The University of Hong Kong
Faculty of Engineering
Department of Mechanical Engineering

Soft Robotic Manipulation for Intraoperative MRI-guided Non-contact Laser Surgery

Fang Ge

Sep 2021

A thesis submitted in partial fulfilment of the requirements of
the Degree of Doctor of Philosophy
at The University of Hong Kong

Abstract of thesis entitled

"Soft Robotic Manipulation for Intraoperative MRI-guided Non-contact Laser Surgery"

Submitted by

Fang Ge

for the degree of Doctor of Philosophy
at The University of Hong Kong
in September 2020

Magnetic resonance imaging (MRI) offers high-contrast soft tissue imaging and the unique capability of temperature sensing inside the tissue, making it superior guidance in image-guided surgery. Laser-based tumor ablation is one of the treatments that has significantly benefited from MRI guidance, with which 3-dimensional (3D) ablation margins alongside thermal distributions can be evaluated in real-time to protect surrounding critical structures while ensuring adequate ablation margins. However, the confined bore of MRI and its high magnetic field significantly limit surgeons' access to patients, which facilitates the development of MRI-guided robotic systems to allow remote control. Although many systems were proposed to enable intra-operative MRI-guided robotic laser ablation, they generally only allow the targeting of rigid and contact-based laser probes inserted to the surgical site along a straight pathway. There is still a gap for developing an MRI-guided robotic system capable of flexible navigation to targeted lesions and performing delicate non-contact laser beam manipulation.

The main focus of this thesis is to develop a robotic system that integrates a soft continuum manipulator and its high-precision control to enable flexible and precise laser beam manipulation in MRI. Soft robots, benefiting from their elastic body, ensure safe interaction with their surroundings, thus allowing noninvasive and flexible access to the deep surgical sites through confined natural orifices, e.g., transoral approach. Soft robots can be fabricated using magnetic resonance (MR) safe materials and driven by pressurized fluid flow, which brings new opportunities to the development of MRI-guided robotics. The

challenge of precise motion control for soft robots is addressed by employing closed-loop control with learning-based modeling and vision feedback. An online learning visual servo control framework is proposed to approximate the nonlinear robot kinematics without prior knowledge of the robot and camera parameters, and enable precise robot navigation even under unknown external disturbances. To achieve delicate laser beam steering using soft robotic manipulators, eye-to-hand visual servo controllers are investigated based on epipolar geometry modeling and machine-learning modeling. A novel miniature soft laser manipulator that ensures MR safety but maintains dexterous manipulation is designed, with its parameters optimized using finite element analysis (FEA). All in all, an integrated soft robotic system is proposed to achieve MRI-guided transoral laser surgery. The robot enables endoscopic laser delivery and operating in the oral and pharyngeal region with sub-millimeter accuracy ($<$0.2 mm). A patient-specific dental guard is designed to create an open-jaw cavity for robot anchorage and access of auxiliary instruments such as a fiberscope. Novel wireless MR markers are incorporated to enable positional tracking in the MRI coordinate. Furthermore, preclinical trials were conducted to evaluate the robot performance both with *ex-vivo* swine tissue and a cadaver model.

(Word count: 433 words)

Declaration

I declare that this thesis represents my own work, except where due acknowledgment is made, and that it has not been previously included in a thesis, dissertation or report submitted to this University or to any other institution for a degree, diploma or other qualifications.

Signed _____

Fang Ge

Dedication

To my parents
who always trust and support me
no matter what I decide to do

To my friends
who have inspired and encouraged me
to keep going and never give up

To my teammates
who have accompanied and helped me
for my research and life

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Ka-Wai Kwok, for his advice, guidance, and encouragement throughout my PhD study. I am deeply motivated by his passion and rigorous attitude towards research. He gave me a comprehensive training, ranging from technical skills, presentation, writing, to communication, which will benefit my whole life.

I am grateful to Dr. Jason Ying-Kuen and Dr. Danny Tat Ming Chan, who has provided great support to my research in the clinical aspect. I am also thankful to Dr. Hing-Chiu Chang, for his professional advice and patient guidance in my MRI experiments.

I am also grateful to all my excellent fellows in the IRIS lab. It is an unforgettable experience with you to discuss questions and conduct experiments. I would like to thank Ziyang Dong, Xiaomei Wang, Kui Wang, Justin Di-Lang Ho, Zhuoliang He, Marco Chow, Wai Lun Tang, Chimlee Cheung, Xiaochen Xie, Becky Thi-Sze Chung. Without your kind assistance, I could not make such achievements.

Thanks for the great support of Hong Kong PhD fellowship, which lightens my financial burden and offers opportunities for me to participate in numerous international academic events.

Lastly, but most importantly, I wish to thank my parents for their unlimited support for my PhD study.

Contents

List of Figures

List of Tables

Nomenclature

Number Sets

$\mathbb{R}$	Real Numbers
$\mathbb{U}$	Actuation Space

Acronyms

1D	1-dimensional
2D	2-dimensional
3D	3-dimensional
6D	6-dimensional
AR	Augmented reality
ASTM	American society for testing and materials
CC	Constant curvature
CO_2	Carbon dioxide
CT	Computed tomography
DICOM	Digital Imaging and Communications in Medicine
DLL	Dynamic link library
DoF	Degree of freedom
EM	Electromagnetic
ENT	Ear, neck and throat
FBG	Fiber Bragg grating
FDA	Food and drug administration
FEM	Finite element modeling
FLASH	Fast low-angle shot

FOV	Field of view
FPC	Flexible printed circuit
FSA	Frozen section analysis
FSE	Fast spin echo
GPR	Gaussian process regression
GRE	Gradient echo
HNC	Head and neck cancer
ICRF	Inductively coupled radio frequency
ID	Inner diameter
ILI	Intensity linear interpolation
intra-op	intra-operative
LGPR	Localized Gaussian process regression
LITT	Laser interstitial thermal therapy
MR	Magnetic resonance
MRI	Magnetic resonance imaging
NEMA	National Electrical Manufacturers Association
NN	Neural network
OCT	Optical coherence tomography
OD	Outer diameter
ONP	Oral, nasal, and pharyngeal
PA	Polycaprolactam
PCC	Piecewise constant curvature
PET	Positron emission tomography
PRE	Proton resonance shift
pre-op	pre-operative
PTFE	Polytetrafluoroethylene
PU	Polyurethane
RCM	Remote-of-center motion

RF	Radiofrequency
RGB	Red-green-blue
RMS	Root-mean-square
RMSE	Root-mean-square error
ROI	Region of interest
SAR	Specific absorption rate
SRC	Spring reinforcement constraint
TLM	Transoral laser microsurgery
TORS	Transoral robotic surgery
UDP	User Datagram Protocol

Chapter 1

Introduction

1.1 Motivation and Objectives

Magnetic resonance imaging (MRI) provides a compelling set of features for the guidance of interventional procedures, including high contrast imaging of soft tissue, detailed visualization of physiological changes, and thermometry. Laser-based tumor ablation stands to benefit significantly from MRI guidance because 3-dimensional (3D) ablation margins alongside thermal distributions can be evaluated in real-time to protect surrounding critical structures while ensuring adequate ablation margins. As a result, an increasing number of studies have introduced intra-operative (intra-op) MRI to guide and evaluate the laser ablation process *in-situ* with rigid and contact-based laser probes, e.g., for percutaneous liver ablation. Comparatively, few researchers investigated MRI usage with projection-based (non-contact) lasers such as those used for transoral laser microsurgery (TLM). This non-contact laser setting requires dexterous laser beam steering at the ablation site, which raises substantial challenges due to the confined bore of MRI and its high magnetic field (i.e., 1.5/3T).

TLM plays a vital role in the treatment of head-and-neck cancers. During current TLM, surgeons must use their empirical technique to estimate the laser ablation depth beyond the crucial artery/nerve while dissecting these tumors. In typical surgical settings, intra-op evaluation can only be performed through time-consuming and costly pathology laboratory procedures such as frozen section analysis (FSA). However, FSA can only assess the gross tumor volumes that are resected, rather than the unresected region left in the oral, nasal, and pharyngeal (ONP) cavity. By introducing MRI guidance, 3D resection margins can

be monitored and evaluated intra-operatively and in real-time to protect critical structures while resecting adequate margins. This may eliminate the need for FSA, smoothening the surgical workflow, shortening operative time, and improving patient outcomes. However, this drives the demand for a robotic system that adapts to the confined ONP anatomy and MRI environment while providing high-precision manipulation.

Soft manipulators driven by pressurized fluid flow can be used to address this challenge through their high conformability, dexterity, and ability to be magnetic resonance (MR) safe. Bio-inspired continuum robots made of elastomer materials are compliant and flexible. They are specifically suitable for tasks involving interactions with surroundings, in which the robots can passively vary their morphology to adapt unstructured environment. However, their flexibility and nonlinear actuated deformation usually hinder their uses in precise manipulation compared to their rigid counterparts.

To address the unmet technical challenges towards MRI-guided transoral laser surgery, this research focuses on developing adaptive controllers for soft robotic manipulation, high-precision soft laser manipulator, and integrated transoral robotic system. The main objectives of this thesis include:

1. To develop an adaptive eye-in-hand visual servoing controller that avoids complicated analytical modeling process of soft continuum robots, while enabling its precise navigation control in constrained environments.

2. To design a miniaturized MR-safe laser manipulator capable of delicate laser beam steering in confined cavities.

3. To propose real-time sensing and control methods that enable precise laser spot targeting using soft manipulators even in the harsh MRI environment.

4. To validate real-time MR-based positional tracking using wireless MR markers and 1-dimensional (1D) projection MRI pulse sequence.

5. To develop a soft robotic system for MRI-guided transoral laser surgery, which can provide flexible and precise delivery of laser energy to lesions even in deep regions.

6. To validate the robot performance through both *ex-vivo* and cadaveric laser ablation trials. Real-time MR thermometry will be investigated to monitor the laser ablation process.

1.2 Structure and Contributions of Thesis

Chapter 2 presents a review of the recent advances in intra-op MRI-guided intervention, in particular for laser surgery. The classification of MR safety is listed, which defines the standard for the development of MRI-guided robotic systems. The current status of non-contact laser surgery in head and neck cancer treatment is reviewed, including transoral laser microsurgery, transoral surgical robotic systems, and endoscopic laser manipulators. In the review of these systems, the clinical demand for efficient intra-op tools, such as intra-op MRI, to assess the laser ablation completeness is highlighted. The existing robotic systems for transoral laser surgery are not MRI compatible. In contrast, the systems with MRI guidance currently are not equipped for controlling non-contact, projection-based lasers, particularly within constrained anatomy such as the oral and pharyngeal cavities. The perspectives regarding the MR-compatible actuation, real-time sensing in MRI, continuum robots for MR-safe/conditional manipulation, and precise motion control for soft continuum robots are discussed.

Chapter 3 proposes an adaptive visual servo controller using nonparametric online learning, enabling precise motion control for a hydraulic-driven soft robot. The controller is constructed by learning the inverse mapping solely from collected camera images, without any prior knowledge of the robot and camera parameters. Excellent accuracy in the learning inverse mapping is assured without having to tune the hyperparameters in the learning approach. Localized Gaussian process regression (LGPR) models enable fast online updates to accommodate new input data that reflect the latest robot status. Sensing data provided by fiber Bragg gratings (FBGs) are combined to enhance motion estimation accuracy under feature-deficient conditions. Augmented reality (AR) is integrated to virtually extending the camera view beyond the normal field-of-view (FOV), thus offering enhanced guidance for operators.

Chapter 4 presents visual servo controllers for automatic laser spot steering using soft manipulators. An epipolar-geometry model is established to acquire the inverse transition mapping from image to actuation space. With this inverse mapping, a feedback control law is derived without dependence on prior information of tissue geometry. A state observer is also designed to provide a dynamic estimation of robot configurations, enabling accurate computation of the inverse transition mapping. A learning-based visual servoing controller is also implemented to achieve automatic targeting of the laser spot that could be projected along any trajectory prescribed *in-situ* on the endoscopic view. The kinematics of soft robot

is approximated using machine learning techniques, which can accommodate modelling uncertainties from the actual operation data and eliminate the need for hand-eye calibration.

Chapter 5 introduces the design of a soft robotic system for MRI-guided transoral laser surgery. The clinical requirements are investigated. A reinforced miniature soft laser manipulator is proposed, with finite element analysis (FEA) conducted to optimize its spring reinforcements. A hydraulic motor is integrated to achieve insertion/retraction of the soft robotic endoscope in the oral and pharyngeal anatomy. A dynamic model is established for the hydraulic transmission system to guide its parametric design. A patient-specific dental guard is proposed to create an open-jaw cavity for robot anchorage and access to auxiliary instruments such as a fiberscope. Wireless MR markers are fabricated and validated to enable real-time positional tracking in MRI.

Chapter 6 presents the preclinical trials of the transoral robotic system with *ex-vivo* swine tissue and a cadaver model. Signal-to-noise tests are firstly conducted to verify the MRI compatibility of the proposed robotic system. The use of anatomical T2 imaging and MR thermometry for monitoring laser ablation is investigated. A cadaver trial is conducted to validate the robot setup/anchorage, intra-op MRI guidance, as well as visual servo-based laser beam manipulation in actual human anatomy. During laser ablation, MR thermometry is implemented to measure the thermal diffusion process in the ablated lesion.

The main contributions of the thesis are summarized as:

1. *First* attempt to address a learning-based visual servo control for a fluid-driven soft robot such that the inverse kinematics can be directly approximated by LGPR. Online updating of LGPR models to compensate *dynamic* disturbances. Point tracking and path following experiments were conducted to demonstrate the precise visual servo of a hyper-elastic soft robot under varying payload.

2. Development of enhancing methods for eye-in-hand soft robotic visual servoing. A hybrid eye-in-hand tracking algorithm is proposed, by combining image-based 2-dimensional (2D) motion estimation with FBGs-based strain sensing to enhance the tracking accuracy in dim/feature-deficient scenarios. An extended view is integrated to allow visualization of virtual features outside the actual camera view.

3. Development of an eye-to-hand vision-based feedback controller for soft robotic laser manipulation based on epipolar geometry modelling. The controller enables precise

laser spot control without prior knowledge of the tissue surface geometry. Dynamic estimation of soft robot configurations is achieved by fusing the model prediction and camera feedback, thus allowing accurate calculation of robot kinematics.

4. Development of a learning-based eye-to-hand visual servoing controller that enables high-precision laser beam steering using a soft laser manipulator. The controller was tested via the path following experiments with various patterns, e.g., batman and zig-zag filling, which demonstrated sub-millimeter accuracy in both boundary tracking and region filling.

5. Design of a reinforced miniature fluid-driven soft manipulator capable of delicate, precise, and repeatable laser spot steering on the mucosa. The MR-safe hydraulic actuation enables simultaneous robot operation and MR imaging while ensuring minimal image artifacts.

6. Development of an MR-safe endoscopic robot system allowing flexible access to lesions in the confined oral and pharyngeal cavities. The system provides a stable platform for laser beam delivery and navigation to the targeted lesions with MRI guidance. The robot is compact and lightweight, enabling mounting on the patient and operation with standard MR imaging coils.

