basis, like a wavelet or shearlet ($N > M$).

In this chapter, we have three significant contributions:

- To our best knowledge, This is the first application of 3D-OCT scan to determine the ablation depth under the separate optical path (SOP) category. This address the limitations of 1D and 2D OCT scans.
- Secondly, we apply compressive sensing on OCT hardware to find the optimal compression ratio. Using the optimal compression ratio we reduced the scanning time without significant loss in the quality of the depth map.
- Finally, we use the depth map to control the exposure time (laser on/off time) to accurately ablate a given depth across different tissue types in a feedback controller.

4.2 OCT Hardware Sampling

To the best of the author's knowledge, prior research on compressive sensing on OCT, notably by Haydar et al. (2020) [37] and McLean and Hendon (2021) [56], primarily concentrated on synthetic sampling. In synthetic sampling, data was initially scanned through a full scan. Then, this full scan was re-sampled at various compression ratios. The re-sampled data was used to reconstruct the full image. However, this method of synthetic sampling does not address the practical need to reduce scanning time. Unlike existing scanning methods, our approach described in Section 3.2.3 is implemented directly on the hardware, reducing scanning time. Moreover, we also vary the compression ratio

from 0.5, 0.8, 0.9, and 1.0. for both B-Scan and C-scan and comparing the reconstruction quality and net reduction factor across 16 different compression conditions.

4.2.1 Experimental Setup

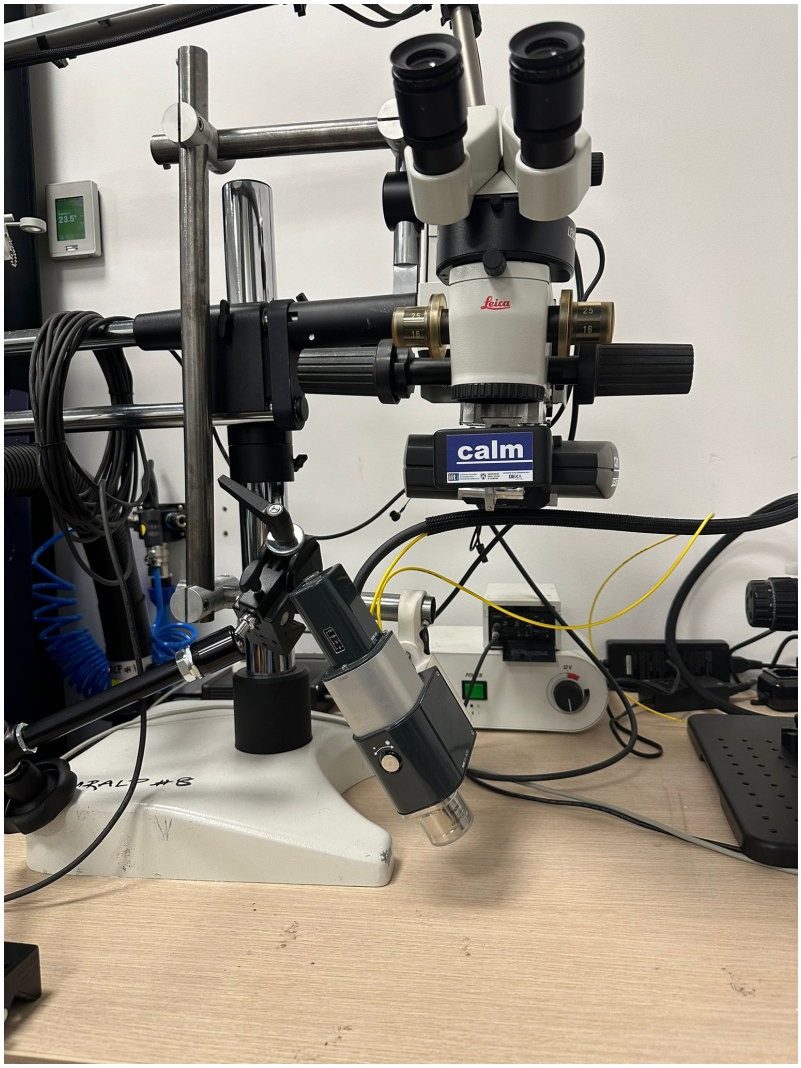

Figure 4.1 OCT-Guided Surgical Laser Setup

In our study, we ablated ex-vivo tissue samples using our laser microsurgery setup. Our laser microsurgery setup consists of a customized motorized laser micromanipulator (CALM), a Leica M651 surgical microscope (Leica, Wetzlar, Germany) with a built-in Leica2 camera, a UniMax 2000EWD (Reliant Technologies, Inc, Forster City, California) laser focusing system, a DEKA SmartXide2 C60 CO2 surgical laser (DEKA M.E.L.A. S.r.l., Calenzano, Italy) with a red aiming laser. We used HiScan (DEKA M.E.L.A.

S.r.l., Calenzano, Italy) to scan the high-power beam [2]. All the trials were done with 3 watts power. The Thorlabs TELESTO (TEL300) system imaged the ablated samples using a spectral domain (SD-OCT) device with an LSM04 objective lens and a nominal central wavelength of 1310 nm. ThorLabs's $C++$ Spectral SDK, version 5.2, is utilized to acquire the data.

4.2.2 Data Acquisition Pipeline

At the start of the process, the OCT device, probe, and processing handles were set up. Next, the scan parameters, such as the acquisition mode and OCT device speed, were set up. Table 4.1 shows the standard parameters.

Table 4.1 Standard Data Acquisition Parameters

Parameter	**Value**
Length of B-Scan	5.0 mm
Width of Volume	5.0 mm
OCT Device Speed	146 kHz
Acquisition mode	Asynchronous Continuous
A-Scan Averaging	3
B-Scan Averaging	N/A

Within each scan, we calculate the number of A-scans per B-scan and B-scans per volume as follows:

$$\begin{aligned} \text{AScansPerBScan} &= 300 \times \text{BscanCR} \\ \text{BScansPerVolume} &= 300 \times \text{CscanCR} \end{aligned} \tag{4.1}$$

BscanCR and CscanCR correspond to the compression ratios of B-Scan and C-Scan, respectively. The compression values chosen are 1.0, 0.9, 0.8, and 0.5 for both B-Scan and C-Scan, resulting in 16 unique scans for each sample. We process the raw data to obtain volumetric data. the volumetric data consisted of 2D array of A-scan. Each A-Scan is filtered by Gaussian filtering and threshold for maximum intensity to find the surface, as shown in Fig. 4.2.

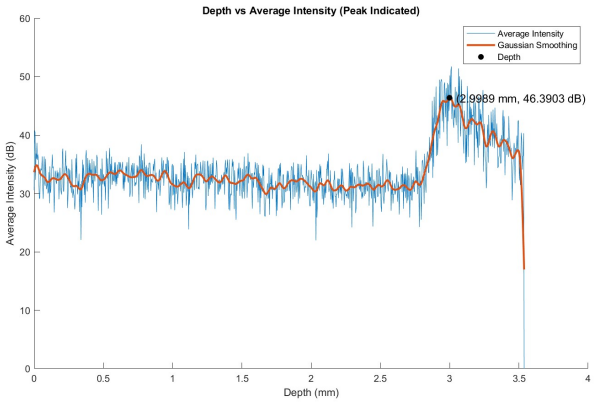

Figure 4.2 Depth Estimataion from Single A-Scan

If both compression ratios are 1.0, the data is at full resolution and is the ground truth. The number of A-scans and B-scans and the data acquisition parameters are limited by the rotational speed of the galvo mirror, which scans the OCT laser beam.

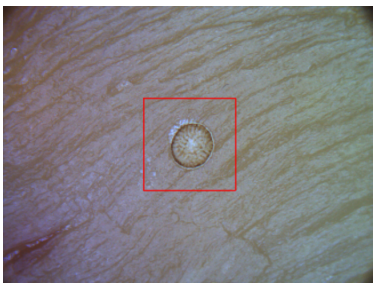

Figure 4.3 Red box indicates the scanned region of the ablated sample

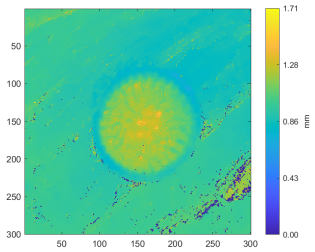

Figure 4.4 Raw Full Surface Map

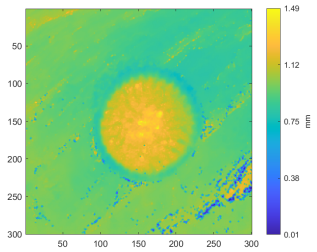

Figure 4.5 Filtered Full Surface Map

4.2.3 Upsampling Strategy

When we perform compressive scanning, we obtain a map with lower resolution as shown in fig. 4.6.

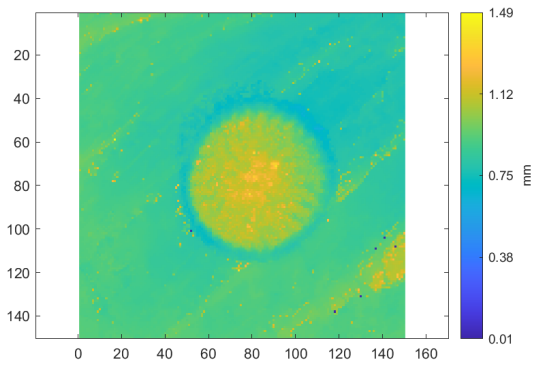

Figure 4.6 Sparse Depth Map of compression ratio 0.5 for both B-Scan and C-Scan

The resolution of this map depends on the compression ratio used during the data acquistion. During compressive reconstruction, missing pixels were estimated at a higher resolution. The process of transforming scanned pixel locations from low resolution to high resolution, where pixels are to be estimated, is called the upsampling process. First, a column-wise upsampling of a compressed matrix by dividing it into intervals and padding zero columns between each interval based on the B-scan compression ratio (BscanCR). The BscanCR, as specified in Table 4.2, determines the size of the interval and the number of zero columns for padding.

Table 4.2 Upsampling Parameters

CR Ratio	**Interval Size**	**Padding Size**
1.0	N/A	N/A
0.9	9	1
0.8	4	1
0.5	1	1

The process is repeated for rows based on the C-scan compression ratio (CscanCR) which decides the interval and padding sizes. This produces upsampled data as seen in Fig. 4.7.

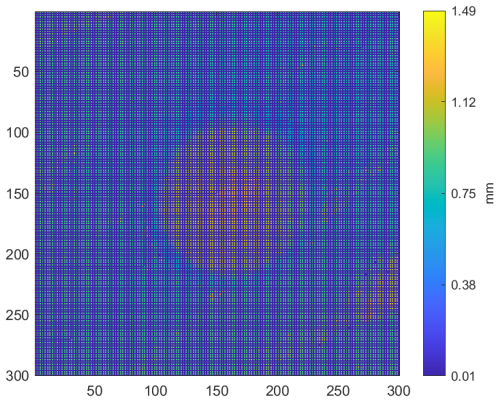

Figure 4.7 Upsampled Surface Map

4.2.4 Compressive Scanning & Reconstruction

The upsampled matrix facilitates the derivation of the sampling matrix (A) and the compressed dataset (y). Subsequently, the sampled dataset (y) and the sampling matrix (A) are employed in conjunction with the compressive sensing algorithm, TVAL3, as delineated in Sections 3.1 and 3.2.2, to reconstruct the image (x), depicted in Figure 4.8. This reconstruction process was executed using MATLAB (R2022b) on an Intel® Xeon® CPU E3-1270 v5 @ 3.60GHz, equipped with 34.29 GB of RAM and Windows 10 Pro (64-bit).

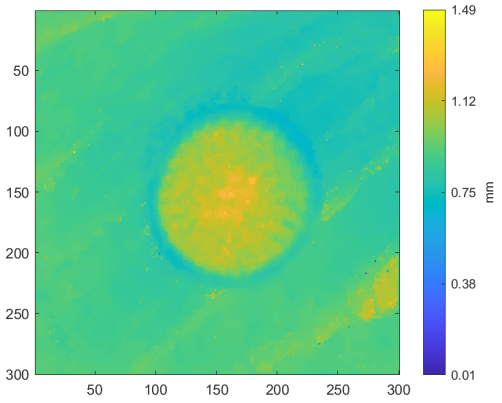

Figure 4.8 Reconstructed Surface Map

4.2.5 Evaluation Metrics

In this chapter, the metric described in Section 3.3 is used to quantify the reconstruction quality. As mentioned in the section 4.2, existing research predominantly employs synthetic sampling, thereby prioritizing reconstruction quality as the primary evaluation metric. Although reconstruction quality is a necessary metric, it proves insufficient for hardware-based compressive sensing applications. Consequently, we introduce a novel metric termed the "Net Reduction Factor," which accounts for scanning and reconstruction times. This metric is crucial, as it highlights whether the increased reconstruction duration does not negate any reduction in scanning time

$$\text{Net Reduction Factor} = \frac{T_{fs} - (T_{cs} + T_R)}{T_{fs}} \tag{4.2}$$

where T_{fs} represents the full scanning time, T_{cs} is the compressive scanning time, and T_R is the reconstruction time. The Net Reduction Factor, which ranges from 0 to 1 is used to quantify the reduction in scanning time.

4.2.6 Results and Discussion

The evaluation of the compressive scan method focused on two key aspects: (1) the quality of the reconstructed depth map and (2) the overall reduction in scanning time, considering both scanning and reconstruction time.

Reconstruction Quality

A reconstruction error of zero signifies an exact match between the compressive depth map and the fully scanned depth map, implying no quality degradation due to compression. This means that the fully scanned image I_{fs} and the compressive scanned image I_{cs} are identical. On the other hand, a reconstruction error of 100% means a poor reconstruction quality, indicating a significant deviation of the compressive depth map from the fully scanned depth map. In essence, a lower reconstruction error is preferable as it signifies a closer match between the compressive and fully scanned depth maps, which suggests better quality preservation during the compression process.