7. Experimental validation of intra-op MRI-guided robotic laser ablation on *ex-vivo* tissue and cadaveric oropharyngeal tissue. Intra-op MR thermometry was used to monitor the laser ablation process. A learning-based controller was employed to enable accurate visual servoing of the laser spot along a path prescribed *in situ* on the camera view.

1.3 Publications and Patents during Ph.D. Study

Journals:

1) G. Fang, M.C.K Chow, J.D.L. Ho, Z. He, K. Wang, T.C. Ng, J.K.H. Tsoi, P.L. Chan, H.C. Chang, D.T.M Chan, Y.H. Liu, F.C. Holsinger, J.Y.K. Chan, K.W. Kwok, *"Soft Robotic Manipulator for Intra-operative MRI-guided Transoral Laser Microsurgery,"* **Science Robotics**, 6(57), 2021.

2) G. Fang, X. Wang, J.D.L Ho, K. Wang , C.K. Chow, K.H. Lee, X. Xie, W.L. Tang, L. Liang, H.C. Chang, C.J. Juan, Y.H. Liu, J.Y.K. Chan, K.W. Kwok, *"Soft robotic manipulation of laser spot using epipolar geometry – an enabling technique for non-contact laser ablation,"* **Soft Robotics**, 2021 (under review).

3) G. Fang, X. Wang, K. Wang, K.H. Lee, J.D.L. Ho, H.C. Fu, D.K.C. Fu, K.W. Kwok, *"Vision-based Online Learning Kinematic Control for Soft Robots using Local Gaussian Process Regression,"* **IEEE Robotics and Automation Letters**, vol. 4, no. 2, pp. 1194-1201, 2019.

4) X. Wang, J. Dai, H.S. Tong, K. Wang, G. Fang, X. Xie, Y.H. Liu, K.W.S. Au, K.W. Kwok, *"Learning-based Visual-Strain Fusion for Eye-in-hand Soft Robot Pose Estimation and Control,"* **IEEE Transactions on Robotics**, 2021 (in the 2nd round of revision).

5) J. Dai, Z. He, G. Fang, X. Wang, Y. Li, C.L. Cheung, L. Liang, I. Iordachita, H.C. Chang, K.W. Kwok, *"A Robotic Platform to Navigate MRI-guided Focused Ultrasound System,"* **IEEE Robotics and Automation Letters**, vol. 6, no. 3, pp. 5137-5144, 2021.

6) C.L. Cheung, G. Fang, J.D.L. Ho, L. Liang, K.V. Tan, H.C. Chang, K.W. Kwok, *"Omnidirectional Monolithic Marker for Intra-operative MRI Position Sensing,"* **IEEE Transactions on Medical Imaging**, 2021 (in the 2nd round of revision).

7) T.K.W. Leung, X.D. Ji, B.Y. Peng, G.K.K. Chik, D.S.H.S. Dai, G. Fang, T. Zhang, X. Cheng, K.W. Kwok, C.O. Tsang, G.K.K. Leung, P.K.L. Chan, *"Micro-electrodes for In Situ Temperature and Bio-impedance Measurement,"* **Nano Select**, 2021.

8) Z. Dong, X. Wang, G. Fang, Z. He, J. D.L. Ho, W.L Tang, X. Xie, C.K. Ching, K.W. Kwok, *"Shape Tracking and Feedback Control of Cardiac Catheter Using MRI-guided*

Robotic Platform," **IEEE Transactions on Robotics**, 2020 (in the 2nd round of revision).

9) Z. He, Z. Dong, G. Fang, J.D.L. Ho, C.L. Cheung, H.C. Chang, C.N. Chong, Y.K. Chan, T.M. Chan, K.W. Kwok, *"Design of a Percutaneous MRI-guided Needle Robot with Soft Fluid-driven Actuator,"* **IEEE Robotics and Automation Letters**, vol. 5, no. 2, pp. 2100-2107, 2020.

10) X. Wang, G. Fang, K. Wang, X. Xie, K.H. Lee, J.D.L. Ho, W.L. Tang, J. Lam, K.W. Kwok, *"Eye-in-hand Visual Servoing Enhanced with Sparse Strain Measurement for Soft Continuum Robots,"* **IEEE Robotics and Automation Letters**, vol. 5, no. 2, pp. 2161-2168, 2020.

11) Z. Dong, Z. Guo, K.H. Lee, G. Fang, W.L. Tang, H.C. Chang, D.T.M. Chan and K.W. Kwok, *"High-performance Continuous Hydraulic Motor for MR Safe Robotic Teleoperation,"* **IEEE Robotics and Automation Letters**, vol. 4, no. 2, pp. 1964-1971, 2019.

12) X. Wang, K.H. Lee, D.K.C. Fu, Z. Dong, K. Wang, G. Fang, S.L. Lee, A.P.W. Lee, K.W. Kwok, *"Experimental Validation of Robot-assisted Cardiovascular Catheterization: Model-based versus Model-free Control,"* **International Journal of Computer Assisted Radiology and Surgery**, vol. 13, no. 6, pp. 797-804, 2018.

Conference proceedings:

1) L. Liang, C.L. Cheung, G. Fang, J.D.L. Ho, C.J. Juan, H.W. Chung, K.W. Kwok, H.C. Chang, *"Head Motion Tracking in MRI Using Novel Tiny Wireless Tracking Markers and Projection Signals,"* **International Society for Magnetic Resonance in Medicine (ISMRM) & Society for MR Radiographers & Technologists (SMRT) Annual Meeting & Exhibition**, 2021

2) Z.Y. Li, G. Fang, J.D.L. Ho, C.I. Lam, Y.W. Yim, J.Y.K. Chan, K.W. Kwok, *"Augmented Reality-Guided Visual Servoing for Flexible Endoscope Control,"* **IEEE International Conference on Robotics and Automation (ICRA) Workshop** - Open Challenges and State-of-the-Art in Control System Design and Technology Development for Surgical Robotics Systems, 2019.

Patents:

1) G. Fang, K.W. Kwok, J.Y.K. Chan, H.C. Chang, X. Wang, J.D.L. Ho, C.K. Chow, *"Visual Servoing of An MR-safe Soft Manipulator for Transoral Laser Microsurgeries,"* **US Prov. Patent**: US 63/021,692. [Filed on 8 May 2020]

2) K.W. Kwok, Z.L. He, Z.Y. Dong, J.D.L Ho, G. Fang, *"Fluid-Driven Robotic Needle Positioner for Image-guided Percutaneous Interventions,"* **US Prov. Patent**: US 63/053, 798. [Filed on 20 July 2020]

3) K.W. Kwok, Z.L. He, J. Dai, G. Fang, X. Wang, *"A Robotic Platform to Navigate MRI-guided Focused Ultrasound System,"* **US Prov. Patent**: US 63/159, 392. [Filed on 10 March 2021]

4) K.W. Kwok, Z.L. He, J.D.L Ho, G. Fang, X. Wang, K. Wang, *"Surgical Instrument with Flexible Steerable Segment,"* **US Prov. Patent**: US 53/141, 613. [Filed on 26 January 2021]

Chapter 2

Robotic Laser Surgery with MRI guidance

2.1 Introduction

Image-guided surgery [54] has become a trend in the development of surgical procedures, with pre-operative (pre-op) or intra-op images being employed to assist surgeons to perform the safer treatment. Among various imaging modalities, MRI is superior in terms of non-invasive, non-ionizing radiation, and high contrast for soft tissues [55]. MRI enables detailed visualization of the physiological changes of tissue, and even temperature perception inside tissue, making it unique and superior guidance for interventional procedures. MRI has been widely investigated in applications such as biopsy [56, 57], cardiovascular intervention [58, 59], targeted drug delivery [60], and thermal therapy for tumor ablation [61, 62].

Lasers are widely used for tumor treatment, accrediting to its capability of very precise energy deposition and thus a clear-defined boundary of the thermal ablation zone. The use of fiber-based laser probes is favorable for usage with MRI due to their intrinsic MRI-compatibility. Recently, there are increasing studies aiming to introduce intra-op MRI to enable *in-situ* evaluation of the laser ablation process, along with the development of MRI-guided robotic systems for laser probe navigation. However, the strong magnetic field in the MRI scanner prevents the use of ferromagnetic materials, thus posing significant challenges to robot system design. There are also limited choices of sensors available in the harsh MRI environment. These limitations make the precise control of the instrument

in the confined workspace of MRI bore a difficult task. This could be even challenging in non-contact laser surgery requiring frequent and dexterous manipulation of the laser beam, rather than contact ablation at local points only. However, the recent advancements of techniques like MR-conditional actuation, fast imaging sequencing, MR-based tracking, and intelligent control bring great potential in extending MRI-guided robotic intervention to more complex procedures.

This chapter will present the recent advances of intra-op MRI technologies and the new developments of MRI-guided robotic platforms. The prior arts of non-contact laser surgery will be reviewed, together with the clinical motivation of introducing MRI guidance. The latest developments of soft continuum robots and their precise motion control will be discussed, showing their potential in surgical applications. Furthermore, the motivation of bringing soft robotic manipulators to MRI-guided laser surgery will be highlighted. This chapter provides the research background for the technical works presented in this thesis.

2.2 Recent Advances of MRI

MRI is a non-invasive imaging technique that has been widely used in clinics for diagnosis, staging, and follow-up of diseases. Up to 2018, there are about 36,000 scanners being in used worldwide, with an increase of 2,500 every year [63]. MRI employs strong magnetic fields, magnetic field gradients, and radio-frequency fields to create image contrast from the spins of hydrogen protons [64], which varies with the types of tissues. There is no ionizing radiation involved, making MRI distinguished from X-ray, computed tomography (CT) and positron emission tomography (PET) scans. In addition, MRI can offer versatile imaging capabilities to provide information about anatomical (**Fig. 2.1a**), functional (**Fig. 2.1b**), and metabolic (**Fig. 2.1c**) changes [65]. These advantages prompt the application of MRI in various areas such as neuroimaging, cardiovascular, and musculoskeletal.

2.2.1 Intra-operative MRI

Compared with ultrasound or CT scans, MRI typically takes a longer time to acquire a high-quality image. Recently, much progress can be seen in terms of accelerating MRI scans. Uecker *et al.* achieved real-time image acquisition at a temporal resolution of 20ms, with the use of radially k-space encoding and fast low-angle shot (FLASH) technique [66]. The rapid development of deep learning provides opportunities to accelerate the MRI image

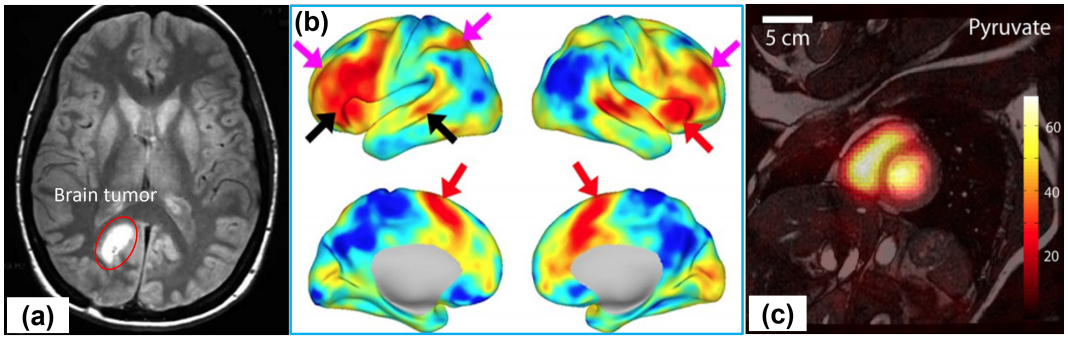


Fig. 2.1: **(a)** MR anatomical image showing the brain tumor. **Image source**: [1]. **(b)** Functional MRI response to word generation. **Image source**: [2]. **(c)** ^{13}C pyruvate image of the blood pool in cardiac chambers. **Image source**: [3].

reconstruction with greater undersampling of the k-space [67, 68]. A commercial platform (RTHawk, Heartvista®) used artificial intelligence to speed up a complete cardiac Ischemia exam from 90 minutes to 15 minutes [69]. Apart from fast imaging, it also allows real-time control of the scanner and streaming out the MRI raw data for automatic feedback. These advances could facilitate the use of MRI intra-operatively to guide interventions.

A discomfort or concern related to an MRI scan is the narrow and confined bore. The development of open MRI may offer a potential solution for future scanner settings. Currently, open MRI achieves the highest field strength of 1.2 Tesla (**Fig. 2.2a**), which was demonstrated with acceptable accuracy for guiding needle insertion [70]. Recently, there is a portable MRI system (Swoop^{TM}, **Fig. 2.2b**) cleared by the U.S. food and drug administration (FDA) for MR imaging of the brain, which is a breakthrough in terms of MRI accessibility [71]. Another advance in MRI technology is the development of a flexible receive coil (**Fig. 2.2c**), which enables application-specific design with compact size [72, 73].

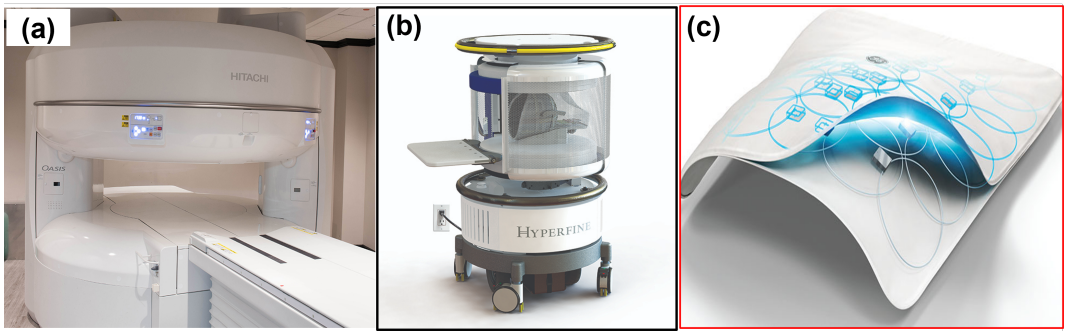


Fig. 2.2: **(a)** 1.2T Open MRI scanner from Hitachi. **Image source**: [4]. **(b)** FDA-cleared portable MR imaging system from Hyperfine. **Image source**: [5]. **(c)** Flexible MRI receive coil from GE. **Image source**: [6].

2.2.2 Setup of Interventional MRI Systems

MRI provides excellent contrast for soft tissues, capable of visualizing most lesions without using contrast agents [74]. Besides, it enables imaging of a slice in any arbitrary position and orientation, making it flexible for accessing complex regions during interventional procedures. To leveraging those benefits, intra-op or interventional MRI has been developed fast in recent years.

The world's first intra-op MRI is the GE Signa System, with a "double donut" design (**Fig. 2.3a**). It provides full patient access by arranging two opposite magnetics horizontally beside the MR table. Accrediting to its open configuration, surgeons could perform surgical procedures conveniently while standing or sitting. Intra-op MR images can be acquired without patient movement. There have been more than 1,000 neurosurgical cases conducted with this system since 1995. However, performing surgery inside the scanner needs specialized non-ferromagnetic instruments, many of which are not available. Hitachi proposed a commercial intra-op MRI system integrating an open scanner in the operating theatre room (**Fig. 2.3b**). Patients can be smoothly transited between surgical state and imaging state. However, the low field (<0.5T) offered by open-type scanners results in reduced image resolution that is not comparable to conventional diagnostic scanners (1.5T and 3T).

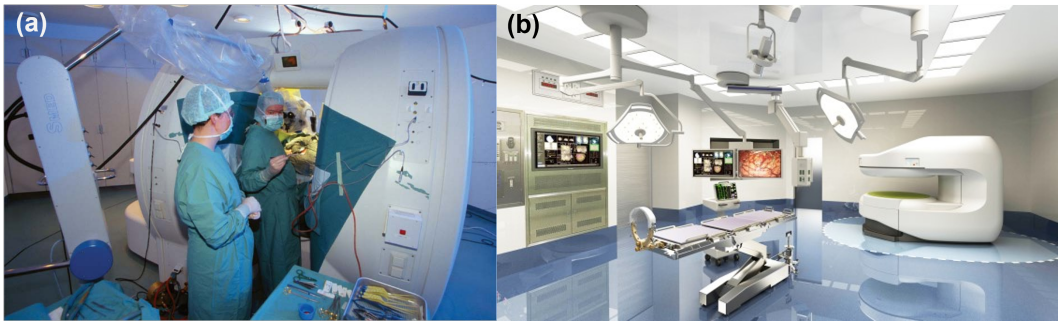


Fig. 2.3: (a) Surgeons perform surgery in a GE Signa SP system. **Image source**: [7]. **(b)** Setup of the Hitachi intra-op MRI system integrating operating theatre and MR scanner in one room. **Image source**: [8].

The other type of intra-op MRI system uses the high-field closed-bore scanner, which can provide higher imaging resolution and faster imaging speed. An example is the IMRIS hybrid operating suite (**Fig. 2.4a**). It is equipped with a ceiling-mounted rail system capable of transporting the magnet between a diagnostic room and adjacent operating rooms. Therefore, diagnostic-quality images can be obtained without transferring the patient.

However, special features, e.g., regulated airflow and radiofrequency (RF) shielding, must be included in the operating rooms. Another example is the advanced multi-modality image-guided operating (AMIGO) suite (**Fig. 2.4b**), which comprises a PET/CT room apart from the MRI room and operating room, thus offering multi-modality image guidance. The MR scanner can traverse on ceiling rails to the operating table, while the PET/CT scanner is stationary. Although these systems enabled improved safety and image qualities, the process of moving "in-and-out" patients could complicate surgical workflow and induces additional operation time.

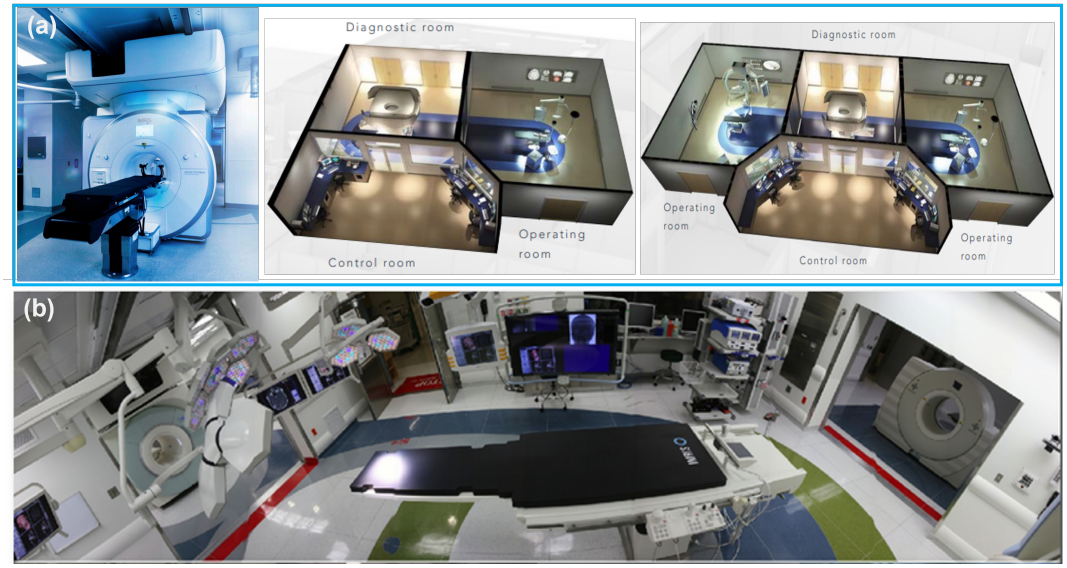


Fig. 2.4: (a) IMRIS intra-op MRI suite with double-room or three-room configurations. **Image source**: [9]. **(b)** AMIGO operating suite combining MRI and PET/CT for image-guided therapy. **Image source**: [10].

2.2.3 MR Compatibility of Interventional Devices

To create a smoother surgical workflow, operations may need to be performed inside or near the MRI scanner, during which MR imaging guidance can even be obtained in real-time to assist the procedure. However, the strong magnetic and radiofrequency field of MRI raises special safety issues to the instruments being used. Potential hazards can be caused by the magnetic forces or torques generated on metallic objects. Besides, the radiofrequency pulses may induce the heating of electrical cables, needles, and implants. Surgical instruments could also introduce inhomogeneity of the magnetic field, resulting in imaging artifacts, which may be hazardous in MRI-guided interventions. Therefore, any devices designed for usage in MRI should undertake rigorous testing following the protocols defined by the American Society for Testing and Materials (ASTM) F2503 [75]. A comprehensive

Table 2.1: Classification of MR labels in ASTM F2503 standard

Label	Term	Description
	MR Safe	An item that poses no known hazards resulting from exposure to any MR environment. MR Safe items are composed of materials that are electrically nonconductive, nonmetallic, and nonmagnetic.
	MR Conditional	An item with demonstrated safety in the MR environment within defined conditions. At a minimum, address the conditions of the static magnetic field, the switched gradient magnetic field and the radiofrequency fields. Additional conditions, including specific configurations of the item, may be required.
	MR Unsafe	An item which poses unacceptable risks to the patient, medical staff or other persons within the MR environment.

classification of the requirements for three categories of devices is also listed (**Table 2.1**), i.e., MR-safe, MR-conditional and MR-unsafe.

In many interventional procedures, e.g., laser ablation, robotic assistance plays an important role in improving targeting accuracy and thus the surgical outcome. However, conventional robotic systems commonly use metallic components and electric motors, hindering their migration to MRI-guided interventions. This prompts the development of MR-safe or conditional robotic systems. In the following section, the representative works and recent advances about MRI-guided interventional devices will be introduced, with our focus on laser surgery.

2.3 Recent Development of MRI-guided Laser Therapy

MRI-guided laser interstitial thermal therapy (LITT) [76–78] has been widely used to treat lesions in various organs such as the brain, liver, and prostate. Laser energy is delivered using an optical fiber to target and ablate lesions selectively. Compared to other ablation techniques, such as ultrasonic energy [79], argon beam coagulator [80] and radiofrequency energy [81], lasers provide precise energy deposition and ablation margin, as well as effective hemostats. Intra-op MRI guidance is proved to bring tremendous benefits to LITT,

in terms of improving targeting accuracy and monitoring the necrotic zone being ablated. There are also commercial laser therapy systems that aim to utilize MR guidance, such as the TRANBERG® [82] and Visualase® [83] system.

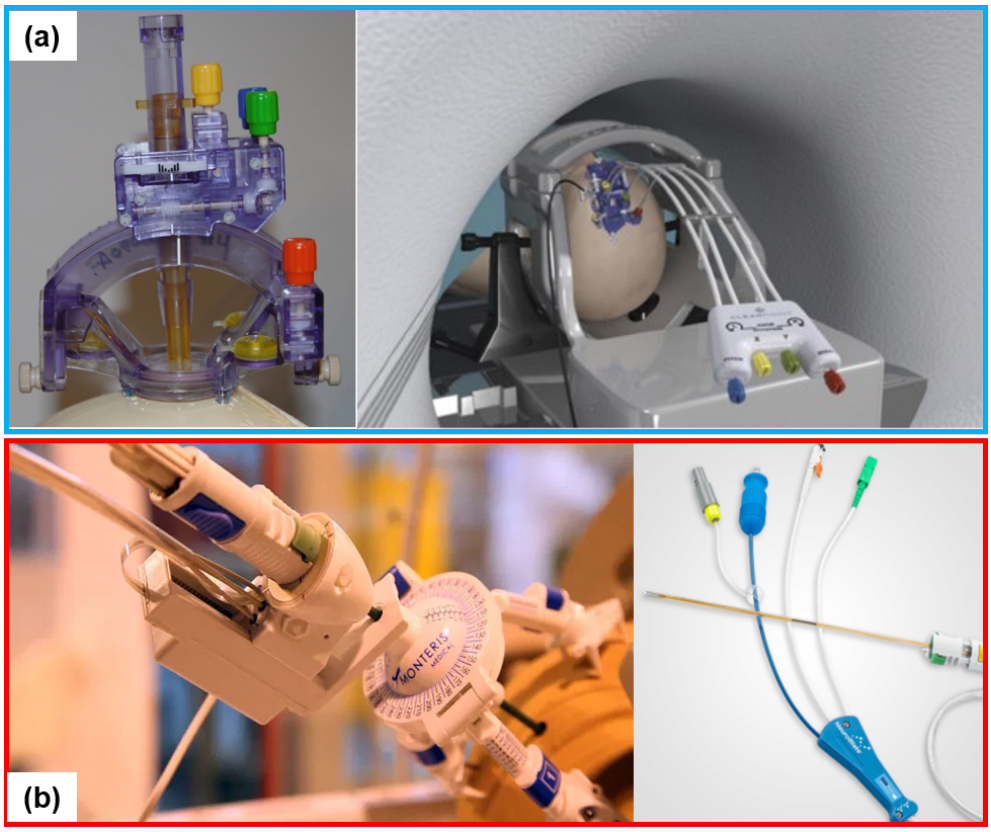


Fig. 2.5: **(a)** Clearpoint® system for stereotactic needle placement. **Image source**: [11]. **(b)** Neuroblate® system featuring robotic manipulation of laser probe. **Image source**: [12].

2.3.1 Neurosurgery

One of the earliest applications of MRI-guided intervention is neurosurgery, including deep brain stimulation [84] and LITT. The common feature of these two procedures is the use of needle-based instruments to deliver electrical impulses or laser energy. To accurately locating 3D targets within the brain, stereotaxy is widely used for positioning the instruments. A number of MR-safe/conditional stereotactic systems are also proposed to enable intra-op MRI-guided stereotaxy.

There are two commercial stereotactic systems that target MRI-guided laser therapy. The first one is the ClearPoint® system (ClearPoint Neuro Inc., USA) (**Fig. 2.5a**), which integrates a single-use frame device for instrument alignment with MRI guidance [85]. It

allows manual adjustment of the frame trajectory from outside of the MRI bore. The other one is the Neuroblate® platform (Monteris Medical Inc., USA) (**Fig. 2.5b**), which enables two degrees-of-freedom (DoFs) robotic manipulation of the laser probe [86]. Piezoelectric motors were employed to achieve rotation and insertion of the probe. The orientation of probe is adjusted manually through a separate stereotactic frame (AXiiiS), comprising three linear translating legs and a ball unit.

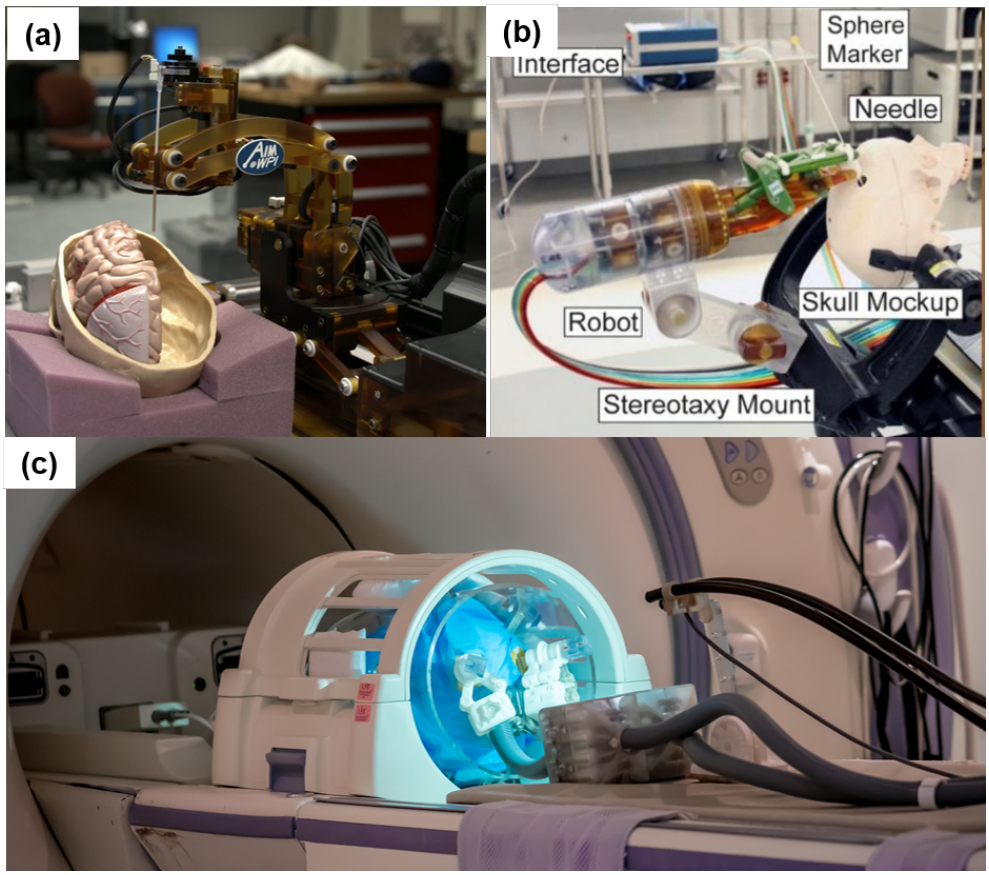


Fig. 2.6: (a) Piezoelectric-driven 7-DoFs robotic system for stereotactic neurosurgery. **Image source**: [13]. **(b)** Pneumatic-driven needle-guide robot allowing 2-DoFs orienting of needle. **Image source**: [14]. **(c)** Hydraulic-driven robotic system enabling bilateral instrument navigation [15].