Table 4.3 presents the reconstruction error across various compression ratios in both dimensions. Initially, it appears counterintuitive that when both B-Scan and C-Scan are lowest, that is 0.5 the reconstruction error is lowest at $13.41 \pm 3.56\%$. Whereas other compression ratios have 50% to 90% error. This is primarily attributed to error caused

Table 4.3 Reconstruction Quality

		C-Scan Compression Ratio			
		0.5	0.8	0.9	1
B-Scan Compression Ratio	0.5	13.41 ± 3.56%	99.95 ± 0.10%	99.92 ± 0.15%	99.88 ± 0.22%
	0.8	99.81 ± 0.32%	99.69 ± 0.45%	99.50 ± 0.59%	99.15 ± 0.87%
	0.9	98.48 ± 1.36%	97.43 ± 1.76%	95.43 ± 2.89%	91.88 ± 3.65%
	1	85.65 ± 4.16%	74.07 ± 4.79%	53.60 ± 4.06%	N/A

by moving the location of the pixel during the upsampling process. To illustrate, consider an simple toy example of scanning A-scans along a line, as shown in Fig. 4.9.

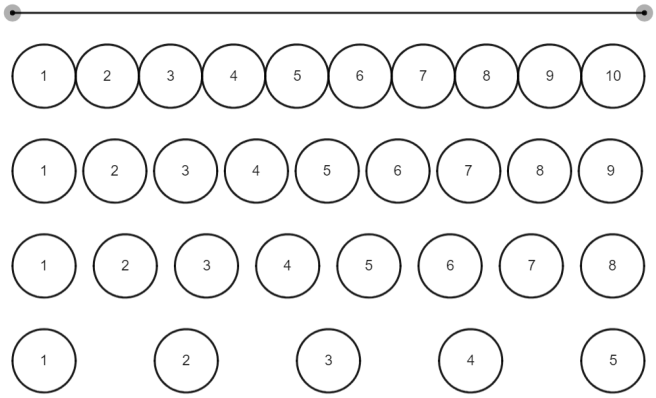

Figure 4.9 Actual scan location

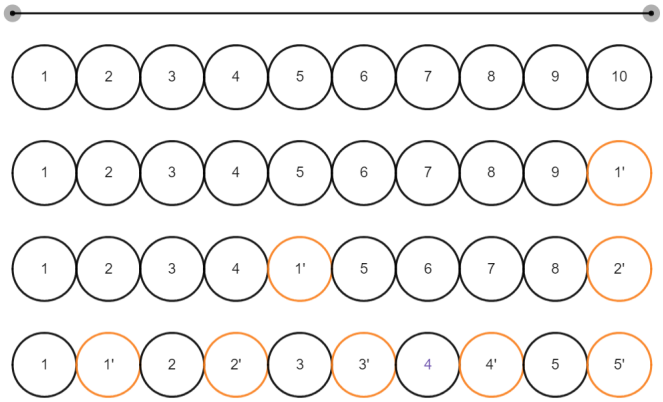

Figure 4.10 Assumed scan location during upsampling. Orange circles are padded pixels which are estimated; black circles are the scanned pixels.

In this example, the first row indicates 10 A-scans along a line, which represent the ground truth or reference. The second row, with a 0.9 compression ratio, comprises 9 A-scans; the third row, at a 0.8 ratio, includes 8 A-scans; and the fourth row, at a 0.5 ratio, consists of 5 scans along the line. These scans are uniformly spaced. However, as depicted in Fig. 4.10, the upsampling process reveals a discrepancy between the actual scan locations and their up-sampled location counterparts for all compression ratios. This mismatch is minimized at a compression ratio of 0.5, leading to a lower reconstruction error. In the future, precise control over the scan locations will be crucial for improved estimation. Unfortunately, ThorLabs $C++$ Spectral SDK, version 5.2, provides functionality to specify the length of line segment to be scanned and number of scans. However, it does not allow precise control over the exact location of each scan.

Finally, when both compression ratios are 0.5, the lowest reconstruction error is 9.85%, and the highest is 16.97%. A standard deviation of 3.56%. This is mainly due to the surface level variation of the tissue sample; if the sample is highly uneven, some portion of the sample is out of focus, thus a higher error.

Scanning Time

Tables 4.4 and 4.5 represent the reconstruction and scanning times across various compression ratios in both dimensions.

The full scanning time is 16.99 seconds, which consistently decreases as the compression ratio increases from 0.9 to 0.5. At a compression ratio 0.5 for both axes, the lowest scan time achieved is only 4.34 seconds. The reconstruction time varies from 633 milliseconds to 893 milliseconds. As the compression ratio increases, the number of pixels to estimate

Table 4.4 Reconstruction Time (milliseconds)

		C-Scan Compression Ratio			
		0.5	0.8	0.9	1
B-Scan Compression Ratio	0.5	893.99 ± 79.86	749.61 ± 48.00	693.96 ± 40.04	649.84 ± 37.55
	0.8	706.86 ± 55.18	718.27 ± 20.33	708.34 ± 39.93	666.45 ± 48.89
	0.9	672.84 ± 33.53	690.92 ± 30.34	704.02 ± 48.63	642.23 ± 26.82
	1	633.09 ± 28.36	649.49 ± 27.04	653.55 ± 19.25	N/A

Table 4.5 Scanning Time (Seconds)

		C-Scan Compression Ratio			
		0.5	0.8	0.9	1
B-Scan Compression Ratio	0.5	4.34 ± 0.12	6.78 ± 0.20	7.51 ± 0.14	8.33 ± 0.16
	0.8	5.96 ± 0.14	9.41 ± 0.08	10.55 ± 0.09	11.69 ± 0.13
	0.9	6.54 ± 0.05	10.21 ± 0.08	11.44 ± 0.12	12.84 ± 0.10
	1	7.05 ± 0.04	11.00 ± 0.13	12.56 ± 0.25	16.99 ± 0.18

increases, leading to a corresponding increase in reconstruction time. A net reduction factor of 0 indicates that there is no decrease in scanning time, meaning that the time required for compressive scanning and reconstruction time ($T_{cs} + T_R$) is equivalent to the full scanning time (T_{fs}). The closer the Net Reduction Factor is to 1, the greater the time savings shown in table 4.6,

Table 4.6 Net Reduction Factor

		C-Scan Compression Ratio			
		0.5	0.8	0.9	1
B-Scan Compression Ratio	0.5	69.11 ± 0.65%	52.67 ± 1.57%	49.47 ± 1.25%	44.69 ± 1.79%
	0.8	58.20 ± 1.89%	38.32 ± 1.43%	32.40 ± 1.07%	25.68 ± 1.09%
	0.9	56.31 ± 0.69%	34.14 ± 1.00%	27.14 ± 1.03%	20.09 ± 1.05%
	1	54.02 ± 0.51%	31.18 ± 0.94%	22.60 ± 1.17%	N/A

A 69.11% savings in scanning time was achieved when the compression ratios of both B-Scan and C-Scan are 0.5. It is very clear that in terms of both net reduction factor and reconstruction quality, a compression ratio of 0.5 for both B-Scan and C-SCacn remains the optimal choice. This is primarily due to the least upsampling error and high saving in time due to skipping alternate rows and columns.

4.3 Feedback Controller

In this section, the optimal compression ratio of 0.5 is used for both B-scan and C-scan to precisely ablate a given depth of 1mm across different tissue types. Both tissues, chicken breast and beef, were ablated under the same conditions, utilizing a 5-watt power setting and identical Kp and Kd values.

4.3.1 Depth Estimation

The depth map generated by 3D Optical Coherence Tomography (OCT) provides absolute depth information, which is directly not useful. The metric of interest is the relative depth, defined as the distance from the surface to the deepest point of ablation. This is determined by calculating the difference between pre-and post-ablation depth maps. Direct computation of this differential, however, can yield inaccurate estimates due to the presence of noise, which primarily originates from two sources. The first noise source arises from surface irregularities. To counteract this, the region of interest (ROI) is limited to the center of the ablated crater, where the median depth is determined. The second noise source is due to the inclined nature of ablation, which introduces points surpassing the OCT's depth of field. A straightforward binary thresholding approach is applied to the ablated crater to address this issue, effectively minimizing the influence of this noise and enhancing the overall accuracy of the relative depth estimation. Compressive sensing acts as an essential filter, filtering out any excessive noise, as seen in Fig.4.8. The primary reason for this is that the process of under-sampling (25%) and estimating the remaining pixels primarily functions with natural images [5], whereas noises do not qualify as natural images. Consequently, compressive sensing inherently serves as an essential filter by default.

4.3.2 PD Controller

The exposure time (laser on/off time) is calculated based on the current depth using a proportional plus derivative controller, as shown in eq 4.3.

$$\text{Exposure Time} = K_p \cdot e + K_d \cdot \frac{de}{dt} \tag{4.3}$$

where K_p and K_d are the proportional and derivative constants. e is the difference between current depth and desired depth, and $\frac{de}{dt}$ is the rate of change of error. The PD controller was manually tuned by trial and error until the desired depth was reached within two cycles.

4.3.3 Results & Discussion

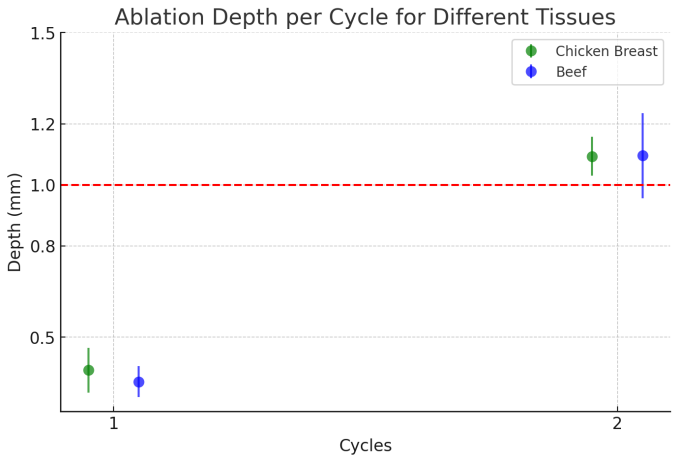

Table 4.7 Ablated Depth Per Cycle

Tissue	Cycle 1	Cycle 2
Chicken Breasts	0.39 ± 0.07	1.09 ± 0.06
Beef	0.35 ± 0.06	1.10 ± 0.16

Irrespective of the tissue type, the overall ablated depth accuracy remained less than 0.1 mm. However, the standard deviation in beef was higher, at 0.16. This might be because the PD controller gains were tuned specifically for chicken breast and not for beef. This issue can be mitigated in two ways: either by having a high-resolution system, i.e., increasing the number of cycles by employing a control system that responds more slowly, thus converging slowly towards the desired depth. Alternatively, using a Model Predictive Controller could help to converge quickly with high precision, since the model provides prior information.

4.4 Conclusion

This work presents a novel approach in laser microsurgery, primarily through the innovative implementation of 3D OCT scanning under SOP category. This addresses the limitations of traditional 1D and 2D OCT scans by accurately determining ablation depths, a crucial factor in surgical precision. Then, we successfully apply compressive sensing on OCT hardware. Our findings reveal that a compression ratio of 0.5 for both B-scan and C-scan strikes the perfect balance, resulting in a substantial reduction

in scanning time by 69.11% while maintaining high-quality depth map. Notably, this approach also enhances noise filtration, eliminating the need for traditional filtering methods. However, the reconstruction error is influenced by the up-sampling process and the depth of focus, suggesting areas for future enhancement. Either software or hardware-based approaches can improve the upsampling methods. Hardware-based approaches could enable low-level API access for precise control of scanning locations. Alternatively, one could explore software-based novel upsampling methods. The depth of focus can be improved with longer objective lenses. However, these objectives are often larger. Hence, other computational sensing methods should explored that could enhance the depth of field without increasing the physical size of the object lens. Additionally, the development of a feedback controller using depth maps for precise laser exposure adjustment has shown effectiveness across various tissue types, potentially enhancing the efficiency and precision of surgical procedures. Looking ahead, integrating more advanced controllers such as (MPC) and adapting our methods to diverse surgical conditions could further expand the applicability of this technology in clinical settings, paving the way for more refined surgical techniques.

Chapter 5

Auto-CALM: Automated Computer-Assisted Laser Microsurgery

Abstract

In the previous chapters, ablation over a given point was performed using a high-speed scanner, but the ablation area remains limited. This chapter extends the capability of CALM to ablate over larger area defined by the surgeon automatically. Auto-CALM allows the surgeon to define the ablation area, which is then precisely ablated by the system while compensating for tissue motions and deformations. This is achieved based on three control blocks: target tracking, laser tracking, and the ablation algorithm. The ablation area, i.e., the target, is defined by the surgeon using a stylus tablet. Then, the target is graphically overlayed on the surgical video. This target is tracked in real time using improvised optical flow. The improved optical flow features a scaling technique for enhanced resilience to tissue deformations and a census transformation for robustness against illumination variations. Based on the laser spot position and target positions, the ablation algorithm generates a trajectory to ablate the target in real-time. This motion compensation increases the accuracy of the system against breathing motions. Auto-CALM was tested on a porcine larynx model simulating breathing, Auto-CALM demonstrated superior performance to manual and other methods such as TAM, achieving high precision under challenging conditions. This system promises significant improvements in surgical ablation tasks, with further studies aimed at clinical translation.

5.1 Introduction

Previously, a computer-assisted laser microsurgery system (CALM) was proposed to enable accurate ablations and resections using an intuitive graphics tablet for precise laser control [2]. CALM is based on a robotic laser micromanipulator and has demonstrated excellent potential for clinical use in transoral laser microsurgery (TLM). The work presented here further enhances CALM with intelligent capabilities by automating the laser movement over given area selected by the surgeon. Also, It adapts to real-time tissue motions and deformations for enhanced laser ablation accuracy. The system components are shown in Fig. 5.1. In this work, automatic capabilities are added to CALM based on the integration of a real-time monocular color vision system. This enables tracking both the laser spot and the target tissue. The laser spot is tracked using both classical approach and pre-trained models. Target tracking combines a scaling strategy with a customized optical flow method featuring pyramid windowing and census transformation to manage tissue deformations and brightness variations as decribed in section 5.3.3. Finally, an controller is proposed to execute user-defined ablation operations based on the real-time laser position and target position. This includes motion compensation, which is implemented to increase precision and robustness against breathing motions. To start the operation, the user draws the desired ablation region using a graphics tablet and a stylus. Then, Auto-CALM performs the entire ablation automatically. This work offers three primary contributions to the advancement of CALM: 1) Laser Spot Tracking. 2) Target Tracking: A method based on optical flow is used to track moving targets. To overcome challenges related to tissue deformation, feature changes, and the assumption of velocity smoothness in optical flow, a scaled lesion contour strategy is proposed. In addition, long-range deformation and brightness variations are addressed using pyramid and census transformations. 3) Ablation Algorithm: This algorithm controls the ablation process on a moving target by updating the laser trajectory in real-time. First, it ablates the contour of the desired ablation area, and then its inside. The internal ablation process uses continuous raster scanning with speed compensation to ensure complete ablation and enhance accuracy. This chapter is organised as follows: First, section 5.2 describes the experimental setup. Section 5.3.1 presents the calibration method that enables laser steering control from a monocular vision system. Then, the methods for laser tracking, target tacking and autonomous ablation are presented in sections 5.3.2, 5.3.3, 5.3.4, respectively. Evaluations metrics are presented in section 5.3.5. Then, the performed experiments are introduced in 5.3.6. Evaluation results of tracking methods

are presented in 5.4. Results from comparative experiments with manual CALM and Track Anything Model (TAM) are presented in 5.4.3. Results from robustness assessment experiments on a porcine larynx model under various conditions of motion and tissue features are presented in section 5.4.4. Finally, conclusions are presented in 5.5.

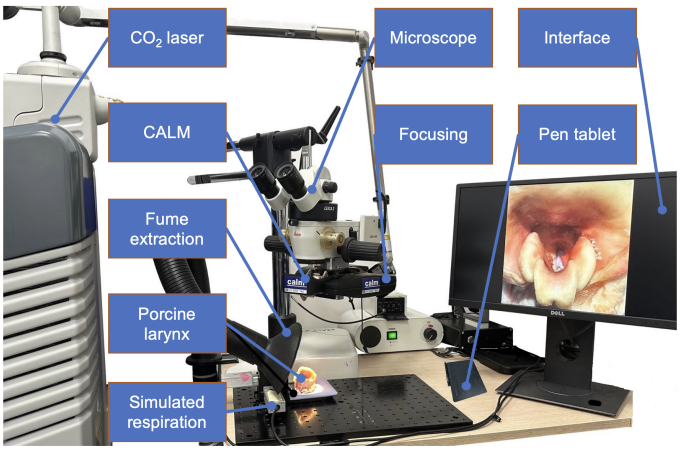

Figure 5.1 Auto-CALM system components.

5.2 Materials

The Auto-CALM system is shown in Fig. 5.1. It includes a custom motorized laser micromanipulator (CALM), a Leica M651 surgical microscope (Leica, Wetzlar, Germany) with a built-in Leica2 camera, a UniMax 2000EWD (Reliant Technologies, Inc, Forster City, California) laser focusing system, a DEKA SmartXide2 C60 CO_2 surgical laser (DEKA M.E.L.A. S.r.l., Calenzano, Italy) with a red aiming laser, a Weller WFE 2ES fume extractor (Weller, Besigheim, Germany), and a Wacom CTL-4100ML graphics tablet (Wacom Co. Ltd., Kazo, Saitama, Japan) used for lesion contour design by the operator. In addition, the setup includes a custom interface circuit for automatic activation of the ablation laser, and a custom motion platform designed to hold the porcine larynx model and simulate human respiration.

CALM is based on a spherical orientation mechanism with anti-backlash gears and high-resolution encoders as in Fig. 5.2. It provides high precision (35 μm), a programmable working area (up to 40 $\times$ 40 mm^2 at 400 mm distance), and controllable scanning speeds up to 100 mm/s [2].

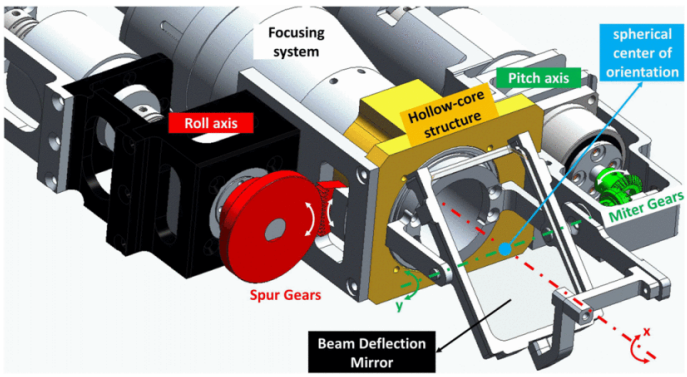

Figure 5.2 CAD Design of CALM [2]

5.3 Methods

Auto-CALM relies on a vision-based controller. Therefore, the first method implemented in the system is responsible for computing the mapping between CALM workspace and the workspace of the microscope camera. The second method implements laser spot tracking, which enables compensating for vision-based controller and laser positioning errors. The subsequent methods implement target tracking and the ablation procedure.

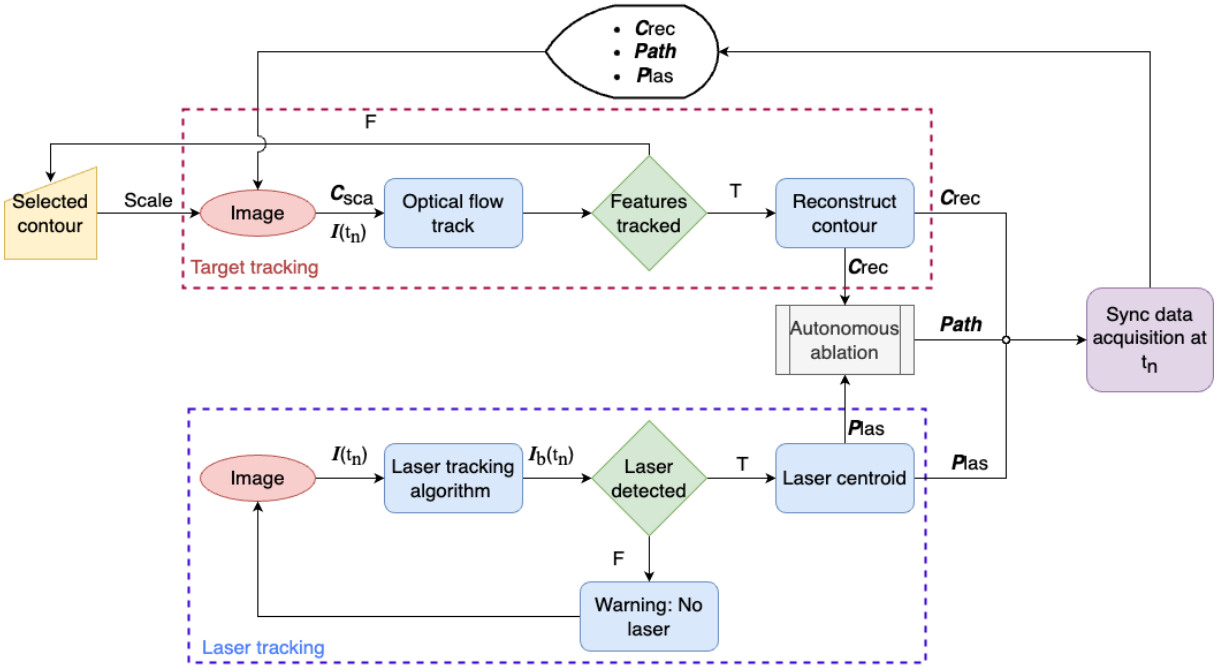

Figure 5.3 Sequence diagram of vision-guided system with three blocks.

The overall framework is shown in Fig. 5.3. The operator begins the procedure by drawing the contour of the desired ablation area on the current video frame using the stylus of the graphics tablet. This contour is then scaled up, generating the tracking contour (i.e., a new contour suitable for tracking). In parallel, the laser spot is tracked on the current video frame, generating a segmentation mask. Then, during operation, the tracking algorithms constantly update the tracking contour position and the laser spot position based on each new video frame. The ablation contour is reconstructed based on the updated tracking contour in parallel. Finally, using this information and the laser spot position, the automatic ablation controller updates the ablation path and the control of CALM.

5.3.1 CALM to RGB Image Mapping

A velocity PID controller controls the CALM laser spot. This is akin to mouse-stylus controllers. Hence, calibration aims to establish a direct mapping that links the laser spot position on the surgical video frame to the positions of CALM's actuators. This enables controlling the aiming of the laser in a closed loop using image coordinates as input. Therefore, this calibration aims to determine the parameters of the affine transformation that maps the RGB image space (i.e., points in the surgical video frame) to CALM's coordinates. Further details regarding calibration are available in this work [50] from our lab.

5.3.2 Laser Spot Tracking

We first implemented laser spot tracking using a classical technique and then developed another implementation using the pre-trained model. Ultimately, we decided to use the pre-trained model approach due to its novelty. Hence, in all the subsequent experiments were conducted using the approach of pre-trained models.

Classical Technique

The laser spot is segmented by thresholding based on the color and area of the laser spot in the hue-saturation-value (HSV) color space. However, this is not sufficient as the laser spot's area changes from frame to frame due to laser diffusion, even if the laser is not moving. To address this, we apply morphological operations [61]. Morphological operations are a set of operations that process images based on shapes. These operations are beneficial for preprocessing tasks such as removing noise, filling holes, or finding

and emphasizing specific shapes within an image. Two basic morphological operations, erosion, and dilation, were used. Furthermore, a Kalman filter [84] is used to get stable laser spot position by optimally fusing model predictions with sensor measurements, addressing inherent uncertainties. Prior predictions are derived from the initial state estimate and simple state transition model. In the simplistic model, we assumed the spot position remains constant from one-time step to the next in the absence of external control inputs. These prior estimates are combined with real-time laser position, factoring in uncertainties quantified by process and measurement covariances. The process covariance represents the uncertainty in the model's predictions, while the measurement covariance represents the uncertainty in the sensor measurements.

Pre-Trained Model

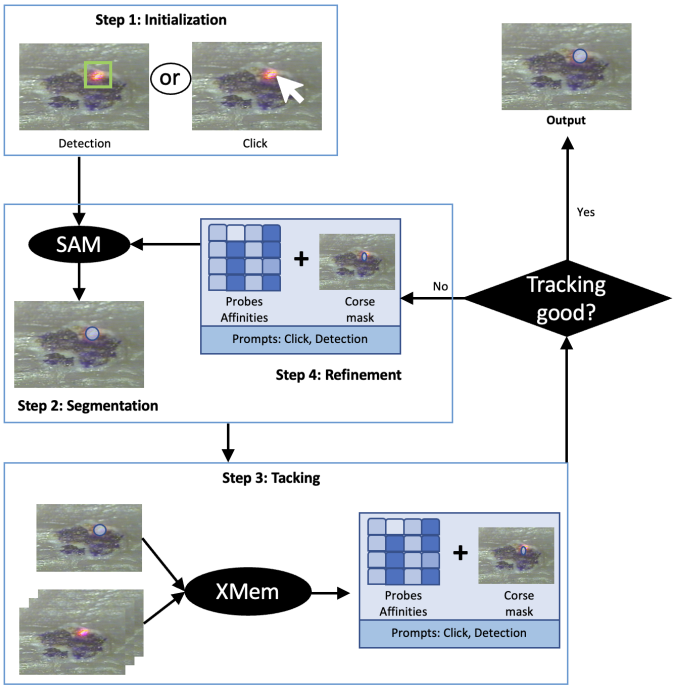

Figure 5.4 Pipeline of laser spot tracking based on SAM and XMem.

The visual laser spot tracking algorithm is a feature-based method. It is designed to be robust to occlusion, background noise, illumination changes, and diffusion at the tissue. Taking inspiration from TAM, a method based on SAM and XMem for tracking the laser spot centroid in the intricate surgical setting is proposed. This method enhanced autonomy during initialization through edge detection, and adjusted probes and affinities

to enhance suitability for intricate surgical procedures. The method's pipeline is shown in the Fig. 5.4 and described below.

Step 1: Initialization - In this step, a mask based on the HSV color space is used to identify potential regions of interest (ROIs) in the initial frame. OpenCV's dilation function is employed to enhance the contour of the laser spot and compensate for any deformation [14]. A subsequent edge detector is used to identify circles within the Regions of Interest (ROIs), taking advantage of the consistent color and brightness exhibited by the laser spot after automatic white balance adjustment within a single frame. To ensure accurate initialization, users have the option to manually select the ROI through interactive clicks, enhancing the reliability of the process.

Step 2: Segmentation with SAM - Exploiting SAM's capabilities to segment ROIs using cues from the automatic detection or the bounding bounding boxes defined by the user, an initial mask of the target object is derived.

Step 3: Tracking with XMem - Leveraging on the mask generated in the previous step, XMem conducts semi-supervised Video Object Tracking (VOS) in the subsequent frames. As XMem performs well in relatively simple scenarios, its predicted masks are predominantly used. However, when mask quality diminishes, XMem's predictions are retained along with intermediate parameters like probes (ROIs) and affinities (features of laser spot), which guide the transition to Step 4.

Step 4: Refinement with SAM - When XMem's predicted mask quality falls short, as mentioned above, SAM is called upon for mask refinement. Concretely, the probes and affinities are adapted as point prompts for SAM, while the predicted mask from Step 2 serves as a mask prompt. Leveraging these prompts, SAM generates a more accurate segmentation mask. This refined mask is also integrated into XMem's temporal correspondence to enhance subsequent object discrimination.

SAM is based on ViT [23] and trained on the large-scale dataset SA-1B [46]. In the case of zero-shot edge detection, the results from BSDS500 [53] show a high recall at 50% precision, albeit with reduced precision. Compared to the ground truth, SAM predicts more edges, including sensible ones that are not annotated in BSDS500. However, given the accurate region of interest (ROI), SAM can accurately segment the contour of the laser spot. To evaluate SAM's performance in our scenario, we calculate the error of the laser spot centroid obtained from SAM in step 2 and compare it to the pre-placed laser position in section III-C. The error is measured to be $29 \pm 6\ \mu m$.