Many research prototypes have also been proposed to achieve robotized stereotaxy under MRI. Fisher *et al.* [13] designed a 7-DoFs robot with a remote-of-center motion (RCM) mechanism (**Fig. 2.6a**), enabling automatic positioning, orientating, and insertion of a needle. Each DoF of motion was driven by a piezoelectric motor. No significant reduction ($\leq 10.3\%$) in signal-to-noise ratio (SNR) was found during the robot's operation. Stoianovici *et al.* [14] developed a needle positioning robot with two motorized rotational DoFs for orientating an RCM mechanism (**Fig. 2.6b**). Each DoF was actuated by a custom-designed pneumatic motor. The RCM location can be adjusted using a passive

arm. Both two robotic systems need to be mounted on the MR table. Guo *et al.* [15] presented a robot allowing bilateral stereotactic procedures (**Fig. 2.6c**). Two needles can be oriented simultaneously using two five-bar parallel mechanisms. The robot was driven in a master-slave manner using 10-meters long hydraulic transmission between the control room and MRI room. Its compact design allows patient-mounting setup, as well as operating inside a standard imaging head coil.

2.3.2 Percutaneous Interventions

Percutaneous ablation is performed by inserting a needle through the patient's skin towards the target anatomy [87]. It is widely employed in the treatment of tumors inside numerous organs, such as the breast [88], kidneys [89], and liver [90]. To ensure accurate ablation margin ($\leq$ 1 cm) and avoid inadvertent organ damage, MRI has been investigated to monitor the ablation procedures, taking advantage of MRI's high soft-tissue contrast. Laser applicators ensure MR safety, enabling MR imaging to be conducted simultaneously with the ablation. However, the success of the ablation procedure still depends on the targeting accuracy of needles, which drives the development of MR safe/conditional needle placement robots.

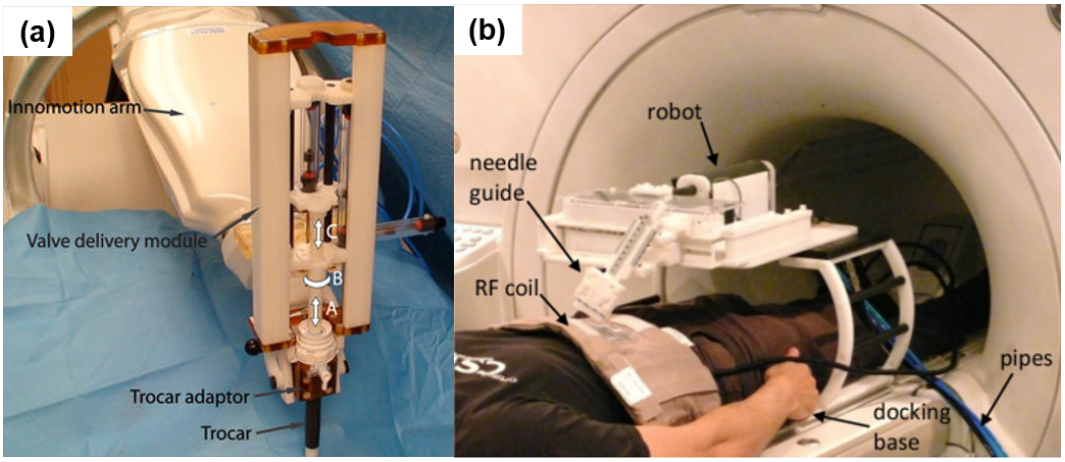


Fig. 2.7: (a) Innomotion robotic system for MRI- and CT-guided needle placement. **Image source**: [16]. **(b)** Pneumatic-actuated robot for MRI-guided laser ablation of liver. **Image source**: [17].

Table-mounted systems are proposed to achieve high-accuracy needle positioning, benefiting from their rigid structure. Innomotion (Innnomeic Inc., Herxheim) is a commercial robotic system designed for MRI- and CT-guided needle positioning [16] (**Fig. 2.7a**). It utilizes pneumatic cylinders to achieve 5-DoFs manipulation of the

needle. Franco *et al.* designed a 4-DoFs pneumatic-actuated needle-positioning robot, and conducted a pilot study of MRI-guided laser ablation on two patients [17] (**Fig. 2.7b**). A gantry was used to support the robot frame above the patients. The multi-imager compatible needle-guided robot proposed by Stoianovici *et al.* also allows MRI-guided percutaneous intervention [91]. However, the table-mounted systems generally have a large footprint and need to work with a custom MRI body coil. Potential needle movement induced by patient respiration may lead to a safety hazard.

Patient-mounted systems can move with the human body, thus offering improved safety related to patient motion. Li *et al.* proposed a body-mounted robotic assistant for MRI-guided low back pain injection [18] (**Fig. 2.8a**). The robot provides 4-DoFs manipulation of the needle with the use of piezoelectric motors. Together with a mounting frame, the robot has a dimension of 219 mm×250 mm×87 mm, and a weight of 1.3 kg. He *et al.* designed a semi-automated robotic system for MRI-guided percutaneous needle procedures [19] (**Fig. 2.8b**). Three soft fluid-driven actuators were employed to achieve fine adjustment of needle orientation. Accredited to the compact size (Ø108 mm × 115 mm height) and lightweight (189g), the robot enables patient mounting and operating with standard loop coils. Several robots can even be used simultaneously to enable multiple needle targeting.

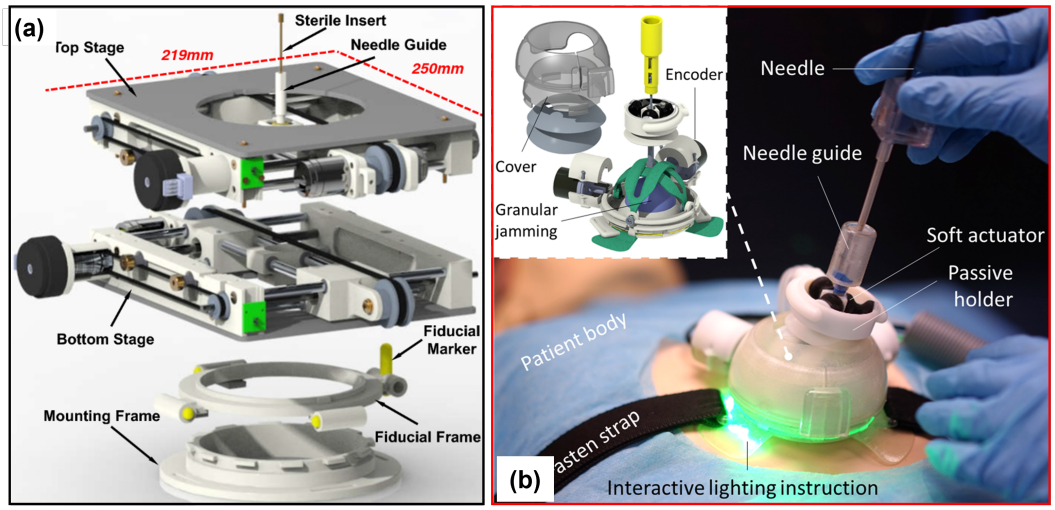


Fig. 2.8: **(a)** Body-mounted robotic assistant offering 2-DoFs translating and 2-DoFs orientating of needle. **Image source**: [18]. **(b)** Compact needle-placement robot allowing manual coarse-adjustment and robotic fine-tuning of needle orientation [19].

2.4 Non-contact Laser Surgery in Head-and-Neck Cancer Treatment

Non-contact laser surgery is conducted by projecting a laser beam to tissues, creating an incision or ablation effect. Optical lenses are commonly used to direct and focus the output beam, thus providing a small laser spot with a high power density. To achieve a precise treatment margin, the laser probes need to be frequently manipulated during the ablation process. Although the working mode of contact (e.g., interstitial) *versus* non-contact lasers may be different, they share the same principle to apply thermal energy to create damage in tissue, which is totally dependent on tissue temperature rise and its duration [92].

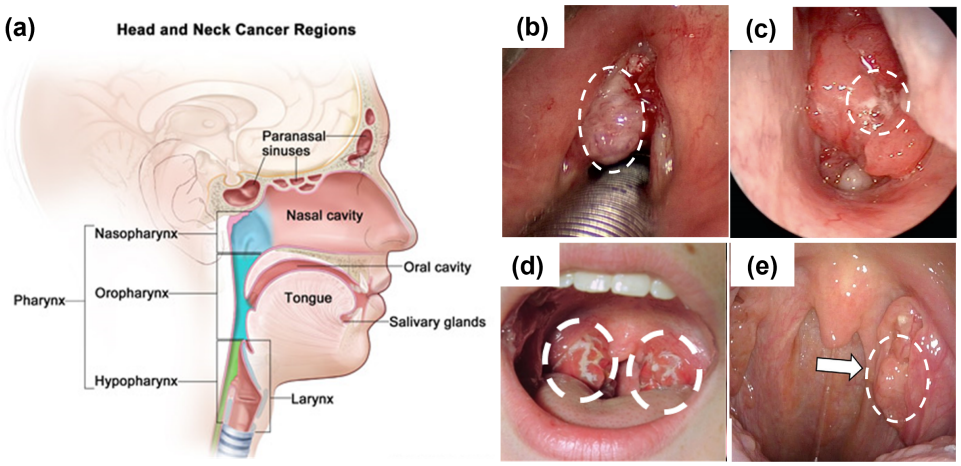


Fig. 2.9: **(a)** Anatomy of head and neck regions. **Image source**: [20]. **(b)** Larynx cancer near the vocal cords. **Image source**: [21]. **(c)** Nasopharyngeal cancer. **Image source**: [22]. **(d)** Pharynx cancer in the tonsil wall. **Image source**: [23]. **(e)** HPV-associated oropharyngeal cancer. **Image source**: [24].

2.4.1 Head and Neck Cancer

Head and neck cancer (HNC) (**Fig. 2.9**) is the seventh most common cancer in the world, causing more than 450k deaths every year [93]. As these cancers are located in regions associated with critical functions (e.g., respiration, phonation, and deglutition), traditional open surgery or adjuvant therapy often results in significant patient morbidity, including dysphagia and dysphonia. Transoral surgery has become a dominant alternative by performing tumor resection through the mouth, primarily aiming at organ preservation [94]. However, this approach can still be associated with considerable detriment to the patient due to the lack of specialized instrumentation to facilitate working in these confined spaces.

Ablation technologies enable local energy delivery to targeted lesions inside the body [95], and thus allowing localized and controlled ablation for minimally invasive surgery. Among them, in particular, for smaller lesions, lasers can provide precise energy deposition and ablation margin, as well as effective hemostasis.

Laser therapy has been used in the treatment of a variety of head and neck cancers [96,97], with laser ablative therapy most commonly used in the treatment of glottic cancers. Laser ablation in glottic cancers with ultranarrow margins has been shown to be effective in preserving vocal function without the need for further adjuvant radiotherapy [98]. This can be used for even bulky tumors that are early stage, with the advantage of minimal thermal spread to the surrounding normal tissue [99]. Now with the Human Papilloma Virus (HPV)-associated oropharyngeal cancer epidemic where improved outcomes and narrower margins for resections are needed, the use of laser for resections has been shown to improve pain scores, reducing hospital stay and time to oral feeding when compared to the use of traditional cautery [100].

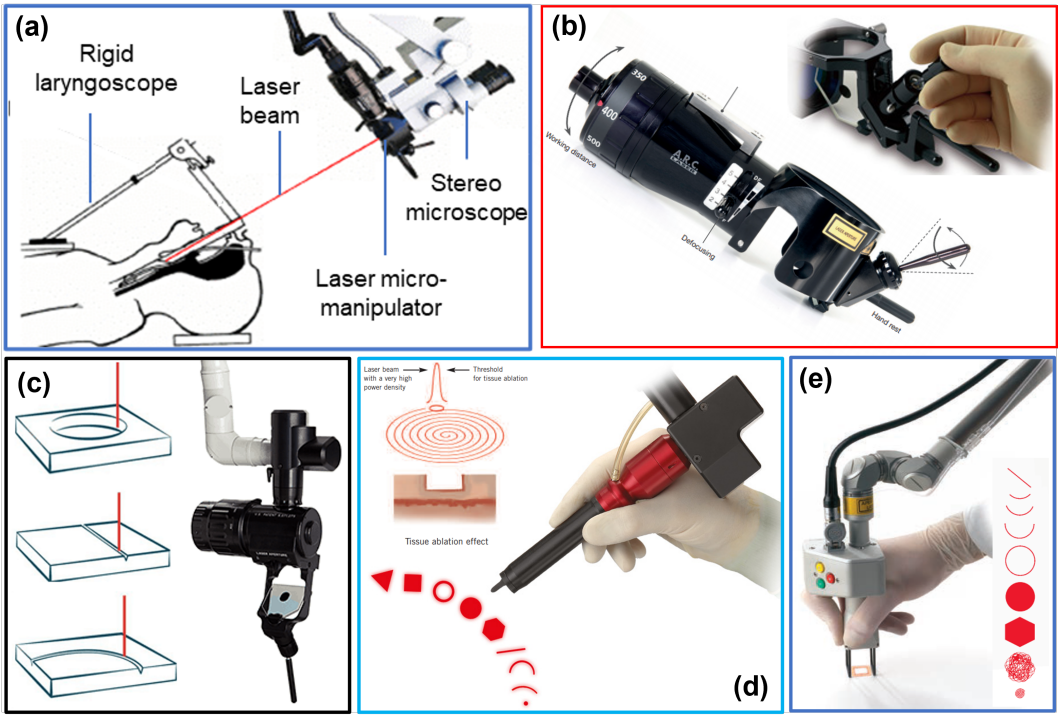


Fig. 2.10: **(a)** Setup of current transoral laser surgery. **Image source**: [25]. **(b)** Laser micro-manipulator with a mirror rotated by joystick. **Image source**: [26]. **(c)** Luminens AcuBlade laser scanner with pre-programmed patterns. **Image source**: [27]. **(d)** KLS Martin SoftScan digital scanner system. **Image source:** [28]. **(e)** DEKA HiScan system enabling high-precision scanning shapes. **Image source**: [29].

2.4.2 Transoral Laser Microsurgery

With the development of minimally invasive surgical techniques and medical lasers, TLM has attracted growing attention in the treatment of HNCs (**Fig. 2.9**) [101]. Previous oncological studies [102–104] have demonstrated TLM's excellent performance in treating laryngeal tumors with the minimal adverse influence of vocal cord function and reasonable functional voice outcomes. TLM performs laser beam projection trans-orally for tumor resection, eliminating external incisions as in conventional open-neck surgeries [105]. The high-power laser beam needs to be projected at a small size ($\sim$250 μm, [106]) and manipulated precisely ($<$1 mm, [107]) in the narrow laryngeal area ($\sim$20×20 mm^2). It is demanding in terms of accuracy due to the high risk of damage to the vocal muscles [108]. This also challenges the development of laser steering devices and corresponding control schemes.