XMem is designed to address the challenges encountered in long-term videos by utilizing a mask description of the target object in the initial frame to track the object

and produce corresponding masks in subsequent frames. The training datasets are YouTubeVOS [87] and DAVIS [65]. The integration of SAM can solve its difficulties of requiring initial accurate mask and recovering from tracking or segmentation failure. To evaluate the predicted mask in step 3, we introduce the average $\mathcal{J}\&\mathcal{F}$ of Jaccard index $\mathcal{J}$ and contour accuracy $\mathcal{F}$ [64]. $\mathcal{J}$ represents the *intersection-over-union* of the estimated segmentation and the groundtruth mask. Given an output segmentation $\mathcal{M}$ and the corresponding ground-truth mask $\mathcal{G}$ it is defined as defined as:

$$\mathcal{J} = |\frac{\mathcal{M} \cap \mathcal{G}}{\mathcal{M} \cup \mathcal{G}}| \tag{5.1}$$

Closed contours $_c(\mathcal{M})$ define the boundary of the mask's spatial area. Consequently, contour-based precision (P_c) and recall (R_c) between the contour points of $_c(\mathcal{M})$ and $_c(\mathcal{G})$ can be calculated using bipartite graph matching. $\mathcal{F}$, denoted as $\mathcal{F} = 2P_cR_c/(P_c + R_c)$, represents a combined metric of precision and recall. Mask quality deterioration in step 3 indicates that $\mathcal{J}\&\mathcal{F}$ is below 85.

5.3.3 Target Tracking

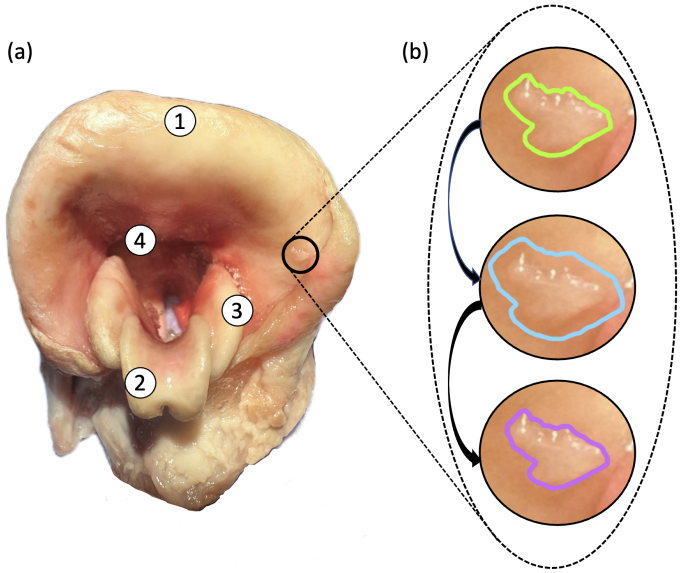

Figure 5.5 (a) Anatomy of the porcine larynx. (b) Scaling strategy: green line represents the drawn target contour, blue line is the scaled contour, and purple line is the reconstructed contour.

Target tracking relies on a scaling strategy integrated with a customized optical flow method. Here, the optical flow tracking incorporates pyramid windowing and census

transformation techniques to effectively address tissue deformations and variations in brightness. Sequence diagram of vision-guided system with three blocks. C_{sca}: Scaled contour; t_n: time sequence; C_{rec}: Reconstructed contour; *Path*: Ablation path generated by algorithm; P_{las}: Laser spot centroid position.

The scaling strategy is employed to prevent tracking failure when the laser spot passes through the original contour and disrupts the feature that is otherwise trackable by optical flow. The scaling strategy involves three steps as shown in Fig. 5.3 and Fig. 5.5 (b):

1. Outlining the lesion contour (Selected contour: green line) with a stylus in the current frame.

2. Scaling the contour to generate the tracking contour (C_{sca}: blue line).

3. Reconstructing the target contour (C_{rec}: purple line) after estimating the tissue motion using the proposed optical flow method.

In the second step, the contour is scaled by 1.3 times the distance between each point of lesion contour and the target centroid to generate the tracking contour, ensuring a minimum distance of 0.26 mm between each point of blue line and its corresponding point on the green line, which is twice the diameter of the CO_2 laser spot.

Consider a point x, y in the image space. Let $\delta x, \delta y$ be the change in the displacement of that point when an object moves. Now, under the assumption that the intensity of a point remains constant and the time and displacement steps are infinitesimally, we arrive at the optical flow constraint equation:

$$I_x u + I_y v + I_t = 0$$

where $u = \frac{\delta x}{\delta t}$ and $v = \frac{\delta y}{\delta t}$, and I_x, I_y, I_t are the partial derivatives of intensity with respect to x, y, and t (time). I_x, I_y, I_t can be easily computed by using the current frame and the previous frame. However, the parallel component of optical flow u, v cannot be computed due to the aperture problem. Hence, it is an ill-posed problem. To obtain a unique solution, we regularize the equation by an additional constraint; the two major methods are (1) Lucas and Kanade and (2) Horn and Schunck. In this work, we adopted Lucas and Kanade's method. This method considers a window around each point within which the optical flow (u, v) is constant. This works well in textured regions. So far, we have considered the object's motion to be small. This might only sometimes be the case.

In order to address this, we employ the pyramidal Luncak Knade algorithm. When an image is downsampled, the coarse object motion is reduced to a few pixels. Hence, we can apply the Lucas Kandae algorithm at a lower resolution. Then, use the optical flow from a lower resolution to wrap the image at a higher resolution. Then, this wrapped image is used to compute the optical flow [73].

The integration of census transformation contributes to mitigate the effects of varying brightness by transforming the pixels surrounding a central pixel into a numerical representation through a kernel size of 3 by 3 [3]. It computes, for every pixel in the image, a binary string (census signature) by comparing its grey value with the grey values in its neighbourhood. This process is based on a transformation that associates a pixel P with a clockwise concatenation of strings representing the set of neighboring pixels whose intensity are different from that of P. Each Census digit is defined as

$$\xi(P,\ P') = \begin{cases} 0 & P - P' > \phi \\ 1 & |P - P'| < \phi \\ 2 & P' - P > \phi \end{cases} \tag{5.2}$$

where ϕ represents the tolerance parameter, defining the threshold for considering pixel intensity differences when comparing neighborhoods. The Census Transform focuses on encoding the relative spatial arrangement of intensities in a local neighborhood, which is a fundamentally different approach aimed at feature description and matching rather than the image filtering of conventional low-pass filtering. Hence, census transformation provides more robust and stable features for optical flow calculation.

5.3.4 Ablation Algorithm

The automatic ablation algorithm (Algorithm 1) controls the state of the CO_2 laser power (ON/OFF) and the velocity vector $v_c = (v_x\ v_y)^T$ of the laser spot. Given the scaled target contour and the tracked visual laser spot, it generates an ablation path plan. This path consist of a series of points P_P^i, where i is from 1 to N at every frame of the RGB image, as shown in Fig. 5.6(a). The laser is on during the entire ablation procedure when moving from one point to the next. The distance between each point is set to be $0.23mm$, which is the diameter of the CO_2 laser spot in our system.

The ablation path is defined to ablate the contour of the target area first to avoid over ablation, as tissue tends to shrink when ablated. Then, the inside area is ablated following a continuous raster scanning path. In each frame, the velocities v_x and v_y are

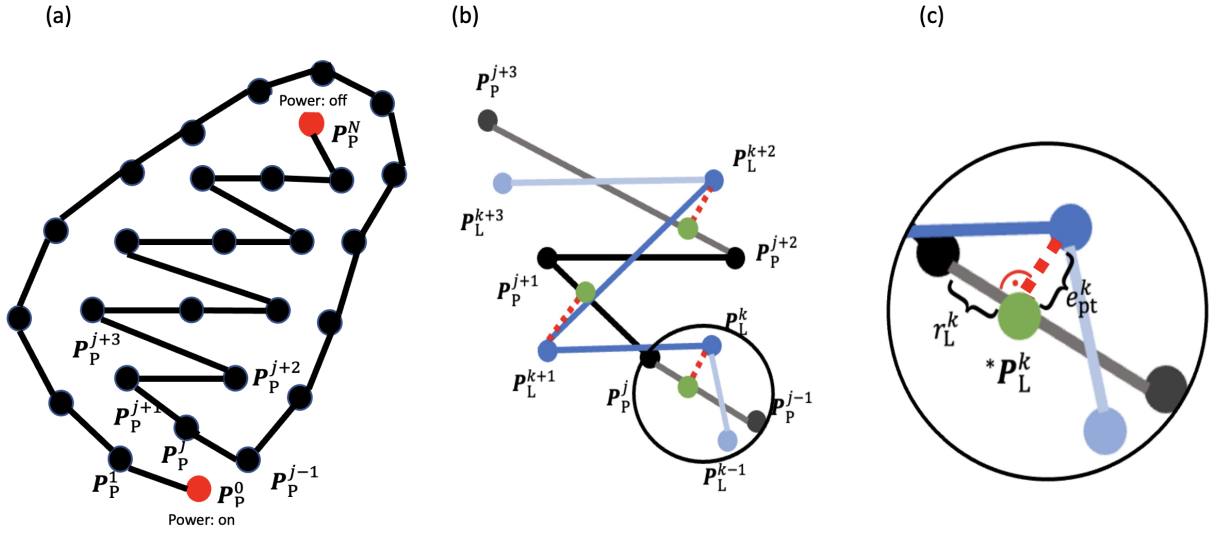

Figure 5.6 (a) Generated path: path inside is a continuous raster scan to ensure complete ablation; the CO_2 laser is activated in the beginning and turned off at the end of scan process. (b)(c) PTE metrics: the determination of the PTE metrics is based on a discrete representation and geometric relationships, as described in [47].

calculated using Eq. 5.3 and Eq. 5.4, respectively, to move the laser spot to the next target point P_{i+1}, where $i = 0, ..., N$ and $N = len(path) - 1$.

$$v_x = sign(\theta_1)|v_y(\Delta x_c/\Delta y_c)| \tag{5.3}$$

$$v_y = sign(\theta_2)|\frac{v(\Delta x_c/\Delta y_c)}{\sqrt{1+(\Delta x_c/\Delta y_c)}}| \tag{5.4}$$

where x_c, y_c are the relative displacement in CLAM workspace required to reached the desired position; v is the speed of CALM, consisting of nominal speed and motion compensation: $v = v_{nominal} + v_{comp}$. In this case, v_{comp} is calculated based on the motion of the target as $v_{comp} = \mathbf{d}/d_t$, where $\mathbf{d}$ represents overall pixel displacement vector as defined in 5.3.3. The velocity compensation compensates for the breathing motion of the target.

Algorithm 1 Ablation algorithm

Input: $\boldsymbol{C}_{rec}, \boldsymbol{P}_{las}$
Output: v_x, v_y, $Power$

$i, Power = 0, Off$
while $i \leq len(\boldsymbol{C}_{rec})$ **do**
  $Path \leftarrow \boldsymbol{C}_{rec}$
  $v_x,\ v_y,\ Dist \leftarrow \boldsymbol{P}_{las}$ **and** $Path[i]$
  if $Dist < threshold$ **then**
    if $i == 0$ **then**
      $Power = On$
    end if
    $i + 1$
  end if
  Publish: $v_x,\ v_y,\ Power$
end while
$v_x, v_y, Power = 0, 0, Off$
Publish: $v_x,\ v_y,\ Power$

where $\boldsymbol{C}_{rec}$ = reconstructed contour; $\boldsymbol{P}_{las}$ = laser spot coordinates; $Path$ = generated ablation path; $Dist$ = distance between laser spot and target point in $Path$; $threshold$ = 0.05 mm (value set to evaluate path following).

5.3.5 Evaluation metrics

The evaluation of the automated ablation system is based on three metrics: Path tracing error (PTE), contour error, and surface error.

Path Tracing Error

The PTE metric represents the deviation between the nominal path and the recorded tracing trajectory as shown in Fig. 5.6 (b) and (c). Both the planned path and the laser spot trajectory are represented in the camera frame (CF). The computation of the orthogonal projection positions ${}^{*}\mathbf{P}_L^k$ of the laser spot $\mathbf{P}_L^k$ onto the segments comprising the planned path is described in the literature [47], where $k = 1, ..., N$. The scaling factor r_L^k is evaluated to determine its validity. If it satisfies the condition $0 \leq r_L^k \leq 1$, the laser spot measurement is projected onto the corresponding path segment and considered for further processing. This results in the computation of the orthogonal path error between

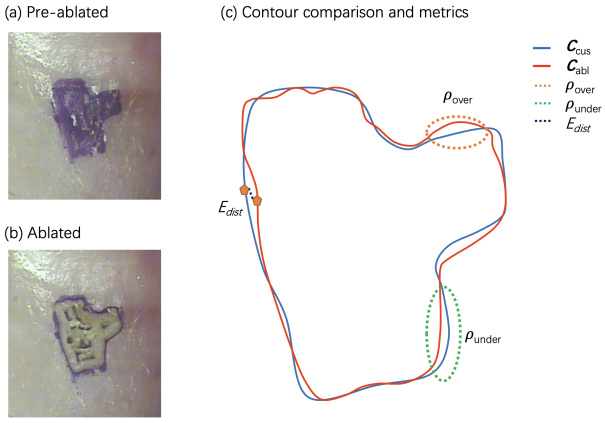

Figure 5.7 An example of contour error measurement.

the k-th spot measurement and its corresponding projection on the planned path as the L^2 norm:

$$e^k_{PT} = ||\mathbf{P}^k_L -^* \mathbf{P}^k_L||_2 \tag{5.5}$$

The RMSE of e^k_{PT} provides the metric PTE as:

$$PTE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(e^k_{PT})^2} \tag{5.6}$$

Contour Error Metrics

For defining the contour error metrics, the distance error E_{dist} illustrated in Fig. 5.7 is used. This error corresponds to the distance between the selected contour $\mathbf{C}_{cus}$ and post-ablation contour $\mathbf{C}_{abl}$ that is measured manually. This is defined as the minimum distance (infimum) from a query point $\mathbf{P}_{cus} \in \mathbf{C}_{cus}$ to $\mathbf{P}_{abl} \in \mathbf{C}_{abl}$:

$$E_{dist} = \text{min\_dist}(C_{cus}, C_{abl}) \tag{5.7}$$

Using E_{dist}, the following two error metrics are defined:

1. Contour Error (CE), defined as the RMSE of E_{dist} over the entire contour.
2. Maximum Absolute Error (MAE), defined as the maximum E_{dist} between $\mathbf{C}_{cus}$ and $\mathbf{C}_{abl}$

Surface Error

The surface error is evaluated using the DICE similarity coefficient, which is a widely used metric to determine the similarity of overlapping regions. Here, DICE is used to compare the ablated area with the target area [91].

$$DICE = \frac{2 \times Area(overlapping)}{Area(Target) + Area(Ablated)} \tag{5.8}$$

The precision and robustness of a laser cutting procedure can also be evaluated by introducing the ratios of overshooting (ρ_{over}) and undershooting (ρ_{under}). These ratios are defined as follows:

$$\rho_{over} = \frac{Area(Overshoot)}{Area(Target)} \tag{5.9}$$

$$\rho_{under} = \frac{Area(Undershoot)}{Area(Target)} \tag{5.10}$$

The denominator for both ρ_{under} and ρ_{over} is the same, allowing for comparison of the undershooting and overshooting effects on system performance as depicted in Fig. 5.7.

In addition, the ablation speed of different experimental conditions was used as a metric for performance comparisons. In the case of Auto-CALM and the automatic system based on TAM, the ablation speed was defined as the average speed between every two points along the planned path. In the case of manual CALM, the ablation speed was computed as the average speed of every movement distance of 0.23 mm. These speed values were then used to calculate the overall ablation speed mean and std for each condition.

5.3.6 Assessment Experiments

The use of a *ex-vivo* porcine larynx specimen in surgical experiments has been validated due to its anatomical, physiological, and functional similarities to the human larynx [79]. Hence, in all experiments *ex-vivo* porcine larynx is used as the specimen. A sinusoidal linear motion with 5 mm displacement and 2 mm/s velocity simulated the physiological motion of the larynx due to respiration [81]. In addition, features on the specimen were marked using a surgical pen, and areas without features were left unmarked.

In all experiments, the perimeter of the target ablation area was drawn using the graphics tablet and stylus. In addition, the CO_2 laser was always operated using the

SmartPulse emission mode (pulsed laser) at a power level of 0.4 W with a frequency of 50 Hz.

Comparative Experiment

In the first experiment, we compared the area ablation performance of Auto-CALM with two other conditions: Manual operation of CALM and autonomous ablation based on TAM replacing target tracking. Four ablation trials were conducted for each method. Breathing motions were simulated, and clear features were present in all the trials, as summarized in Table 5.1. The ablation was performed at the epiglottis of the ex-vivo porcine larynx mode. Here, clear features referred to easily identifiable diseased tissue (i.e., the coloured areas) which are marked by the surgical pen.

Table 5.1 Comparative Experiments Conditions

Experiment	**Location**	**Motion**	**Clear Features**
AUTO-CALM	Ep	Simulated	Present
Manual CALM	Ep	Simulated	Present
TAM	Ep	Simulated	Present

Robustness Experiment

In the second experiment, we examined the robustness of Auto-CALM considering four conditions:

1. No motion and no clear features (NM-NF);
2. Clear features but no motion (NM-F);
3. Motion but no clear features (M-NF);
4. Motion and clear features (M-F).

These conditions are summarized in Table 5.2.

In this experiment, CALM's velocity was initialized at 3 mm/s with compensation and each condition was assessed through ablation tests at four different locations in the larynx as shown in Fig. 5.5:

1. Epiglottis (Ep);

Table 5.2 Robustness Experiments Conditions

Experiment	Location	Motion	Clear Features
NM-NF	Ep PIC PA IFP	None	None
NM-F	Ep PIC PA IFP	None	Present
M-NF	Ep PIC PA IFP	Simulated	None
M-F	Ep PIC PA IFP	Simulated	Present

2. Paired interarytenoid cartilages (PIC);
3. Paired arytenoid (PA);
4. Intralaryngeal fat pad (IFP)

5.4 Results & Discussion

5.4.1 Laser Tracking Result

To evaluate the efficacy of this approach, we performed controlled laser ablations along a predetermined trajectory while documenting the visual laser centroid's path for later comparison with the pre-defined trajectory. In addition, to test the system robustness, we conducted these trials under four distinct conditions: tissue material under low illumination level (T-L), tissue material under high illumination level (T-H), plaster of Paris (PP) under low illumination level (P-L), and PP under high illumination level (P-H). Here, low and high illumination levels correspond to the minimum and maximum settings of the microscope's light source.

Laser tracking was assessed by following a ground truth trajectory across four test scenarios, which included different illumination levels and materials, as shown in Fig. 5.8

and Table 5.3. The difference between the average PTE for the low and high illumination conditions on biological tissue was 12 µm, suggesting that the laser tracking performance is only slightly affected by illumination levels. The maximum PTE for all conditions does not exceed the laser spot radius (114 µm in our system), which indicates that the system maintains errors within a potentially clinically acceptable range. Finally, the spread between the median and maximum values across all conditions is only 57 µm, indicating accurate tracking.

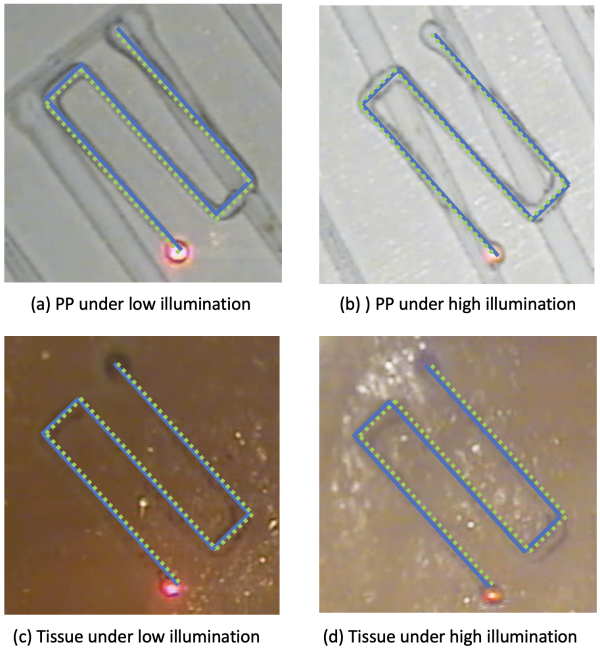

Figure 5.8 Laser tracking results. Green and blue represent path recorded by proposed method and planned path, respectively.

Table 5.3 Path tracing error metrics of laser tracking in μm

Condition	**PTE**	**Mdn**	**Max**
$P-L$	67	55	98
$P-H$	72	66	107
$T-L$	63	69	89
$T-H$	75	81	112

5.4.2 Target Tracking Result

To evaluate tracking accuracy, we recorded the estimated motion and velocity of a target tissue during 8 trials conducted under tissue motion simulation. The direction and speed of motion are controlled by the human respiration platform. In our case, the speed is set to 2 mm/s, simulating the physiological motion of the larynx due to respiration. For each trial, we recorded the moving direction and speed of ten distinct points on the target contour, estimated using the proposed tracking method. Subsequently, we computed two key metrics: Root Mean Squared Error (RMSE) and Maximum Absolute Error (MAE) with respect to the known motions of the target. We then calculated the mean and standard deviation (std) of these metrics.

Table 5.4 shows the RMSE and MAE of the target algorithm for direction and speed estimation during motion, both in feature and no-feature conditions.

Table 5.4 Target tracking assessment

Condition	**Direction (°)**		**Speed (mm/s)**	
	RMSE	**MAE**	**RMSE**	**MAE**
M-NF	1.41 ± 0.0785	1.65 ± 0.2231	0.31 ± 0.0235	0.38 ± 0.08792
M-F	1.16 ± 0.0723	1.29 ± 0.1790	0.28 ± 0.0402	0.34 ± 0.07469

As expected, the tracking error is lower when features are present compared to the no-feature condition. However, the difference in error is relatively small both in direction estimation (RMSE: 0.25°, MAE: 0.36°) and speed estimation (RMSE: 0.03 mm/s, MAE: 0.04 mm/s). This demonstrates the robustness of the target tracking algorithm even in feature-poor conditions.

5.4.3 Comparative Result

Table 5.5 Comparison experiments

Method	**Contour error (mm)**		**Surface error (%)**			**Mean ablation speed** (mm^2/s)
	CE	**MAE**	$\boldsymbol{\rho}_{under}$	$\boldsymbol{\rho}_{over}$	**DICE**	
CALM	1.24 ± 0.0863	1.84 ± 0.0561	4.83 ± 0.383	15.69 ± 2.103	80.02 ± 4.648	0.35 ± 0.1724
TAM	1.71 ± 0.0941	2.14 ± 0.1554	27.83 ± 4.899	0 ± 0	78.77 ± 4.367	1.37 ± 0.5471
Auto-CALM	0.33 ± 0.0354	0.51 ± 0.0972	0.99 ± 0.260	5.49 ± 0.524	96.235 ± 1.648	1.43 ± 0.1623

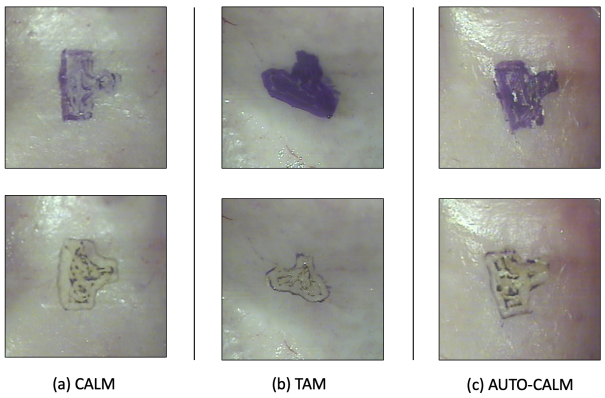

Figure 5.9 Comparative Experiment Conditions: The first row represents the target area, while the second row displays the results after ablation. (a) Ablation using manual control of CALM, (b) automatic ablation using TAM, and (c) automatic ablation using Auto-CALM.

The results presented in Table 5.5 offer a comparative analysis of the efficacy of Auto-CALM with respect to the manually controlled CALM and automatic ablation based on TAM. Sample results are illustrated in Fig. 5.9.

Auto-CALM demonstrates a notable improvement in the contour metric with respect to the other two condition, both in terms of CE and MAE. Specifically, it achieves errors as low as 0.33 mm and 0.51 mm for CE and MAE respectively. These reductions are statistically significant both when compared to CALM condition (CE p-value = 1.34×10^{-4}, MAE p-value = 1.20×10^{-4}) and to the TAM condition (CE p-value = 1.12×10^{-4}, MAE p-value = 1.29×10^{-4}). This substantiates Auto-CALM's enhanced fidelity to the prescribed ablation trajectory.

In terms of surface error, TAM presents an elevated ρ_{under} of 27.83% due to shrinkage of tissue during ablation, whereas CALM is characterized by a large ρ_{over} of 15.69% due to manual errors. In contrast, Auto-CALM registers a modest error in both ρ_{under} of 0.99%, and ρ_{over} of 5.49%. Statistical analyses attest the significant difference between the conditions: When compared to CALM, (ρ_{under} p-value = 1.519×10^{-3}, ρ_{over} p-value = 1.696×10^{-3}); when compared to TAM, (ρ_{under} p-value = 2.304×10^{-3}, ρ_{over} p-value = 3.66×10^{-4}).

Auto-CALM's small errors in both undershooting and overshooting demonstrate the benefit of the proposed control method based on a scaling strategy for target tracking, which handles tissue deformations, and on the implemented laser tracking method, which handles the challenges in tracking the laser spot on real tissue even during ablation.

The congruence between the ablated and target areas, measured by the DICE coefficient, is also maximized by the use of Auto-CALM, reaching 96.22%. This is primary due to the minimized undershooting and overshooting. The statistical significance of this result is confirmed relative to CALM (p-value = 6.873×10^{-3}) and TAM (p-value = 1.4437×10^{-2}).

Observing Fig. 5.9, one can note dark spots on the ablated areas. These spots are, for the most part, small tissue carbonizations resulting from the laser-tissue interactions. Nonetheless, in some spots where the ink had penetrated deeper into the tissue, the ink can indeed still be seen. Further laser passes over the target ablation area should completely remove residues of this “diseased" tissue. This highlights the importance of both knowing how deep we want to ablate the tissue and of devising ways to control the ablation depth. These topics are part of our current research efforts. For what regards the experiments performed herein, the experimental data demonstrated the proposed automatic ablation method could provide a complete coverage of the target ablation area.

Finally, the results show that automatic operations can generate a uniform laser ablation speed that is also overall faster than the mean speed achieved through manual operation. The mean ablation speed under Manual CALM was 0.3±0.1724 mm/s. Both Auto-CALM and TAM exhibited higher mean speeds of 1.43 mm/s and 1.37 mm/s, respectively. In particular, Auto-CALM demonstrated a lower standard deviation, indicating a more uniform motion.

5.4.4 Robustness Result

The results of the robustness assessment experiment are depicted in Fig. 5.10. The data is presented as heatmaps representing the individual mean value of the error metrics for each experimental condition across the four anatomical locations. In addition, the overall error metrics considering all different anatomical locations for each condition are displayed as error bar graphs.

A comparison of planned path and actual path is shown in Fig. 5.11(a), and the corresponding path tracing error (PTE) over length of path is shown in Fig. 5.11(b). The largest mean error was recorded under the M-NF condition at 0.29 mm, closely followed by M-F at 0.28 mm. This stands in contrast with both the NM-NF and NM-F conditions, for which smaller mean PTE errors of 0.20 mm were recorded as expected (due to the absence of tissue motion).