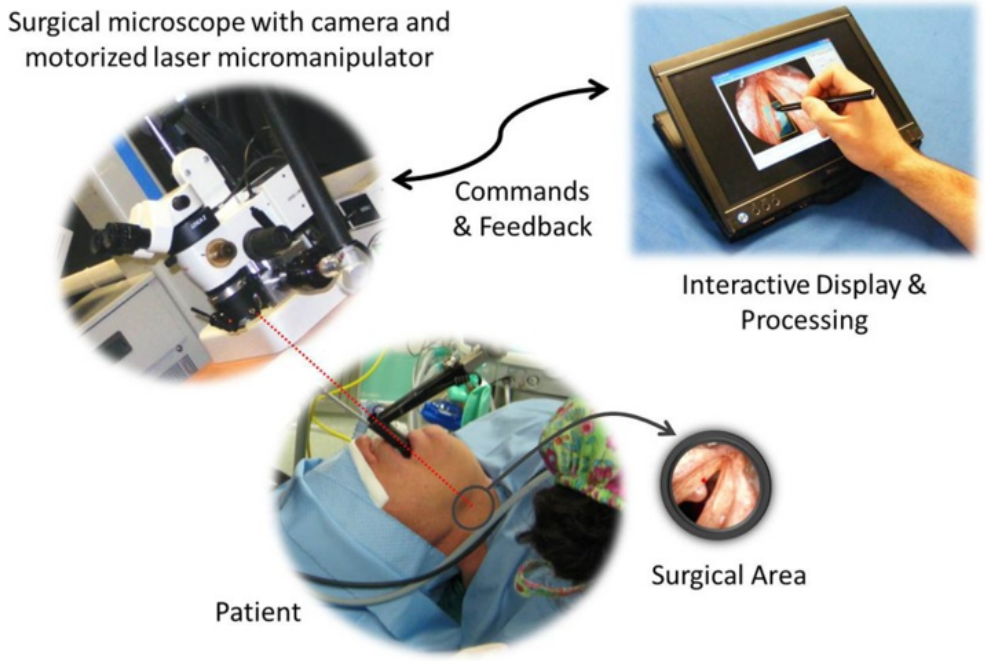


Fig. 2.11: Virtual scalpel concept using a stylus for interactive display and laser control. **Image Source**: [30, 31]

Carbon dioxide (CO_2) laser ($\lambda \simeq$10.6 μm) with a high absorption rate in soft tissues is the most common laser source in TLM. The lasers are typically delivered and maneuvered via a set of external prisms and mirrors, which provides laser beam projection and steering at the surgical site. In a conventional TLM setup, surgeons manually operated a joystick to manipulate a beam-splitter mirror that projects the laser beam at the targeted lesions [105] (**Fig. 2.10a-b**). However, such manual manipulation requires the surgeons

to master high psychomotor skills to achieve dexterous hand-eye-foot coordination. Recently, the motorization of external manipulators was introduced to improve the laser aiming precision and efficiency [109]. Motorized laser scanners (e.g., Lumenis AcuBlade (**Fig. 2.10c**), KLS Martin SoftScan (**Fig. 2.10d**), DEKA HiScan (**Fig. 2.10e**)) allowing pre-programmed scanning patterns are commercially available, providing improved laser incision quality. However, these systems still need surgeons to position the pattern using a manual micromanipulator. New motorized micromanipulators were proposed to enable the incorporation of a stylus-based machine-surgeon interface to release surgeons from looking through the microscope during laser manipulation (**Fig. 2.11**), thus offering improved ergonomics [30, 31]. However, these laryngeal laser surgery setups are mainly based on the design of straight-line laser beam projection from an external manipulator to the surgical site (**Fig. 2.10a**). To ensure sufficient exposure of tumors to the laser beam, patients may need to maintain a high extension of the neck. Moreover, the long working distance ($\sim$400 mm) causes a leveraging effect on the laser beam, which can amplify inaccuracies in laser steering.

2.4.3 Transoral Endoscopic Surgery assisted with Robots

Lasers with a shorter wavelength, e.g., thulium ($\lambda \simeq$2,013 nm) and blue light ($\lambda \simeq$445 nm) laser, were introduced to overcome the limitations of "line-of-sight". They can be coupled into a flexible optical fiber to enhance access to surgical targets and achieve efficient tissue dissection as well as excellent hemostasis [110]. The development of hollow-core optical fiber also enabled the coupling of the commonly-used CO_2 laser. Fiber-optic lasers are already employed in clinical practice [111–113]. However, the difficulties in manual manipulation of flexible optical fibers have also led to its incorporation with surgical robotics systems.

Currently, there is a single transoral surgical robotic platform that is FDA-approved for transoral minimally invasive surgery (da Vinci, Intuitive Surgical Inc., USA) [114]. Compared to the daVinci SI system using four individual rigid instruments (**Fig. 2.12a**), the da Vinci SP system integrates four articulated instruments into a single cannula (**Fig. 2.12b-c**), thus enabling deep and narrow access. Although these systems offer high quality vision and control for surgeons, the high setup cost and long training time limit their availability. An alternative hybrid approach using endo-laparoscopic instruments and a robotic Flex® camera (Medrobotics, USA) (**Fig. 2.12d**) has also been approved for

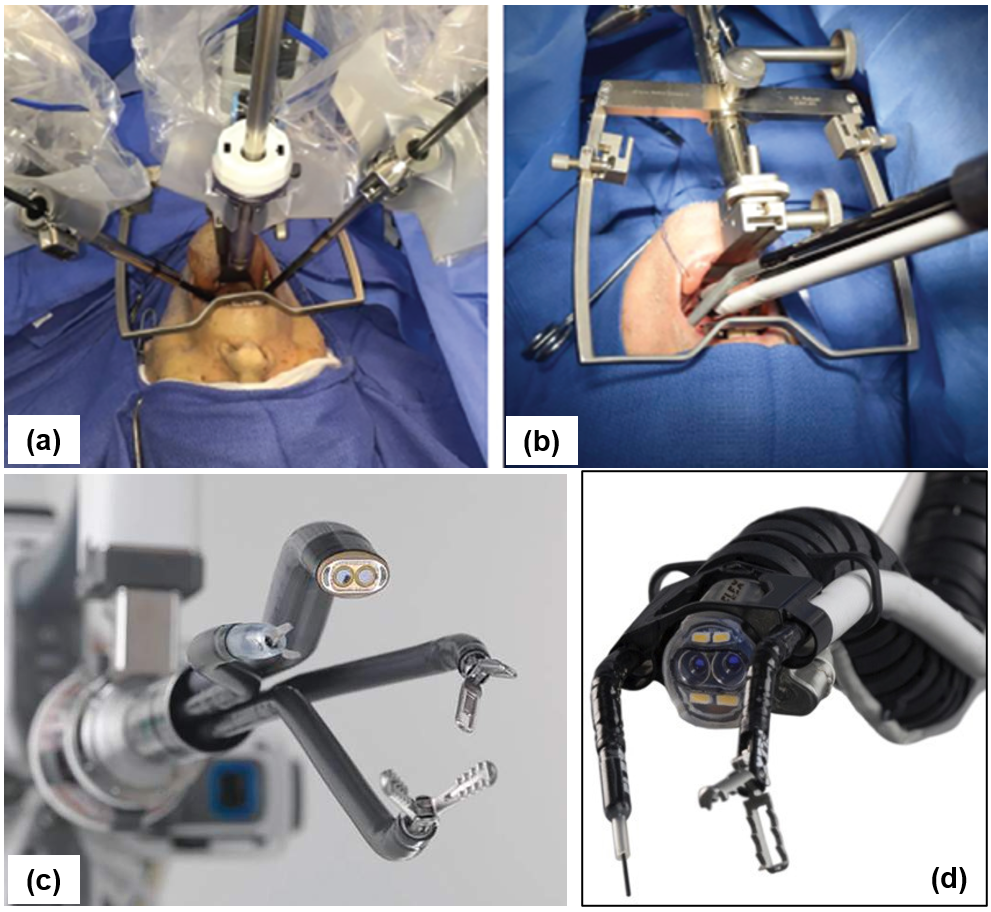


Fig. 2.12: (a) Setup of da Vinci SI system for transoral surgery. **Image source**: [32]. **(b)** Transoral setting with da Vinci SP robot. **Image source**: [32]. **(c)** Endowrist instruments of da Vinci SP surgical system. **Image source**: [33]. **(d)** Flex® robotic endoscope comprising two flexible instruments. **Image source**: [34].

clinical use for head and neck surgery [115]. Two pilot studies [114, 116] have evaluated the feasibility of combining transoral robotic surgery (TORS) with flexible fiberoptic laser technology, demonstrating enhanced precision, fine cutting margins, reduced postoperative pain, and increased operation safety.

2.4.4 Endoscopic Laser Manipulator

The advances of flexible medical laser bring the opportunity to access lesion targets located in the deeper regions, which also promotes the development of endoscopic laser manipulators. Renevier *et al.* [35] designed a two-DoFs actuated micromirror based on a micro-electro-mechanical system (**Fig. 2.13a**), which could be integrated into an endoscopic tip. However, the inclination of reflecting mirrors may result in an elliptical laser spot. Acemoglu *et al.* [36] proposed a flexible robotic laser scanner, which used four electromagnetic (EM) coils to manipulate a laser fiber (**Fig. 2.13b**), while the projection

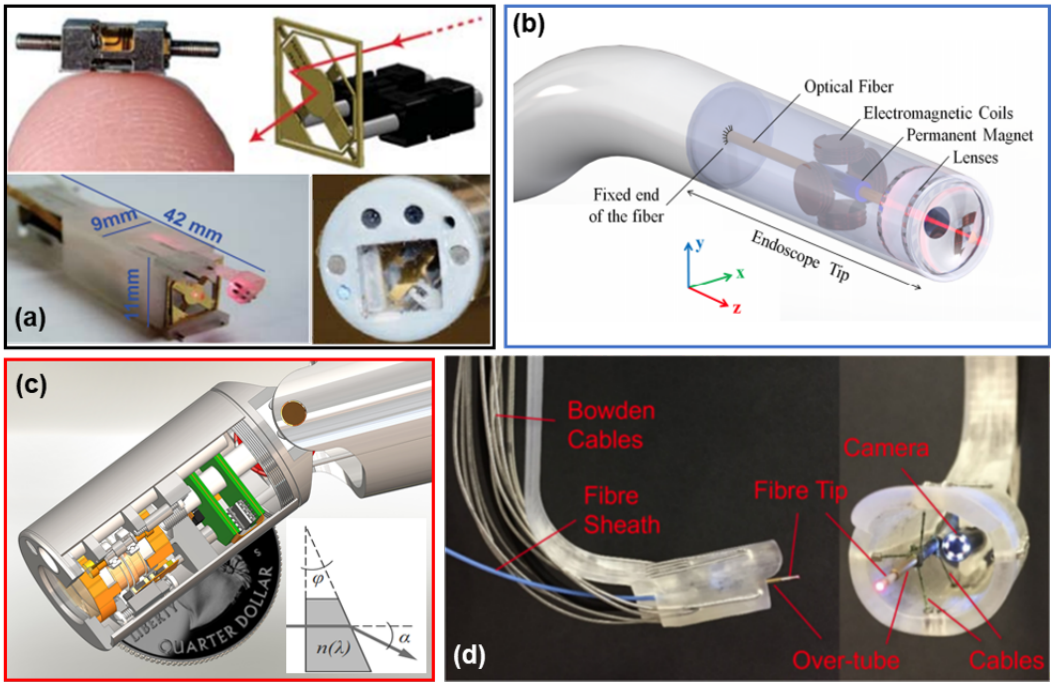


Fig. 2.13: (a) Endoscopic laser steering using an actuated mirror. **Image source**: [35]. **(b)** Laser scanner manipulating a laser fiber by four electromagnetic coils. **Image source**: [36]. **(c)** Laser scalpel with Risely prism beam steering system. **Image source**: [37]. **(d)** Cable-driven mechanism for optical fiber manipulation. **Image source**: [38].

workspace of laser spot was restricted in a 5 × 5 mm square. A 5-cm long rigid laser scalpel using the Risley prism beam steering mechanism was designed (**Fig. 2.13c**), which could be mounted on a modified laryngoscope [37]. An extensible continuum robot has been investigated to achieve focal adaptation for non-contact laser surgery [117]. A cable-driven parallel mechanism was proposed to achieve 5-DoFs manipulation of an optical fiber tip (**Fig. 2.13d**), which has to maintain a close distance to tissue, e.g., 1 mm for incision and 2-5 mm for ablation and hemostasis [38].

2.4.5 Clinical Motivations of MRI-guided Transoral Laser Surgery

Despite robot assistance, surgeons must still depend on their experience to approximate a safe 'deep margin' while avoiding injury to crucial neurovascular or muscular structures. In normal surgical settings, such intra-op evaluation can only be performed through time-consuming and costly pathology laboratory procedures such as FSA [118] (**Fig. 2.14a-b**). While frozen section performed immediately after resection is the current gold-standard approach, margin assessment performed simultaneously and in real-time as the surgeon treats the tumor would present a significant advance.

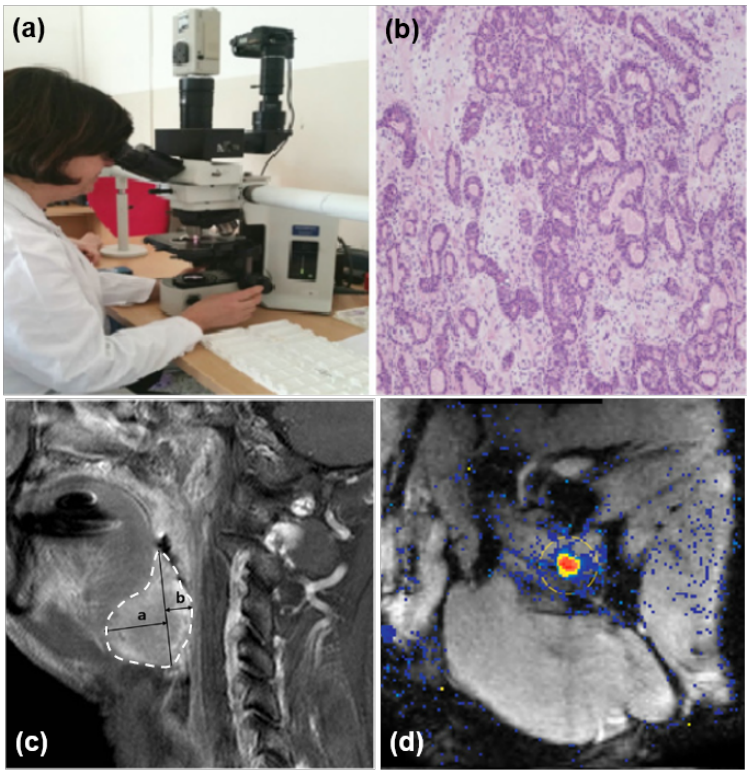


Fig. 2.14: (a) Microscopic setup for intra-op FSA. **Image source**: [39]. **(b)** Pathology result of a frozen section. **Image source**: [39]. **(c)** Tumor depth measured on an anatomical MR image of tongue base. **Image source**: [40]. **(d)** Temperature map acquired by MR thermometry during high intensity focused ultrasound ablation. **Image source**: [41].