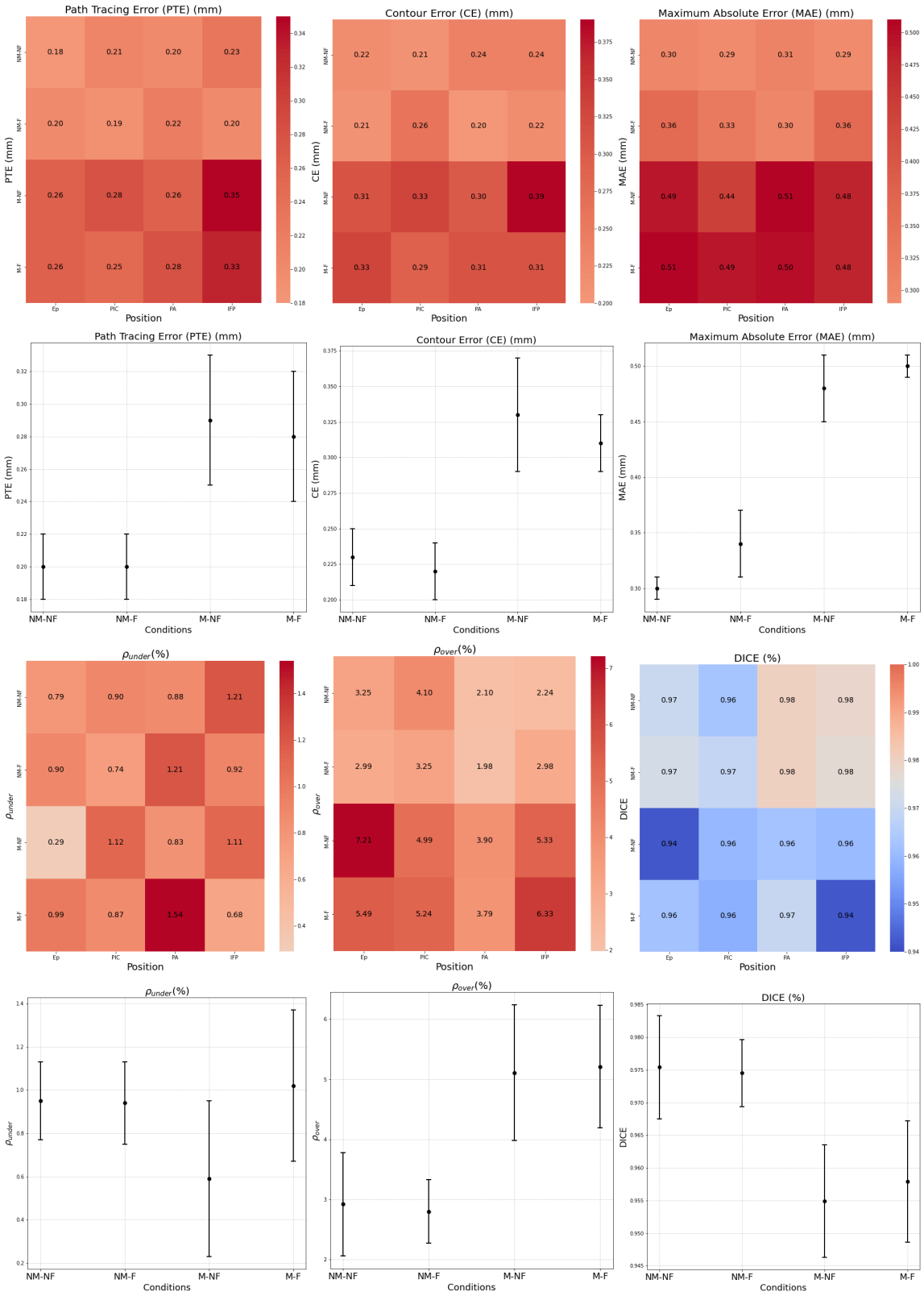

Figure 5.10 Heatmaps representing the individual error metrics for four conditions across four anatomical locations on the specimen, and error bar graphs representing the mean and standard deviation of the errors over all anatomical locations.

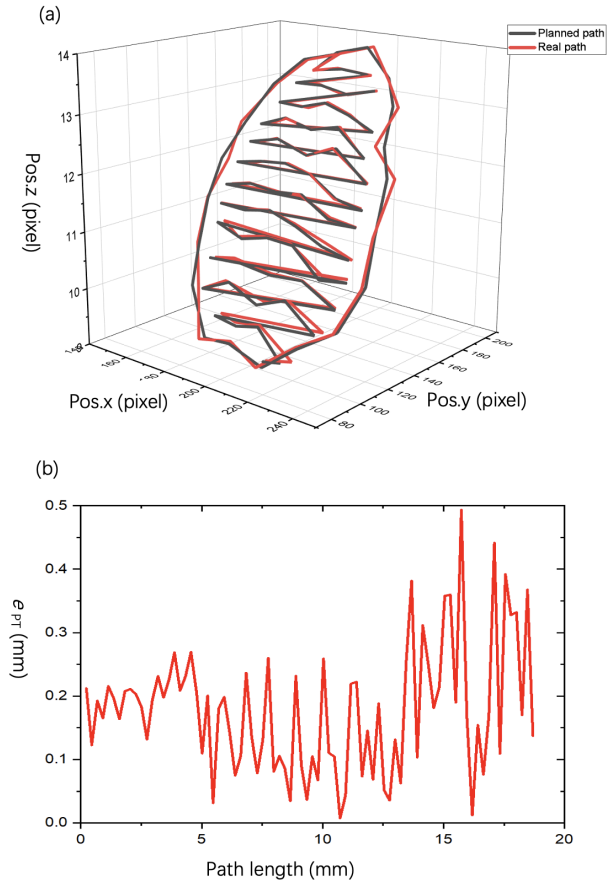

Figure 5.11 (a) Comparison of planned and actual laser path, (b) Path tracing errors over the length of the nominal path.

Overall, the maximum difference between the mean PTE of the difference conditions was only of 0.09 mm, which is a fairly small value considering that the CO_2 laser spot in our system presents a 0.23 mm diameter.

Similar tendencies are also evident in CE and Maximum Absolute Error (MAE) metrics, which are related to contour errors. The M-NF and M-F conditions showed the highest mean CE and MAE errors (respectively 0.33 mm and 0.31 mm for CE; and 0.48 mm and 0.50 mm for MAE). These were contrasted by the lower mean errors in the NM-NF and NM-F conditions: 0.23 mm and 0.22 mm for CE, and 0.30 mm and 0.34 mm for MAE. The maximum differences between the motion and no-motion conditions were only 0.11 mm for CE and 0.20 mm for MAE. Overall, these errors may fall within clinically acceptable ranges given the typical 1–2 mm resection margins [32].

The graphs in Fig. 5.10 also show that the impact of having clear features or no clear features is not strong on the results. This corroborates the good performance of the proposed target tracking algorithm. In addition, a noteworthy observation is the

smaller mean ρ_{under} of 0.88% compared to the mean ρ_{over} at 4.01%. This discrepancy suggests a higher risk of ablating healthy tissue, particularly during target movement. This pattern is more pronounced under the M-NF condition, with the Ep anatomical location displaying the highest ρ_{over} of 7.21%.

The DICE metric, which gauges the matching between the ablated and target areas, was 97.54% for NM-NF, 97.45% for NM-F, 95.49% for M-NF, and 95.79% for M-F. These overall high DICE coefficients demonstrate the ability of Auto-CALM to reliably ablate the desired area under different conditions. Nonetheless, it is clear that the no-motion conditions perform significantly better than motion conditions. This trend is consistent across the different anatomical locations. However, the maximum difference between the mean DICE values for the different conditions is only 2,05%. Therefore, while motion conditions tend to present greater error magnitudes and variability, the differences are relatively minor, attesting the robustness of the system against tissue motions. Also in this case the impact of feature quality on the system performance was very small.

5.5 Conclusion

In Chapter 5 Auto-CALM, an Automated computer-assisted laser microsurgery system specifically tailored for Transoral Laser Microsurgery (TLM), was introduced. The system incorporates real-time visual sensing, providing automated laser ablation capabilities through the integration of novel algorithms for laser tracking, target tracking, and ablation control. A thorough assessment of the system was performed through realistic target ablation experiments conducted on four distinct anatomical locations of *ex-vivo* porcine larynx models. Notably, Auto-CALM demonstrated a DICE coefficient improvement of more than 16% over both manual operation of CALM and automatic ablation based on Track Anything Model (TAM). Furthermore, its performance was assessed using another three key metrics: path tracing error, contour error, and maximum absolute error. The results demonstrated that Auto-CALM performed well across different anatomical locations and under different motion conditions and visual features quality, underscoring its potential to be an effective and reliable system for accurate ablations during transoral laser microsurgery. Future research will include the development of novel techniques for ablation depth control, and the optimization of the proposed algorithms for faster performance. These next developments should provide sufficient data to validate the system's safety and expedite its application in clinical scenarios.

Chapter 6

Conclusion

6.1 Summary

In my PhD thesis, I address the intricate challenges and advancements in the field of laser microsurgery and optical imaging. The first chapter begins by exploring the complexities of laser microsurgery and delves into the use of advanced imaging methods like Optical Coherence Tomography (OCT) and Image Scanning Microscope (ISM). I then focus on the critical factors in designing OCT-guided laser surgery systems, emphasizing the prevention of OCT detector saturation, the need for precise focusing optics, and the alignment of OCT and therapy lasers, with a special emphasis on Separate Optical Paths (SOP) due to their compatibility with existing surgical workflows.

In the third chapter, I present a novel approach in laser scanning microscopy, specifically in ISM, by employing a single-photon avalanche diode (SPAD) array detector for compressive sensing. This approach significantly improves image quality and reduces data acquisition time, addressing major limitations in LSM. The fourth chapter introduces an innovative method for determining ablation depths using 3D OCT in laser microsurgery under SOP. By incorporating compressive sensing into OCT hardware, this method efficiently assesses depth while a feedback controller adjusts the laser exposure time across different tissue types, ensuring accurate ablation.

The final chapter of my thesis introduces Auto-CALM, a new controller for real-time dynamic laser ablation. This system allows for precise ablation, accounting for tissue motions and deformations, and is validated through extensive experiments, demonstrating promising results for clinical application. Overall, my thesis represents a substantial contribution to the field of laser microsurgery, combining sophisticated imaging techniques

with advanced signal processing methodologies like to enhance surgical precision and efficiency.

6.2 Future Directions

There are multiple frontiers in which further research can be pursued to accelerate clinical translation.

Extended focus of depth: Optical imaging devices have a tight depth of focus. An ideal, perfectly flat surface would have no issue. However, biological tissues, especially in in-vivo applications, will not be flat. Plus, there is the involuntary and voluntary motion of muscles has to be handled.

Occlusion Challenges in Laser Surgery: Laser surgery often takes place within confined spaces, where smoke is generated during the procedure. Additionally, the combination of low light conditions and the presence of surgical instruments can lead to occlusion issues that need to be addressed for effective optical imaging.

In my third chapter, I utilized SPAD micro-images to achieve better quality images. However, SPAD is capable of photo-counting. More specifically, SPAD generates a lot of data. Efficiently handling this data should also be considered. Hence, further research must be conducted to explore what more is possible. For example, consider examining the device from UBICEPT, a startup based at MIT and the University of Wisconsin-Madison.

In the field of endoscopic probes, despite several advancements, there still exists a technological gap, particularly in integrating OCT-guided laser surgery within a single probe. This challenge can be approached through various methods: first, by utilizing optical fiber combined with multiplexing techniques; second, by employing mirrors along with multiple light sources; and third, by fusing two separate probes together in a co-aligned fashion. Among these methods, the use of mirrors and the concept of dual probes are especially promising. Mirrors have the advantage of transmitting light across a wide spectrum, effectively accommodating both high and low power wavelengths. This capability is crucial in surgical contexts where varying intensities of light are needed. A notable example of this application is the CO2 laser produced by Deka, which operates invisibly but is co-aligned with a visible aiming laser. Such innovations, leveraging the inherent properties of optical components, offer exciting new avenues for exploration in endoscopic probe design. I strongly suggest users explore Ansys Zemax OpticStudio for new designs and highly recommend the Optical Engineering Course by the University of Colorado Boulder on Cousera.

References

[1] Polarization-sensitive oct. http://obel.ee.uwa.edu.au/research/techniques/polarization-sensitive-oct/.

[2] Alperen Acemoglu, Nikhil Deshpande, Jinoh Lee, Darwin G Caldwell, and Leonardo S Mattos. The calm system: New generation computer-assisted laser microsurgery. In *2019 19th International Conference on Advanced Robotics (ICAR)*, pages 641–646. IEEE, 2019.

[3] G Allevi, L Casacanditella, L Capponi, R Marsili, and G Rossi. Census transform based optical flow for motion detection during different sinusoidal brightness variations. In *Journal of Physics: Conference Series*, volume 1149, page 012032. IOP Publishing, 2018.

[4] Silke Aumann, Sabine Donner, Jörg Fischer, and Frank Müller. *Optical Coherence Tomography (OCT): Principle and Technical Realization*, pages 59–85. Springer International Publishing, Cham, 2019. ISBN 978-3-030-16638-0. doi: 10.1007/978-3-030-16638-0_3. URL https://doi.org/10.1007/978-3-030-16638-0_3.

[5] Author's First name Initials. Author's Last name. Natural image statistics for human and computer vision. CiteSeerX, Year of Publication. Available online at: https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.215.9114.

[6] Beina Azadgoli and Russell Y Baker. Laser applications in surgery. *Annals of Translational Medicine*, 4(23):452, Dec 2016. doi: 10.21037/atm.2016.11.51.

[7] Peter O. Bayguinov, Dennis M. Oakley, Chien-Cheng Shih, Daniel J. Geanon, Matthew S. Joens, and James A. J. Fitzpatrick. Modern laser scanning confocal microscopy. *Current Protocols in Cytometry*, 85(1), June 2018. ISSN 1934-9300. doi: 10.1002/cpcy.39. URL http://dx.doi.org/10.1002/cpcy.39.