Recently, online estimation of tissue temperature [119] and ablation depth [120] based on modeling the laser-tissue thermal interaction were investigated. However, this method requires collecting actual data from the laser ablation process on *ex-vivo* tissues, which could be inaccurate for *in-vivo* applications. An alternative is to introduce an intra-op sensing modality to monitor the laser ablation process. MRI has gained prevalence as an intra-op image guidance modality because it can clearly contrast cancerous tissue as well as critical neurovascular structures (**Fig. 2.14c**). MRI can form high-resolution 3D surgical roadmaps for pre-operative planning, delineating early stage (T1/T2) carcinoma [121], while real-time MRI can readily visualize the physiological changes of tissue due to successful or incomplete laser cutting/ablation. The live visualization of fine changes in tissue temperature ($<$1 °C) is made possible by MR thermometry [122], which has been widely adopted for thermal dose monitoring in MRI-guided focused ultrasound procedures [123], radiofrequency ablation [124], and interstitial laser ablation [125]. Closed-loop control has also been demonstrated by utilizing MR thermometry feedback [126, 127]. Some examples of commercial laser therapy systems that utilize real-time MR thermometry include the Tranberg system (Clinical Laserthermia Systems Americas Inc.) and the

NeuroBlate system (Monteris Medical, Inc.). Pairing anatomical and thermal imaging, MRI stands to be the unique modality that can provide accurate intra-op monitoring of ablation depth (**Fig. 2.14c**) relative to the underlying critical tissue, as well as the extent of thermal diffusion (**Fig. 2.14d**), which could facilitate the modulation of laser delivery through different parameters, e.g., laser pulse width, distance, and power.

During the laser ablation, real-time (up to sub-second [128]) MR thermometry can be conducted to measure the temperature diffusion in tissues, offering online monitoring of tumor ablation margin, which can be even overlaid on the camera view to alert overheating to surroundings. After each ablation run, an intra-op anatomical MR scan can be performed for the *in-situ* assessment of the completeness of lesion cutting/ablation. This may eliminate the need for FSA, smoothening the surgical workflow, shortening operative time, and improving patient outcomes.

2.5 Perspectives towards MRI-guided Non-contact Laser Manipulation

Although many MRI-guided robotic systems have been proposed for procedures like percutaneous ablation and neurosurgical surgery (**Section 2.3**), these systems are primarily designed for contact or needle-based ablation wherein the laser tip is inserted through a single entry point in the skin and directly into target lesions where it remains stationary during the ablation process. Any positional adjustments are made between subsequent ablations. At present, no robotic platform is able to leverage real-time intra-op MRI guidance together with continuous control of laser beam projection for tumor ablation. This would be even challenging within constrained anatomy such as the oral and pharyngeal cavities. A new generation of MR safe/conditional robots that features compact size, high accuracy, and dexterous manipulation is expected. In this section, the enabling technologies for MRI-guided non-contact laser manipulation will be discussed, which also shape the main focus of this thesis.

2.5.1 MR-compatible Actuation

The performance of actuators can determine the accuracy, responsiveness, and payload of the whole robotic system. However, the high magnetic field in the MRI environment

prohibits the use of ferromagnetic materials, thus posing great challenges to the choice of actuation methods. Electromagnetic motors commonly used in conventional robotics are not allowed. Besides, the actuators should maintain high performance after long-range ($\sim$10 m) power transmission between the control room and MRI room.

To ensure a safe actuation for robotic systems in the MR environment while minimizing the interferences, various MR conditional/safe actuators were developed, such as piezoelectric motors [129, 130]. However, the high-frequency current exciting the motor stepping action may inevitably induce artifacts on intra-op imaging. This can be solved by putting actuators away from the imaging site or iso-center during their operation, which, however, increases the robot footprint. Inherently MR-safe motors powered by fluid, e.g., pneumatic motors [131–133], have also been proposed. However, the high compressibility of air may cause mechanical transmission delay and thus control inaccuracies. Hydraulic actuation [134, 135] was also investigated. The hydraulic system is accomplished with incompressible fluid as transmission media, offering quick response, accurate control, and high-power density. However, the fluid transmission system via long pipelines should be optimized to reduce any bubbles and leakage. In general, the response of fluid actuation could be nonlinear, which is hard to model, raising difficulties to the precise control of robot end-effector.

2.5.2 Real-time Positional Tracking in MRI

Accurate localization of interventional tools in MRI enables closed-loop control of robotic intervention, which plays an important role in improving targeting accuracy. However, conventional instruments are difficult to be tracked from MR images due to the image artifacts, while image acquisition is also time-consuming. The harsh MR requirement also results in limited choices of external sensors. To address this challenge, MR-based tracking markers were developed to achieve positional localization in MRI.

MR markers can be classified into three types: passive, active, semi-active. Passive tracking markers contain contrast agents that alter local MR signal intensity by shortening the T1 or T2 value [136, 137], or paramagnetic materials that introduce local magnetic field distortion [138, 139]. Although passive marker can avoid RF safety hazards, it generally involves complicated MR sequences to identify marker signals unambiguously from background signals [140], or distinguish signals from multiple markers when they are closed to each other. Active tracking marker utilizes MR-compatible radiofrequency

circuit wired to MR scanner. The major advantage of active markers is that fast and accurate tracking can be conducted with multiple markers connected to the scanner receiver channels. However, the long electrically conductive wire connection would induce RF heating [141], or cause difficulties in clinical implementations [142]. Recently, semi-active markers with tuned circuits have attracted increasing interest for MR-based tracking [143, 144]. They do not require an electrically conducting wire connection to the MRI scanner, thus reducing the implementation complexity, and the risk of RF-induced heating due to wiring.

2.5.3 Continuum Robotic Manipulator

Taking the confined MRI bore into account, a compact manipulator is in demand to enable dexterous manipulation. To access the deep regions within the narrow oral and pharyngeal anatomy, the manipulator should also provide safe and flexible navigation. These challenges drive us to investigate the development of flexible and continuum robots.

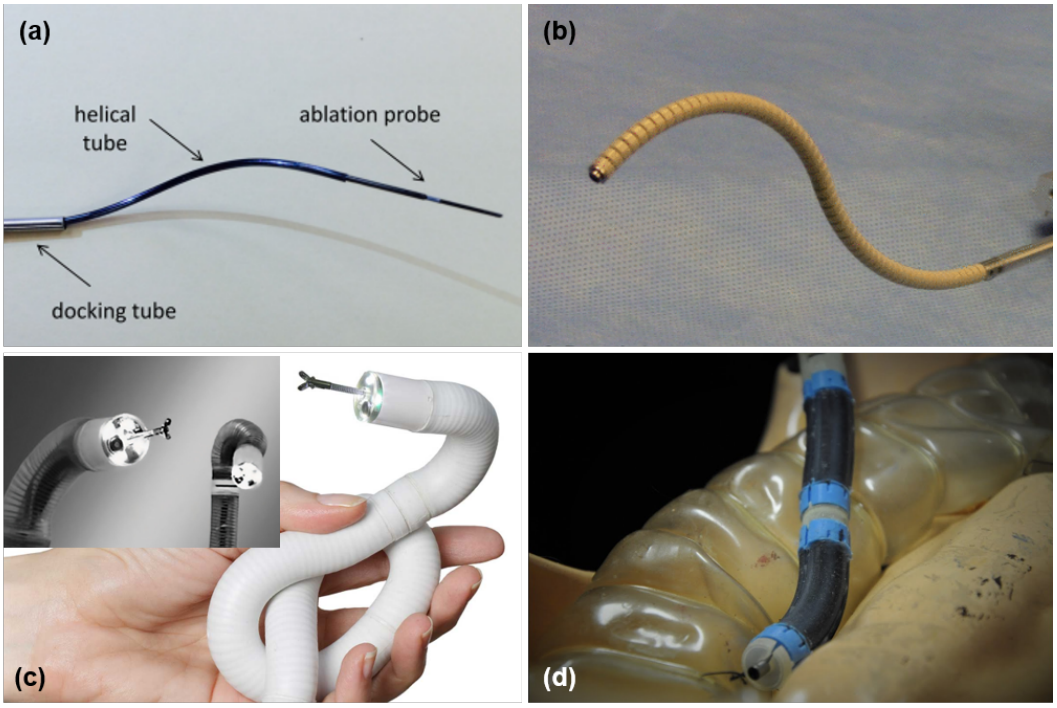


Fig. 2.15: (a) Concentric tube robot for laser probe navigation. **Image source**: [42]. **(b)** Tendon-driven continuum robot for neuroendoscopy. **Image source**: [43]. **(c)** Endotics soft robotic colonoscopy. **Image source**: [44]. **(d)** STIFF-FLOP soft robotic endoscopy for keyhole surgery. **Image source**: [45].

Robot designs based on concentric tubes (**Fig. 2.15a**) were introduced to provide laser probe targeting in MRI [42, 145]. Concentric tube manipulators generally require sophisticated actuation settings to generate both rotational and linear motion, which also

have to be located close to the robot end-effector to minimize losses in transmission. This induces difficulty in the design of compact robots and their actuator setups, particularly when the robot must be accommodated and operated inside the standard MRI head/body coil. This is made further challenging by the common use of piezoelectric-based motors, which should not be placed near the imaging site to prevent EM-artifact generation on images. This makes the external actuation system relatively heavy and large, which may require mounting to the MRI table, impacting surgical workflow and patient access. Tendon-driven snake-like robots [38, 107, 146] (**Fig. 2.15b**) face similar challenges with actuator compactness, and are typically not MRI-compatible.

Soft robots made of elastomeric materials [147, 148] have attracted increasing research interest. This is accredited not only to their high-power density actuation [149], but also their adaptability with confined and unstructured surroundings. Soft robots are low-cost, disposable, and easily sterilized, making them a viable choice in single-use endoscopic applications. Furthermore, soft robotic manipulators can be fabricated [150] using MR-safe materials such as polymers [19], and allow rapid prototyping with customizable, patient-specific designs. Commercial products such as Endotics [151, 152] (**Fig. 2.15c**) and Aer-O-Scope [153] are examples of exploiting soft robotic technologies for disposable, safe, and cost-effective colorectal screening purposes. Recently, a stiffness controllable soft robotic endoscope (STIFF-FLOP, **Fig. 2.15d**) [45, 154] was validated in human cadavers, demonstrating a superior field of vision in comparison to the use of standard laparoscopic vision in a mesorectal excision task [155].

In the case of MRI-guided non-contact laser surgery, soft robotic manipulators driven by pressurized fluid flow show great potential. Hydraulic actuation can be employed to ensure intrinsic MR safety as well as fast mechanical transmission response. Integrated with optical laser fiber, soft robots can provide flexible laser energy delivery and laser beam manipulation.

2.5.4 Precise Motion Control for Soft Continuum Robots

There are two main reasons posing challenges to the precise motion control of soft robots. First of all, the robot fabrication process could introduce uncertainties to robot geometry as well as its mechanical property. Secondly, soft robots feature nonlinear deformation upon actuation due to the intrinsic nonlinear strain-stress relation of elastomeric materials. To develop effective control strategies, several models have been investigated to approximate

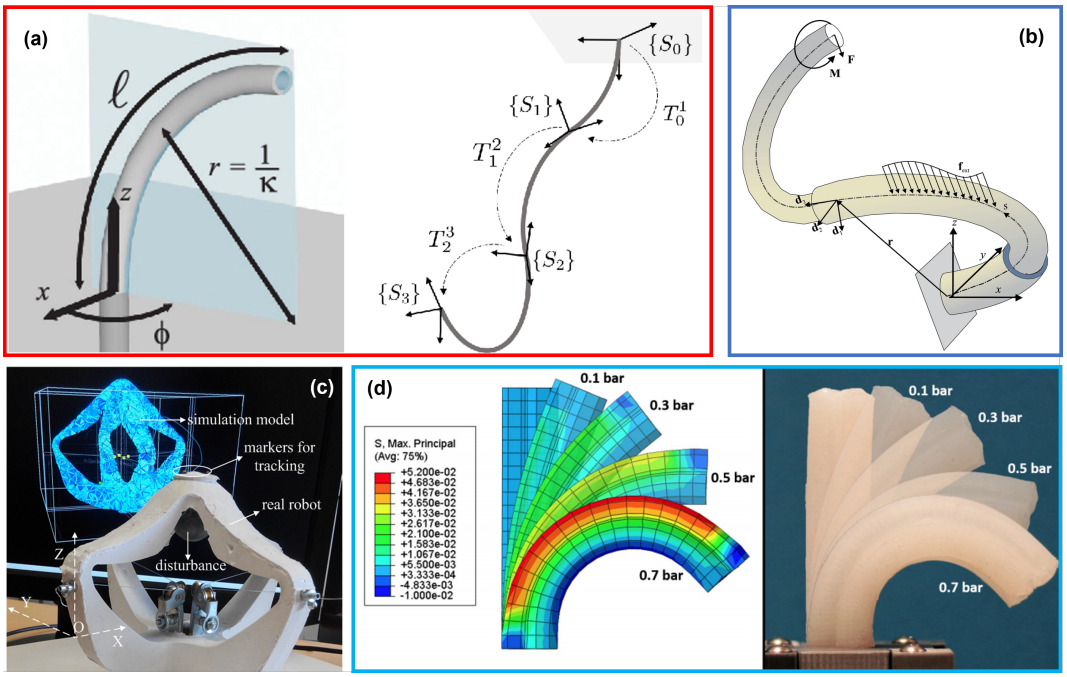


Fig. 2.16: (a) Piecewise constant curvature assumption. **Image source**: [46, 47]. **(b)** Establishing force equilibrium based on Cosserat rod theory. **Image source**: [48]. **c** Real-time finite element modelling for soft robot control. **Image source**: [49]. **(d)** Characterizing soft robot kinematics based on FEM. **Image source**: [50].

the kinematics behavior of limber manipulators without skeletons [156]. The piecewise constant curvature (PCC) assumption (**Fig. 2.16a**) was popularly applied to simplify the bending kinematics of continuum robots with uniform shape and symmetrical actuation [47,157]. Combined with positional sensors, e.g., EM tracking probes [158], the PCC-based geometric solution enabled real-time closed-loop control of the robot pose in free space. Recent work utilized the PCC assumption and a self-contained curvature sensor to control the locomotion of a soft robotic snake [159]. Parallel kinematics was employed to achieve position control of a soft robot integrated with elastomer strain sensors [160]. Other modeling approaches, such as those based on the Cosserat rod theory [161, 162] (**Fig. 2.16b**), have been used to investigate the kinematics mapping by establishing force equilibrium, which can account for gravity and external load. Nevertheless, unknown disturbance to the robot, such as unpredictable payload and interaction with surroundings, can promptly deteriorate the model. Finite element modeling (FEM) (**Fig. 2.16c-d**) was also applied to accurately estimate complex robot deformations, by which the kinematics mapping could be generated and incorporated in the soft robot control [50]. Recent works described that asynchronous FEM could be combined with a quadratic programming algorithm to achieve real-time control of soft robots [163] (**Fig. 2.16c**). However, the modeling accuracy is sensitive to geometric and material parameters, of which the searches

are also heuristic. Moreover, the models mentioned above are design-specific to particular robot structures.

Data-driven control approaches avoid complicated analytical modeling by approximating the kinematics mapping or control policies from acquired sensing data. Neural networks (NNs) have been studied to approximate the global inverse mapping of non-redundant soft continuum robots [164, 165]. NNs could be specifically designed to learn global mapping accurately, but it is not efficient for online learning because all network parameters have to be updated in every iteration. A novel data-based approach [166] was also applied optimal control to estimate the kinematic Jacobian matrix online. It demonstrated stable control of a tendon-driven continuum robot in a 2D statically constrained environment. Recently, a locally weighted online learning controller [167] was proposed to achieve the 3D orientation control of a fluid-driven soft robot. It could compensate for externally applied disturbance; however, the need for heuristic tuning of multiple data-dependent parameters becomes the major weakness of this approach [168].

2.6 Conclusion

This chapter introduces the state-of-art techniques for intra-op MRI-guided intervention, in particular for laser surgery. The requirements on MR safety are clarified, which sets the standard for the development of MRI-guided robotic systems. The current status of non-contact laser surgery in head and neck cancer treatment is also reviewed, including the transoral laser microsurgery, transoral surgical robotic system, and endoscopic laser manipulators. However, there is still clinical demand for efficient intra-op tools, such as MRI, to assess the laser ablation completeness. The current robotic systems for transoral laser surgery are not MRI compatible, while existing MRI-guided robotic systems are not equipped for controlling non-contact, projection-based lasers, particularly within constrained anatomy such as the oral and pharyngeal cavities. To this end, some crucial techniques are investigated, including the MR-compatible actuation, real-time positional localization in MRI, continuum robotic manipulator, and precise motion control for soft continuum robots.

In the following chapters, investigations on the aforementioned important techniques are introduced in detail. In **Chapter 3**, a fluid-driven soft robotic endoscope is presented, with visual servoing controllers proposed to enable its precise navigation control. **Chapter 4**

introduces the high-precision laser beam steering control with a soft robotic manipulator, which serves as the fundamental basis for employing soft robots in transoral laser surgery. Then a complete prototype of an MRI-guided transoral robot integrating hydraulic actuation, soft robotic navigation and laser beam steering, and MR-based positional localization is presented in **Chapter 5**. Finally, **Chapter 6** presents the results of preclinical trials that were conducted to validate the robot's feasibility and performance in actual human anatomy.

Chapter 3

Eye-in-hand Vision-based Online Learning Control for Soft Robotic Navigation

3.1 Introduction

As introduced in **Chapter 2**, soft robots, owing to their elastomeric material, ensure safe interaction with their surroundings. These robot compliance properties inevitably impose a trade-off against precise motion control, as to which conventional model-based methods were proposed to approximate the robot kinematics. However, too many parameters, regarding robot deformation and external disturbance, are difficult to obtain, even if possible, which could be very nonlinear. Sensors self-contained in the robot are required to compensate for modeling uncertainties and external disturbances. Camera (eye) integrated at the robot end-effector (hand) is a common setting. To this end, a learning-based controller is proposed based on the camera feedback to accomplish more precise robotic tasks.

In this chapter, an adaptive eye-in-hand visual servo control framework based on local online learning techniques is presented. The controller is constructed by learning the inverse mapping solely from collected camera images, without the need for calibrating robot and camera parameters. Promising accuracy in the learning of inverse mapping is assured without having to tune the hyperparameters in the learning approach. LGPR models enable fast online updates to accommodate new input data that reflect the latest robot status. As a result, precise manipulation can be achieved even when the robot encounters unknown and

varying external disturbances. The major contributions of this work are:

1. *First* attempt to address a learning-based visual servo control for a fluid-driven soft robot such that the inverse kinematics (IK) can be directly approximated by Gaussian process regression (GPR);
2. Efficient update of the trained inverse motion mapping to compensate *dynamic* disturbance by adjusting the most relevant local GPR model;
3. Novel experimental validations demonstrating precise point-to-point tracking and path following of a hyper-elastic low-stiffness soft robot with *variable* tip loading.

3.2 Eye-in-hand Visual Servoing and Applications

Visual servoing is a control strategy that utilizes visual feedback to close the control loop of robotic systems. It is able to resist distance-related limitations in robotic end-effector control, with applications extending to the manufacturing industry, military field, automobile steering, and even aircraft landing [169]. Visual servoing can be classified into eye-in-hand and eye-to-hand configurations based on the camera's location. Eye-in-hand cameras are embedded at the robot end-effector and enable more flexible viewing of the workspace. Using this approach for rigid-link robots, tele-operated object/feature tracking and obstacle avoidance in imperfectly-modeled environments can be achieved [169,170]. Applications also include robotic minimally invasive surgery, where surgical instruments and specific tissues or organs can be tracked intra-operatively [171], reducing the burden of manual endoscope control.

Vision-based systems are a viable choice for integration with soft robots, as they can be small and self-contained. Making use of camera feedback, visual servoing has been extensively studied over the last decades, and many approaches have been proposed [172,173]. Wang et al. [174] first achieved eye-in-hand visual servo control of a cable-driven soft robot based on analytical kinematic modeling and an interaction matrix, where the intrinsic and extrinsic camera parameters are estimated beforehand. Other studies addressed the visual servo of a concentric-tube robot [175] and series pneumatic artificial muscles [176] by estimating the task space Jacobian matrix from image feedback. But these controllers were only validated in free space. Recently, a PCC-based adaptive visual servo controller was proposed for a cable-driven robot in a constrained environment [177]. It could

handle the control of the robot statically constrained by physical interaction. However, the model-based controller could only adapt to a specific type of constraint, i.e., static "stuck", by estimating the "stuck" length online. Other factors, e.g., modeling uncertainties and varying external forces, could also deteriorate the model and introduce control inaccuracies.

3.3 Online Learning Visual Servoing Control Framework

This section describes the proposed online learning control framework for soft robotic navigation based on camera feedback. Utilizing the learning method, the nonlinear motion mapping from task space to robot actuation can be directly modeled. A feedback controller is constructed based on the trained inverse motion mapping to achieve targeting objectives defined in the task space.

3.3.1 Task Space Definition

A camera mounted at robot end-effector allows image-based control strategy, namely eye-in-hand visual servo. Mappings from the spaces of actuation, configuration to task spaces have to be defined successively. The actuator input (at equilibrium) is represented as $\boldsymbol{\alpha}(k) \in \mathbb{U}^m$ at time step k, where $\mathbb{U}^m$ denotes the m-dimensional actuation space. Let $\boldsymbol{s}(k)$ be the manipulator configuration space parameters under input $\boldsymbol{\alpha}(k)$, which corresponds to an end-effector position $\boldsymbol{p}(k) \in \mathbb{R}^3$ and orientation normal $\boldsymbol{n}(k) \in \mathbb{R}^3$ in the Cartesian space. The orientation normal is a vector of unit length. The collective variable $\boldsymbol{\theta}(k) = [\boldsymbol{p}(k) \;\; \boldsymbol{n}(k)] \in \mathbb{R}^6$ depends on robot configuration $\boldsymbol{s}(k)$:

$$\boldsymbol{\theta}(k) = h(\boldsymbol{s}(k)). \tag{3.1}$$

Assume the robot motion is quasi-static, we can obtain the forward transition model:

$$\Delta\boldsymbol{s}(k) = f(\boldsymbol{s}(k), \Delta\boldsymbol{\alpha}(k)) \;, \tag{3.2}$$

where the incremental input from time step k to $k+1$ is denoted by $\Delta\boldsymbol{\alpha}(k) = \boldsymbol{\alpha}(k+1) - \boldsymbol{\alpha}(k)$, and the change of robot configuration induced by the incremental input $\Delta\boldsymbol{\alpha}(k)$ is denoted by $\Delta\boldsymbol{s}(k) = \boldsymbol{s}(k+1) - \boldsymbol{s}(k)$.

The task space is defined in the camera frame (**Fig. 3.1a**), with the incremental displacement denoted as $\Delta\boldsymbol{z}(k) \in \mathbb{R}^2$. The frame is always perpendicular to the camera

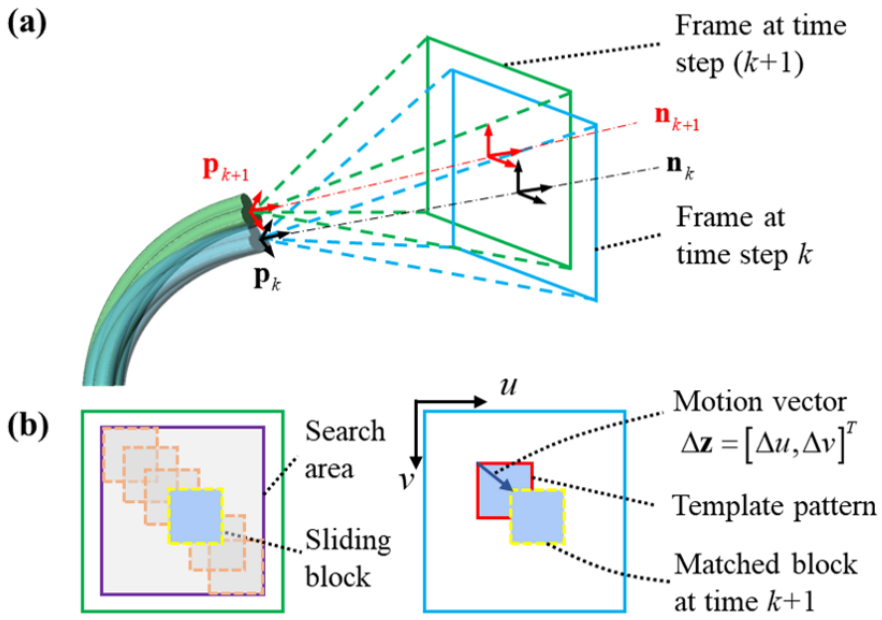


Fig. 3.1: 2D motion estimation in the camera view. **(a)** Image frames at two successive time steps. **(b)** Incremental displacement between two matched blocks with the highest similarity. An red block is selected from the frame at time step k as the template pattern, which is matched by the yellow block searched from the frame at time step $k+1$.

principal axis, which aligns with the robot tip normal. The change of end-effector states $\Delta\boldsymbol{\theta}(\cdot)$ will be mapped to the motion in the camera frame, $\Delta\boldsymbol{z}(k)$. Combining this mapping, **Eq. 3.1** can thereby be extended to:

$$\Delta\boldsymbol{z}(k) = g(\boldsymbol{s}(k), \Delta\boldsymbol{\alpha}(k)), \tag{3.3}$$

The control objective is to generate the actuation command, achieving a desired movement $\Delta\boldsymbol{z}^*(k)$ in task space. Thus, we need to obtain the inverse transition mapping of **Eq. 3.3**, i.e.,

$$\Delta\boldsymbol{\alpha}(k) = \widehat{\Phi}(\boldsymbol{s}(k), \Delta\boldsymbol{z}^*(k)). \tag{3.4}$$

The inverse mapping $\widehat{\Phi}$ is dependent on the current robot states, $\boldsymbol{s}(k)$, which is unknown due to the lack of configuration-related sensing data. However, the robot configuration $\boldsymbol{s}(k)$ is uniquely determined by the actuator input $\boldsymbol{\alpha}(k)$ based on the quasi-static assumption of robot movement. Note that this relation could be altered by external forces. Hence, we can use the actuator input to represent the robot states in the inverse mapping:

$$\Delta\boldsymbol{\alpha}(k) = \Phi(\boldsymbol{\alpha}(k), \Delta\boldsymbol{z}^*(k)), \tag{3.5}$$

which is an approximation of the actual inverse kinematics $\widehat{\Phi}$ in **Eq. 3.4**. The inverse

mapping should be modeled based on the actual operation data comprising image feedback $\Delta \boldsymbol{z}(\cdot)$, and the actuation input $\boldsymbol{\alpha}(\cdot)$, $\Delta \boldsymbol{\alpha}(\cdot)$. Learning-based methods will be employed to avoid complicated analytical modeling.

3.3.2 Motion Estimation on Image Plane

To process the 2D camera image feedback for the learning of inverse transition mapping, the effective estimation of incremental displacement $\Delta \boldsymbol{z}$ between two successive images is first addressed. As shown in **Fig. 3.1b**, a reference block from the image frame at time step k is selected as the template pattern, which is matched by searching from the image frame at time step $(k+1)$ through block sliding. In this calculation, only the 2D movements along the u and v axes are considered. As the time interval between two successive time steps is very short ($\sim$20 Hz), it can be assumed that the change of robot tip orientation is small. Otherwise, the orientation change could introduce inaccuracies to the 2D motion estimation based on template matching. It was evaluated by calculating the movement error of 100 images generated by rotating an original image from -10° to 10° (**Fig. 3.2a**). As the displacement was calculated between successive frames at tiny time intervals, the twisting angle reflected in the camera view is $< 5^\circ$, referring to the recorded twisting angles during pre-training data exploration (**Fig. 3.2b**). Therefore, the movement error caused by the camera rotation would be < 3 pixels, which is acceptable relative to the 10-pixel tolerance. Besides, the orientation-induced error is generally correlated to the robot configuration, which could be accurately resolved by the instant update of the actuation status in the learning model.

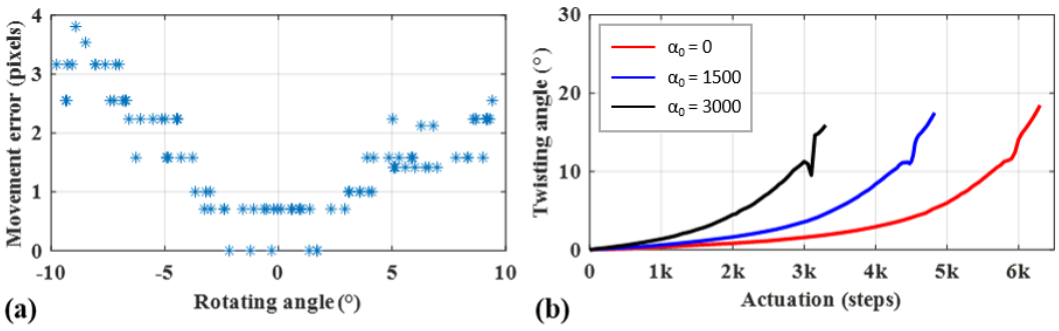


Fig. 3.2: **(a)** Motion estimation errors sampled at angular resolution of 0.25°. **(b)** Twisting angle of robot end-effector (camera) under varying actuation inputs when one chamber is inflated. All three chambers are pre-pressured with actuation input $\boldsymbol{\alpha}_0$.