[8] Kathy Beaudette, Hyoung Won Baac, Wendy-Julie Madore, Martin Villiger, Nicolas Godbout, Brett E. Bouma, and Caroline Boudoux. Laser tissue coagulation and concurrent optical coherence tomography through a double-clad fiber coupler. *Biomed. Opt. Express*, 6(4):1293–1303, Apr 2015. doi: 10.1364/BOE.6.001293. URL http://www.osapublishing.org/boe/abstract.cfm?URI=boe-6-4-1293.

[9] M Bertero, P Brianzi, and E R Pike. Super-resolution in confocal scanning microscopy. *Inverse Problems*, 3(2):195–212, May 1987. ISSN 1361-6420. doi: 10.1088/0266-5611/3/2/006. URL http://dx.doi.org/10.1088/0266-5611/3/2/006.

[10] M Bertero, P Boccacci, M Defrise, C De Mol, and E R Pike. Super-resolution in confocal scanning microscopy: Ii. the incoherent case. *Inverse Problems*, 5(4): 441–461, August 1989. ISSN 1361-6420. doi: 10.1088/0266-5611/5/4/003. URL http://dx.doi.org/10.1088/0266-5611/5/4/003.

[11] Stephen A. Boppart, Juergen Herrmann, Costas Pitris, Debra L. Stamper, Mark E. Brezinski, and James G. Fujimoto. High-resolution optical coherence tomography-guided laser ablation of surgical tissue. *Journal of Surgical Research*, 82(2):275–284, 1999. ISSN 0022-4804. doi: https://doi.org/10.1006/jsre.1998.5555. URL https://www.sciencedirect.com/science/article/pii/S002248049895555X.

[12] Stephen A. Boppart, Juergen M. Herrmann, Costas Pitris, Debra L. Stamper, Mark E. Brezinski, and James G. Fujimoto. Real-time optical coherence tomography for minimally invasive imaging of prostate ablation. *Computer Aided Surgery*, 6(2): 94–103, January 2001. doi: 10.3109/10929080109145996. URL https://doi.org/10.3109/10929080109145996.

[13] Caroline Boudoux. *Fundamentals of biomedical optics: from light interactions with cells to complex imaging systems*. Pollux, Montréal, 2016. ISBN 978-1-366-45119-4 978-1-366-58826-5 978-1-366-44619-0.

[14] Gary Bradski. The opencv library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 25(11):120–123, 2000.

[15] Mauro Buttafava, Federica Villa, Marco Castello, Giorgio Tortarolo, Enrico Conca, Mirko Sanzaro, Simonluca Piazza, Paolo Bianchini, Alberto Diaspro, Franco Zappa, Giuseppe Vicidomini, and Alberto Tosi. Spad-based asynchronous-readout array detectors for image-scanning microscopy, 2020. URL https://arxiv.org/abs/2002.11443.

[16] Marco Castello, Colin J. R. Sheppard, Alberto Diaspro, and Giuseppe Vicidomini. Image scanning microscopy with a quadrant detector. *Optics Letters*, 40(22):5355, November 2015. doi: 10.1364/ol.40.005355. URL https://doi.org/10.1364/ol.40.005355.

[17] Marco Castello, Giorgio Tortarolo, Mauro Buttafava, Takahiro Deguchi, Federica Villa, Sami Koho, Luca Pesce, Michele Oneto, Simone Pelicci, Luca Lanzanó, Paolo Bianchini, Colin J. R. Sheppard, Alberto Diaspro, Alberto Tosi, and Giuseppe Vicidomini. A robust and versatile platform for image scanning microscopy enabling super-resolution FLIM. *Nature Methods*, 16(2):175–178, January 2019. doi: 10.1038/s41592-018-0291-9. URL https://doi.org/10.1038/s41592-018-0291-9.

[18] W. Chang, Y. Fan, X. Zhang, and H. Liao. An intelligent theranostics method using optical coherence tomography guided automatic laser ablation for neurosurgery. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3224–3227, 2018. doi: 10.1109/EMBC.2018.8513016.

[19] Daniel S Choy. History of lasers in medicine. *Thoracic Cardiovascular Surgery*, 36 (Suppl 2):114–117, Jun 1988. doi: 10.1055/s-2007-1022985.

[20] José-Angel Conchello and Jeff W Lichtman. Optical sectioning microscopy. *Nature Methods*, 2(12):920–931, November 2005. ISSN 1548-7105. doi: 10.1038/nmeth815. URL http://dx.doi.org/10.1038/nmeth815.

[21] I Cox, Colin Sheppard, and T Wilson. Super-resolution by confocal fluorescent microscopy. *Optik - International Journal for Light and Electron Optics*, 60:391–396, 04 1982.

[22] Nikhil Deshpande, Jesus Ortiz, Darwin G. Caldwell, and Leonardo S. Mattos. Enhanced computer-assisted laser microsurgeries with a virtual microscope based surgical system. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2014. doi: 10.1109/icra.2014.6907469. URL http://dx.doi.org/10.1109/ICRA.2014.6907469.

[23] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[24] Wolfgang Drexler, Mengyang Liu, Abhishek Kumar, Tschackad Kamali, Angelika Unterhuber, and Rainer A. Leitgeb. Optical coherence tomography today: speed, contrast, and multimodality. *Journal of Biomedical Optics*, 19(7):1 – 34, 2014. doi: 10.1117/1.JBO.19.7.071412. URL https://doi.org/10.1117/1.JBO.19.7.071412.

[25] Marco F. Duarte, Mark A. Davenport, Dharmpal Takhar, Jason N. Laska, Ting Sun, Kevin F. Kelly, and Richard G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, 2008. doi: 10.1109/MSP.2007.914730.

[26] Hans Edmund Eckel and Walter Franz Thumfart. Laser surgery for the treatment of larynx carcinomas: Indications, techniques, and preliminary results. *Annals of Otology, Rhinology & Laryngology*, 101(2):113–118, February 1992. ISSN 1943-572X. doi: 10.1177/000348949210100202. URL http://dx.doi.org/10.1177/000348949210100202.

[27] Vania V. Estrela, Hermes Aguiar Magalhaes, and Osamu Saotome. Total variation applications in computer vision. 2016. doi: 10.48550/ARXIV.1603.09599. URL https://arxiv.org/abs/1603.09599.

[28] Yingwei Fan, Boyu Zhang, Wei Chang, Xinran Zhang, and Hongen Liao. A novel integration of spectral-domain optical-coherence-tomography and laser-ablation system for precision treatment. *International Journal of Computer Assisted Radiology and Surgery*, 13(3):411–423, September 2017. doi: 10.1007/s11548-017-1664-8. URL https://doi.org/10.1007/s11548-017-1664-8.

[29] Elin Farnell, Henry Kvinge, Julia R. Dupuis, Michael Kirby, Chris Peterson, and Elizabeth C. Schundler. Total variation vs l1 regularization: a comparison of compressive sensing optimization methods for chemical detection. 2019. doi: 10.48550/ARXIV.1906.10603. URL https://arxiv.org/abs/1906.10603.

[30] Loris Fichera, Diego Pardo, Placido Illiano, Jesùs Ortiz, Darwin G. Caldwell, and Leonardo S. Mattos. Online estimation of laser incision depth for transoral microsurgery: approach and preliminary evaluation. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 12(1):53–61, 2016. doi: https://doi.org/10.1002/rcs.1656. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rcs.1656.

[31] Bibin Francis, Manoj Mathew, and Muthuvel Arigovindan. Multiresolution-based weighted regularization for denoised image interpolation from scattered samples with application to confocal microscopy. *J. Opt. Soc. Am. A*, 35(10):1749–1759, Oct 2018. doi: 10.1364/JOSAA.35.001749. URL https://opg.optica.org/josaa/abstract.cfm?URI=josaa-35-10-1749.

[32] Sabrina Garofolo, Cesare Piazza, Francesca Del Bon, Stefano Mangili, Luca Guastini, Francesco Mora, Piero Nicolai, and Giorgio Peretti. Intraoperative narrow band imaging better delineates superficial resection margins during transoral laser microsurgery for early glottic cancer. *Annals of Otology, Rhinology & Laryngology*, 124(4):294–298, October 2014. doi: 10.1177/0003489414556082. URL https://doi.org/10.1177/0003489414556082.

[33] Ajay Gunalan and Leonardo S. Mattos. Towards oct-guided endoscopic laser surgery—a review. *Diagnostics*, 13(4):677, February 2023. ISSN 2075-4418. doi: 10.3390/diagnostics13040677. URL http://dx.doi.org/10.3390/diagnostics13040677.

[34] Ilker Hacihaliloglu, Elvis C.S. Chen, Parvin Mousavi, Purang Abolmaesumi, Emad Boctor, and Cristian A. Linte. *Interventional imaging: Ultrasound*, page 701–720. Elsevier, 2020. ISBN 9780128161760. doi: 10.1016/b978-0-12-816176-0.00033-8. URL http://dx.doi.org/10.1016/B978-0-12-816176-0.00033-8.

[35] Ilker Hacihaliloglu, Elvis C.S. Chen, Parvin Mousavi, Purang Abolmaesumi, Emad Boctor, and Cristian A. Linte. Chapter 28 - interventional imaging: Ultrasound. In S. Kevin Zhou, Daniel Rueckert, and Gabor Fichtinger, editors, *Handbook of Medical Image Computing and Computer Assisted Intervention*, The Elsevier and MICCAI Society Book Series, pages 701–720. Academic Press, 2020. ISBN 978-0-12-816176-0. doi: https://doi.org/10.1016/B978-0-12-816176-0.00033-8. URL https://www.sciencedirect.com/science/article/pii/B9780128161760000338.

[36] Karl-Jürgen Halbhuber and Karsten König. Modern laser scanning microscopy in biology, biotechnology and medicine. *Annals of Anatomy - Anatomischer Anzeiger*, 185(1):1–20, January 2003. ISSN 0940-9602. doi: 10.1016/s0940-9602(03)80002-x. URL http://dx.doi.org/10.1016/S0940-9602(03)80002-X.

[37] Bassel Haydar, Stéphane Chrétien, Adrien Bartoli, and Brahim Tamadazte. Three-dimensional oct compressed sensing using the shearlet transform under continuous trajectories sampling. *Informatics in Medicine Unlocked*, 19:100287, 2020. ISSN 2352-9148. doi: 10.1016/j.imu.2019.100287. URL http://dx.doi.org/10.1016/j.imu.2019.100287.

[38] Chuanzhen Hu, Xianli Wang, Ling Liu, Chuanhai Fu, Kaiqin Chu, and Zachary J. Smith. Fast confocal raman imaging via context-aware compressive sensing. *Analyst*,

146:2348–2357, 2021. doi: 10.1039/D1AN00088H. URL http://dx.doi.org/10.1039/D1AN00088H.

[39] Joseph A. Izatt, Michael A. Choma, and Al-Hafeez Dhalla. *Theory of Optical Coherence Tomography*, pages 65–94. Springer International Publishing, Cham, 2015. ISBN 978-3-319-06419-2. doi: 10.1007/978-3-319-06419-2_3. URL https://doi.org/10.1007/978-3-319-06419-2_3.

[40] Geza J. Jako. Laser surgery of the vocal cordsan experimental study with carbon dioxide lasers on dogs. *The Laryngoscope*, 82(12):2204–2216, December 1972. ISSN 1531-4995. doi: 10.1288/00005537-197212000-00009. URL http://dx.doi.org/10.1288/00005537-197212000-00009.

[41] Jamil Jivraj, Yize Huang, Ronnie Wong, Yi Lu, Barry Vuong, Joel Ramjist, Xijia Gu, and Victor X. D. Yang. Coaxial cavity injected OCT and fiber laser ablation system for real-time monitoring of ablative processes. In *Optical Techniques in Neurosurgery, Neurophotonics, and Optogenetics II*, volume 9305, pages 26 – 29. SPIE, 2015. doi: 10.1117/12.2080606. URL https://doi.org/10.1117/12.2080606.

[42] Jamil Jivraj, Chaoliang Chen, Yize Huang, Joel Ramjist, Yi Lu, Barry Vuong, Xijia Gu, and Victor X. D. Yang. Smart laser osteotomy: integrating a pulsed 1064nm fiber laser into the sample arm of a fiber optic 1310nm oct system for ablation monitoring. *Biomed. Opt. Express*, 9(12):6374–6387, Dec 2018. doi: 10.1364/BOE.9.006374. URL http://www.osapublishing.org/boe/abstract.cfm?URI=boe-9-12-6374.

[43] Nitesh Katta, Austin B. McElroy, Arnold D. Estrada, and Thomas E. Milner. Optical coherence tomography image-guided smart laser knife for surgery. *Lasers in Surgery and Medicine*, 50(3):202–212, 2018. doi: https://doi.org/10.1002/lsm.22705. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/lsm.22705.

[44] Nitesh Katta, Arnold D Estrada, Austin B McElroy, Aleksandra Gruslova, Meagan Oglesby, Andrew G Cabe, Marc D Feldman, RY Declan Fleming, Andrew J Brenner, and Thomas E Milner. Laser brain cancer surgery in a xenograft model guided by optical coherence tomography. *Theranostics*, 9(12):3555–3564, 2019. doi: 10.7150/thno.31811. URL https://doi.org/10.7150/thno.31811.

[45] Ensieh Khalkhal, Majid Rezaei-Tavirani, Mohammad Reza Zali, and Zahra Akbari. The evaluation of laser application in surgery: A review article. *Journal of Lasers in Medical Sciences*, 10(5):S104–S111, December 2019. ISSN 2228-6721. doi: 10.15171/jlms.2019.s18. URL http://dx.doi.org/10.15171/jlms.2019.S18.

[46] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything. *arXiv:2304.02643*, 2023.

[47] Dennis Kundrat, R Graesslin, Andreas Schoob, DT Friedrich, MO Scheithauer, TK Hoffmann, Tobias Ortmaier, Lueder Alexander Kahrs, and PJ Schuler. Preclinical performance evaluation of a robotic endoscope for non-contact laser surgery. *Annals of Biomedical Engineering*, 49:585–600, 2021.

[48] Ben Y.C. Leung, Paul J.L. Webster, James M. Fraser, and Victor X.D. Yang. Real-time guidance of thermal and ultrashort pulsed laser ablation in hard tissue using inline coherent imaging. *Lasers in Surgery and Medicine*, 44(3):249–256, January 2012. doi: 10.1002/lsm.21162. URL https://doi.org/10.1002/lsm.21162.

[49] Chengbo Li, Wotao Yin, Hong Jiang, and Yin Zhang. An efficient augmented lagrangian method with applications to total variation minimization. *Computational Optimization and Applications*, 56(3):507–530, July 2013. doi: 10.1007/s10589-013-9576-1. URL https://doi.org/10.1007/s10589-013-9576-1.

[50] Shunlei Li, Ajay Gunalan, Muhammad Adeel Azam, Veronica Penza, Darwin G. Caldwell, and Leonardo S. Mattos. Auto-calm: Autonomous computer-assisted laser microsurgery. *IEEE Transactions on Medical Robotics and Bionics*.

[51] Zhuoyan Li, Jin H. Shen, John A. Kozub, Ratna Prasad, Pengcheng Lu, and Karen M. Joos. Miniature forward-imaging b-scan optical coherence tomography probe to guide real-time laser ablation. *Lasers in Surgery and Medicine*, 46(3):193–202, March 2014. doi: 10.1002/lsm.22214. URL https://doi.org/10.1002/lsm.22214.

[52] Raphaël Maltais-Tariant, Caroline Boudoux, and Néstor Uribe-Patarroyo. Real-time co-localized oct surveillance of laser therapy using motion corrected speckle decorrelation. *Biomed. Opt. Express*, 11(6):2925–2950, Jun 2020. doi: 10.1364/BOE.385654. URL http://www.osapublishing.org/boe/abstract.cfm?URI=boe-11-6-2925.

[53] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 2, pages 416–423. IEEE, 2001.

[54] Leonardo S. Mattos, Nikhil Deshpande, Giacinto Barresi, Luca Guastini, and Giorgio Peretti. A novel computerized surgeon–machine interface for robot-assisted laser phonomicrosurgery. *The Laryngoscope*, 124(8):1887–1894, 2014. doi: https://doi.org/10.1002/lary.24566. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/lary.24566.

[55] Jessica Mavadia-Shukla, Jiefeng F. Xi, and Xingde D. Li. *Multi-modal Endoscopy: OCT and Fluorescence*, pages 1599–1613. Springer International Publishing, Cham, 2015. ISBN 978-3-319-06419-2. doi: 10.1007/978-3-319-06419-2_54. URL https://doi.org/10.1007/978-3-319-06419-2_54.

[56] James P. McLean and Christine P. Hendon. 3-d compressed sensing optical coherence tomography using predictive coding. *Biomed. Opt. Express*, 12(4):2531–2549, Apr 2021. doi: 10.1364/BOE.421848. URL https://opg.optica.org/boe/abstract.cfm?URI=boe-12-4-2531.

[57] C.A. MILFORD and P.E. O'FLYNN. Management of verrucous carcinoma of the larynx. *Clinical Otolaryngology*, 16(2):160–162, April 1991. ISSN 1749-4486. doi: 10.1111/j.1365-2273.1991.tb01968.x. URL http://dx.doi.org/10.1111/j.1365-2273.1991.tb01968.x.

[58] Claus B. Müller and Jörg Enderlein. Image scanning microscopy. *Physical Review Letters*, 104(19), May 2010. ISSN 1079-7114. doi: 10.1103/physrevlett.104.198101. URL http://dx.doi.org/10.1103/PhysRevLett.104.198101.

[59] F.W. Neukam and F. Stelzle. Laser tumor treatment in oral and maxillofacial surgery. *Physics Procedia*, 5:91–100, 2010. ISSN 1875-3892. doi: https://doi.org/10.1016/j.phpro.2010.08.125. URL https://www.sciencedirect.com/science/article/pii/S1875389210005511. Laser Assisted Net Shape Engineering 6, Proceedings of the LANE 2010, Part 1.

[60] Masato Ohmi, Manabu Tanizawa, Atsushi Fukunaga, and Masamitsu Haruna. In-situ observation of tissue laser ablation using optical coherence tomography. *Optical and Quantum Electronics*, 37(13-15):1175–1183, December 2005. doi: 10.1007/s11082-005-4189-2. URL https://doi.org/10.1007/s11082-005-4189-2.

[61] OpenCV. Morphological operations in opencv. https://docs.opencv.org/4.x/d9/d61/tutorial_py_morphological_ops.html, 2024. Accessed: 2024-02-19.

[62] C. K. N. Patel. Continuous-wave laser action on vibrational-rotational transitions of co_2. *Phys. Rev.*, 136:A1187–A1193, Nov 1964. doi: 10.1103/PhysRev.136.A1187. URL https://link.aps.org/doi/10.1103/PhysRev.136.A1187.

[63] N. Pavillon and N. I. Smith. Compressed sensing laser scanning microscopy. *Opt. Express*, 24(26):30038–30052, Dec 2016. doi: 10.1364/OE.24.030038. URL https://opg.optica.org/oe/abstract.cfm?URI=oe-24-26-30038.

[64] Federico Perazzi, Jordi Pont-Tuset, Brian McWilliams, Luc Van Gool, Markus Gross, and Alexander Sorkine-Hornung. A benchmark dataset and evaluation methodology for video object segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 724–732, 2016.

[65] Jordi Pont-Tuset, Federico Perazzi, Sergi Caelles, Pablo Arbeláez, Alex Sorkine-Hornung, and Luc Van Gool. The 2017 davis challenge on video object segmentation. *arXiv preprint arXiv:1704.00675*, 2017.

[66] Dan Z. Reinstein, Timothy J. Archer, and Marine Gobbe. The history of lasik. *Journal of Refractive Surgery*, 28(4):291–298, April 2012. ISSN 1081-597X. doi: 10.3928/1081597x-20120229-01. URL http://dx.doi.org/10.3928/1081597x-20120229-01.

[67] Andrew M. Rollins and Joseph A. Izatt. Optimal interferometer designs for optical coherence tomography. *Opt. Lett.*, 24(21):1484–1486, Nov 1999. doi: 10.1364/OL.24.001484. URL http://www.osapublishing.org/ol/abstract.cfm?URI=ol-24-21-1484.

[68] Marc Rubinstein and William B. Armstrong. Transoral laser microsurgery for laryngeal cancer: A primer and review of laser dosimetry. *Lasers in Medical Science*, 26(1):113–124, September 2010. ISSN 1435-604X. doi: 10.1007/s10103-010-0834-5. URL http://dx.doi.org/10.1007/s10103-010-0834-5.

[69] H Rudert. Laser-chirurgie in der hno-heilkunde [laser surgery in ent surgery]. *Laryngol Rhinol Otol (Stuttg)*, 67(6):261–8, 1988.

[70] C.J.R. Sheppard and A. Choudhury. Image formation in the scanning microscope. *Optica Acta: International Journal of Optics*, 24(10):1051–1073, October 1977. ISSN 0030-3909. doi: 10.1080/713819421. URL http://dx.doi.org/10.1080/713819421.

[71] Colin Sheppard. Super-resolution in confocal imaging. *Optik - International Journal for Light and Electron Optics*, 80:53, 02 1988.

[72] Yoni Sher. Review of algorithms for compressive sensing of images. 2019. doi: 10.48550/ARXIV.1908.01642. URL https://arxiv.org/abs/1908.01642.

[73] Shree K. Nayar. First Principles of Computer Vision. https://www.youtube.com/playlist?list=PL2zRqk16wsdoYzrWStffqBAoUY8XdvatV, 2024. [Online; accessed 24-March-2024].

[74] Emanuel M. Skolnik, Lawrence Martin, King F. Yee, and Michael A. Wheatley. Radiation failures in cancer of the larynx. *Annals of Otology, Rhinology & Laryngology*, 84(6):804–811, November 1975. ISSN 1943-572X. doi: 10.1177/000348947508400612. URL http://dx.doi.org/10.1177/000348947508400612.

[75] Ki Uk Song. Footprints in laser medicine and surgery: Beginnings, present, and future. *Medical Lasers*, 6(1):1–4, June 2017. ISSN 2288-0224. doi: 10.25289/ml.2017.6.1.1. URL http://dx.doi.org/10.25289/ML.2017.6.1.1.

[76] . Steiner. *Endoscopic Laser Surgery of the Upper Aerodigestive Tract. With Special Emphasis on Cancer Surgery*. Georg Thieme Verlag Stuttgart, New York, 2000. ISBN 9783131627018. doi: 10.1055/b-006-161074. URL http://dx.doi.org/10.1055/b-006-161074.

[77] M. Stuart Strong. Laser excision of carcinoma of the larynx. *The Laryngoscope*, 85(8):1286–1289, August 1975. ISSN 1531-4995. doi: 10.1288/00005537-197508000-00003. URL http://dx.doi.org/10.1288/00005537-197508000-00003.

[78] Jingjing Sun and Huikai Xie. Mems-based endoscopic optical coherence tomography. *International Journal of Optics*, 2011:825629, 2011. ISSN 1687-9384. doi: 10.1155/2011/825629. URL https://doi.org/10.1155/2011/825629.

[79] M Makin Swindle, Andrew Makin, Alan J Herron, Fred J Clubb Jr, and Kendall S Frazier. Swine as models in biomedical research and toxicology testing. *Veterinary pathology*, 49(2):344–356, 2012.

[80] Xu Teng, Feng Li, and Chao Lu. Visualization of materials using the confocal laser scanning microscopy technique. *Chemical Society Reviews*, 49(8):2408–2425, 2020. ISSN 1460-4744. doi: 10.1039/c8cs00061a. URL http://dx.doi.org/10.1039/C8CS00061A.

[81] Aggeliki Tsoli, Naureen Mahmood, and Michael J Black. Breathing life into shape: Capturing, modeling and animating 3d human breathing. *ACM Transactions on graphics (TOG)*, 33(4):1–11, 2014.

[82] Charles W. Vaughan, M.Stuart Strong, and Geza J. Jako. Laryngeal carcinoma: Transoral treatment utilizing the co2 laser. *The American Journal of Surgery*, 136 (4):490–493, October 1978. ISSN 0002-9610. doi: 10.1016/0002-9610(78)90267-2. URL http://dx.doi.org/10.1016/0002-9610(78)90267-2.

[83] Paul J. L. Webster, Joe X. Z. Yu, Ben Y. C. Leung, Mitchell D. Anderson, Victor X. D. Yang, and James M. Fraser. In situ 24 khz coherent imaging of morphology change in laser percussion drilling. *Opt. Lett.*, 35(5):646–648, Mar 2010. doi: 10.1364/OL.35.000646. URL http://ol.osa.org/abstract.cfm?URI=ol-35-5-646.

[84] Greg Welch, Gary Bishop, et al. An introduction to the kalman filter. 1995.

[85] Joel M. White, Harold E. Goodis, and Charles L. Rose. Use of the pulsed nd:yag laser for intraoral soft tissue surgery. *Lasers in Surgery and Medicine*, 11(5): 455–461, January 1991. ISSN 1096-9101. doi: 10.1002/lsm.1900110511. URL http://dx.doi.org/10.1002/lsm.1900110511.

[86] Rebecca M. Willett, Roummel F. Marcia, and Jonathan M. Nichols. Compressed sensing for practical optical imaging systems: a tutorial. *Optical Engineering*, 50(7): 072601, 2011. doi: 10.1117/1.3596602. URL https://doi.org/10.1117/1.3596602.

[87] Ning Xu, Linjie Yang, Yuchen Fan, Dingcheng Yue, Yuchen Liang, Jianchao Yang, and Thomas Huang. Youtube-vos: A large-scale video object segmentation benchmark. *arXiv preprint arXiv:1809.03327*, 2018.

[88] Joseph Zammit and Ian J. Wassell. Adaptive block compressive sensing: Toward a real-time and low-complexity implementation. *IEEE Access*, 8:120999–121013, 2020. doi: 10.1109/access.2020.3006861. URL https://doi.org/10.1109/access.2020.3006861.

[89] Yaokun Zhang, Tom Pfeiffer, Marcel Weller, Wolfgang Wieser, Robert Huber, Jörg Raczkowsky, Jörg Schipper, Heinz Wörn, and Thomas Klenzner. Optical coherence tomography guided laser cochleostomy: Towards the accuracy on tens of micrometer scale. *B ioMed Research International*, 2014:1–10, 2014. doi: 10.1155/2014/251814. URL https://doi.org/10.1155/2014/251814.

[90] Yin Zhang. *An Efficient Algorithm For Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing*. PhD thesis.

[91] Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare MC Tempany, Michael R Kaus, Steven J Haker, William M Wells III, Ferenc A Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189, 2004.

[92] A. Zunino, E. Slenders, F. Fersini, et al. Open-source tools enable accessible and advanced image scanning microscopy data analysis. *Nature Photonics*, 2023. doi: 10.1038/s41566-023-01216-x. URL https://doi.org/10.1038/s41566-023-01216-x.

[93] Alessandro Zunino, Marco Castello, and Giuseppe Vicidomini. Reconstructing the image scanning microscopy dataset: an inverse problem, 2022. URL https://arxiv.org/abs/2211.12510.

[94] Alessandro Zunino, Marco Castello, and Giuseppe Vicidomini. Reconstructing the image scanning microscopy dataset: an inverse problem. *Inverse Problems*, 39 (6):064004, April 2023. ISSN 1361-6420. doi: 10.1088/1361-6420/accdc5. URL http://dx.doi.org/10.1088/1361-6420/accdc5.

Author Bio

Ajay Gunalan, born on December 20, 1995, in Chennai, India, is a Ph.D. student at the Italian Institute of Technology, where he is pioneering the development of novel laser technologies for micro-surgical robots. With a Bachelor's degree in Mechanical Engineering from B.S.A. Crescent Institute of Science and Technology in Chennai, India, he has a solid foundation in engineering principles. Ajay's research is focused on enhancing healthcare affordability by integrating medical and engineering disciplines. His expertise spans computational imaging, image processing, machine learning, and robotic software development. Ajay is deeply committed to working with diverse teams to solve new challenges each day, aiming to transform healthcare through innovative technology. His previous experiences include developing Medical Device, Humanoid, Legged Robot, and Tele-Robotics.

Contact: ajay.gunalan@pm.me
ajaygunalan.com