Therefore, during the continuous robot motion, the camera frames can be assumed to move on the same plane between two successive time steps, thus mainly moving in planar

(2D) motion. A template matching algorithm based on a square of image intensity feature is employed to estimate the 2D translations on the image plane (**Fig. 3.1b**). The algorithm is implemented using the "*matchTemplate*" function in OpenCV [178]. A reference block containing the template pattern is first prescribed in the previous image frame. Then a block is searched from the current image frame by sliding the block pixel-by-pixel and calculating its similarities to the template. A metric function "TM_CCORR_NORMED" is used to quantify the similarity, i.e.,

$$R(\xi,\eta)=\frac{\sum_{i,j}\left(T(i,j)\cdot I(i+\xi,j+\eta)\right)}{\sqrt{\sum_{i,j}T(i,j)^2\cdot\sum_{i,j}I(i+\xi,j+\eta)^2}}, \tag{3.6}$$

where the matrices T and I denote, respectively, intensities in the template block and sliding block. The local indices of pixels in the blocks are denoted by i and j. The location of the sliding block in the image frame is represented by (ξ, η). By finding the block with the highest similarity (**Eq. 3.6**) to the template using "*minMaxLoc*" function, we can obtain a motion vector (**Fig. 3.1**), $\Delta \boldsymbol{z}(k)=[\Delta u,\Delta v]^{\mathrm{T}}$, with respect to the template:

$$[\Delta u,\Delta v]=\arg\max_{\xi,\eta}R(\xi,\eta). \tag{3.7}$$

The forward/backward movement of the camera may induce the scaling effect of the tracked features. However, compared to the lateral camera movement, this is assumed to have minimal effect on the estimation of translation displacements through template matching. Besides, the axial movement of the camera will be small given the short time interval between successive camera imaging ($\geq$ 20 Hz). Similarly, for optical aberration-caused geometrical distortions, its effect on relative motion estimation could be ignored, as the compared templates extracted from successive frames are supposed to be with little changes of distortion. Note that the absolute distortion of the template is not considered in this algorithm. The template matching process is performed continuously alongside the robot movement, thus offering a real-time motion estimation for robot kinematics learning as well as its feedback control.

3.3.3 Local GPR-based Online Learning Control

To "learn" the inverse kinematics, a precise regression approach is required to map the estimated 2D motion to the corresponding actuation input. Given that soft robotic actuation is nonlinear, a nonparametric method, GPR [179], is employed. This data-driven method

can accommodate the unpredictable characteristics caused by the fabrication uncertainty of soft robots and the calibration of the camera intrinsic/extrinsic. After pre-training, the GPR model can be online updated using the new sensing data collected during robot manipulation. A locally weighted learning scheme is utilized to enable fast model prediction and update, thus real-time robot control. There are three key steps as bellow:

Gaussian Process Regression

Training: Our aim is to model the inverse mapping $\Delta\boldsymbol{\alpha} = \Phi(\boldsymbol{\alpha}, \Delta\boldsymbol{z}^*)$ as defined in **Section 3.3.1**, with both the current robot states and desired camera motion, $\boldsymbol{x} = [\ \boldsymbol{\alpha}^{\mathrm{T}} \quad \Delta\boldsymbol{z}^{*\mathrm{T}}\ ]^{\mathrm{T}} \in \mathbb{R}^n$, as input, and the required actuation, $\boldsymbol{y} = \Delta\boldsymbol{\alpha} \in \mathbb{R}^m$, as output. For soft robots, m is the number of actuation chambers. The robot is first commanded to explore the workspace with a predefined actuation sequence, during which a training dataset is collected for model initialization. Given a set of input data $\boldsymbol{X} = \{\boldsymbol{x}_i\}$ and output data $\boldsymbol{Y} = \{\boldsymbol{y}_i\}$, $i = 1, 2, \ldots, N$, each dimension of the output $\boldsymbol{y}^s = \{\boldsymbol{y}_i^s\}, s = 1, 2, ..., m$ is independently trained. GPR assumes that the input and output of the training data satisfies a nonlinear mapping $\boldsymbol{y}_i^s = G(\boldsymbol{x}_i) + \varepsilon$, where ε is a white Gaussian noise with zero mean and variance σ_n^2. The output is modeled as a Gaussian distribution $\boldsymbol{y}^s \sim \mathrm{N}\left(\mathbf{0}, \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_{\mathbf{n}}^{\mathbf{2}}\boldsymbol{I}\right)$, where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X})$ is a covariance matrix. The size of the matrix $\boldsymbol{K}$, $N \times N$, depends on the size of training data, N. Here, the zero-mean prior is adopted, since the change of actuation $\Delta\boldsymbol{\alpha}$ should have zero mean. The i^{th}-row, and j^{th}-column element $k_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ in covariance matrix $\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X})$ is a customized function. Here, squared-exponential kernel function [179] is used:

$$k_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma_s^2 \exp\left(-0.5(\boldsymbol{x}_i - \boldsymbol{x}_j)^T \boldsymbol{\Lambda}(\boldsymbol{x}_i - \boldsymbol{x}_j)\right) \tag{3.8}$$

where ${\sigma_s}^2$ is the signal variance, and $\boldsymbol{\Lambda} = diag(\boldsymbol{\lambda})$ is a diagonal matrix with characteristic length-scales $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]^T$ acting on each dimension of the input $\boldsymbol{X}$ individually. In our experiments, hyperparameters σ_s, σ_n, and $\boldsymbol{\lambda}$ were initialized as 0.1, 0.5, and $0.5\boldsymbol{I}$ respectively. Then they can be updated by maximizing the negative log marginal likelihood. The hyperparameters can be found by standard optimization methods such as conjugate gradient, which is an automatically seeking procedure without the need of heuristic intervention. With these, GPR can generate a global nonlinear mapping model ready for prediction.

Prediction: Given a query input set $\widehat{\boldsymbol{x}}$, the joint distribution of the observed target values $\boldsymbol{y}^s$ and predicted value $g(\widehat{\boldsymbol{x}})$ are expressed as [179]:

$$\begin{bmatrix} \boldsymbol{y}^s \\ g(\widehat{\boldsymbol{x}}) \end{bmatrix} \sim \mathrm{N}\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) + \sigma_n^2 \boldsymbol{I} & \boldsymbol{k}(\boldsymbol{X},\widehat{\boldsymbol{x}}) \\ \boldsymbol{k}(\widehat{\boldsymbol{x}},\boldsymbol{X}) & k(\widehat{\boldsymbol{x}},\widehat{\boldsymbol{x}}) \end{bmatrix}\right) \tag{3.9}$$

Based on **Eq. 3.9**, we can solve the mean $\bar{g}(\widehat{\boldsymbol{x}})$ and covariance $V(\widehat{\boldsymbol{x}})$ of the prediction as:

$$\bar{g}(\widehat{\boldsymbol{x}}) = \boldsymbol{k}^T(\boldsymbol{X},\widehat{\boldsymbol{x}})\left(\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) + \sigma_n^2\boldsymbol{I}\right)^{-1}\boldsymbol{y}^s = \boldsymbol{k}^T(\boldsymbol{X},\widehat{\boldsymbol{x}})\boldsymbol{\beta} \tag{3.10}$$

$$V(\widehat{\boldsymbol{x}}) = k(\widehat{\boldsymbol{x}},\widehat{\boldsymbol{x}}) - \boldsymbol{k}^T(\boldsymbol{X},\widehat{\boldsymbol{x}})\left(\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) + \sigma_n^2\boldsymbol{I}\right)^{-1}\boldsymbol{k}(\boldsymbol{X},\widehat{\boldsymbol{x}}) \tag{3.11}$$

where $\boldsymbol{\beta}$ is the prediction vector with a dimension of $N \times 1$. For the visual servoing task, the predicted value $\bar{g}(\widehat{\boldsymbol{x}})$ and $V(\widehat{\boldsymbol{x}})$ give, respectively, the mean and covariance of the actuation command that is required to achieve the desired camera motion queried in $\widehat{\boldsymbol{x}}$.

Localized GPR Model

During GPR predicting or updating, the inversion of matrix $\left(\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) + \sigma_n^2\boldsymbol{I}\right)$ has to be calculated, which is time-consuming due to the complexity of $O\left(N^3\right)$. Decreasing the size of training dataset N should be an effective way to accelerate the computation. Therefore, instead of grouping the collected training data in one matrix, k-means algorithm is applied to divide them into M clusters, $D_j, j = 1,2,\ldots,M$, with a Gaussian-based kernel (**Eq. 3.8**) utilized to quantify the similarity between each pair of data. For each data, its similarity to the centers of clusters will be calculated, with the closest one selected to include the data. Clusters from 1 to M will be iteratively updated so that the sum of data similarities to their assigned clusters are minimized. As discussed in **Section 3.3.1**, the actuator input $\boldsymbol{\alpha}$ can be used to represent the robot state, which is thus selected as the basis for clustering. The sample size of the j^{th} cluster is denoted by N_j, with the corresponding center represented as $\boldsymbol{c_j} \in \mathbb{U}^m$. A local GPR model Φ_j is trained for each cluster, with a upper limit of sample size $N_j^{\max}$ imposed to ensure efficient computation. The limit is also maintained during model updating with newly collected data. At each time step, **Eq. 3.10** can be used to calculate a prediction $\widehat{\boldsymbol{y}}_j$ from each local model Φ_j. Taking the weighted average of the predictions of M local models, the combined actuation input [180] can be obtained as:

$$\Delta\boldsymbol{\alpha}(k) = \textstyle\sum_{j=1}^{M} \omega_j \widehat{\boldsymbol{y}}_j \Big/ \sum_{j=1}^{M} \omega_j, j = 1,2,\ldots,M, \tag{3.12}$$

where the weight ω_j is given by the similarity between current actuation input $\boldsymbol{\alpha}(k)$ and the center of j^{th} cluster. The similarity is measured using a Gaussian kernel (**Eq. 3.8**), assuming each cluster of data satisfies a Gaussian distribution. The steps of data processing

and training for model initialization are summarized in **Algorithm. 1**.

Algorithm 1 Pre-training of LGPR models

Inputs:

$\boldsymbol{X}$: input samples;
$\boldsymbol{Y}$: output samples;
$\omega(\cdot)$: similarity kernel function;

1: Clustering training dataset into M local groups with k-means clustering algorithm.
2: **for** each cluster $j = 1, 2 \ldots, M$ **do**
3: Train j^{th} Local GPR model **Eq. 3.8**;
4: **end for**

Incremental Learning

During robot execution, its actuation property may be changed by external disturbances, which result in observations different from the training data collected in free space. To enhance the adaptability of the controller to those uncertainties, online refining of the pre-trained models is implemented. At each time step k, new sample data pairing the input $\boldsymbol{x} = [\ \boldsymbol{\alpha}(k)^{\mathrm{T}} \quad \Delta\boldsymbol{z}(k)^{\mathrm{T}}\ ]^{\mathrm{T}} \in \mathbb{R}^n$ and output $\boldsymbol{y} = \Delta\boldsymbol{\alpha}(k)$ can be acquired and grouped into the nearest cluster D_r for model updating. As the online data can reflect the latest robot operation condition, the updated model will adapt to unknown external disturbances, thus providing more accurate control. During the dataset updating, the size limits of each cluster are imposed to ensure fast model prediction. If the current size of local model $N_r < N_r^{\max}$, the new sample can be directly added into the cluster with other data kept unchanged; otherwise $N_r \geq N_r^{\max}$, the oldest sample $[\boldsymbol{x}_1, \boldsymbol{y}_1]$ will be deleted while adding the new sample. A new local model Φ_r will be trained based on the updated dataset.

The key steps of model updating is to calculate the new prediction vector $\boldsymbol{\beta}$ as in **Eq. 3.10**, which involves the inversion of matrix $\left(\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma^2 \boldsymbol{I}\right)$. To enable efficient computation of this inversion, Cholesky decomposition $\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$ is employed to represent the matrix. Then the prediction vector can be solved from the linear equations, $\boldsymbol{y} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\beta}$. The advantage is that the Cholesky decomposition can be directly adjusted to accommodate the newly added data [180], by inserting a new row in matrix $\boldsymbol{L}$:

$$\boldsymbol{L}_{new} = \begin{bmatrix} \boldsymbol{L} & \boldsymbol{0} \\ 1^{\mathrm{T}} & l_* \end{bmatrix}, \tag{3.13}$$

$$1 = \boldsymbol{L}^{-1}\boldsymbol{k}(\boldsymbol{X}, \boldsymbol{x}_{new}), \quad l_* = \sqrt{k(\boldsymbol{x}_{new}, \boldsymbol{x}_{new}) - \|1\|^2}. \tag{3.14}$$

where $k(\cdot)$ is the kernel function as given in **Eq. 3.8**.

Discarding the oldest sample from matrix $\boldsymbol{L}$ can be achieved by two steps: The first one is to re-arrange the oldest data to the last row. This can be achieved by multiplying $\boldsymbol{L}$ with a permutation matrix:

$$\boldsymbol{R} = \boldsymbol{I} - (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_{N_r})(\boldsymbol{\delta}_1 - \boldsymbol{\delta}_{N_r})^{\mathrm{T}}, \tag{3.15}$$

i.e., $\boldsymbol{RL}$. In **Eq. 3.15**, a vector $\boldsymbol{\delta}_i \in \mathbb{R}^{N_r}$ contains all zero elements except for the i^{th} element that is equal to one. We can then obtain the new matrix $\boldsymbol{L}_{new}$ by deleting the last row of the re-arranged matrix, $\boldsymbol{RL}$. The procedures for updating the pre-trained models incrementally are given in **Algorithm. 2**. **Fig. 3.3** presents the online learning control framework integrating four key units: command input, LGPR model prediction and updating, soft robot actuation, and image processing.

Algorithm 2 Incremental learning of LGPR Models

for each new sample $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ **do**
  for each LGPR model $\Phi_j, j = 1, 2 \ldots, M$ **do**
    Calculate the similarity between the sample and cluster center $\omega_j = \omega(\boldsymbol{x}_i, \boldsymbol{c}_j)$;
  end for
  Select the nearest model: $r = \arg\max_j \omega_j$;
  if $\omega_r >$ similarity threshold **then**
    if $N_r > N_r^{\max}$ **then**
      Remove the oldest data in model Φ_r;
    end if
    Add $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ into the most relevant local cluster D_r: $\boldsymbol{X}_r = \boldsymbol{X}_r \cup \boldsymbol{x}_i$, $\boldsymbol{Y}_r = \boldsymbol{Y}_r \cup \boldsymbol{y}_i$;
    Update the center of local cluster: $\boldsymbol{c}_r = \mathrm{mean}(\boldsymbol{X}_r)$;
    Calculate the new Cholesky matrix and prediction vector using **Eq. 3.13**;
  else
    Generate a new cluster $\boldsymbol{c}_{M+1} = \boldsymbol{x}_i, \boldsymbol{X}_{M+1} = \boldsymbol{x}_i, \boldsymbol{Y}_{M+1} = \boldsymbol{y}_i, M = M + 1$;
    Train a new local GPR model;
  end if
end for

3.4 Experiments of Image-based Visual Servoing

In this section, experimental validation of the online learning visual servoing controller is conducted using a pneumatic-driven hyper-elastic soft robot. Point-to-point tracking and path following tasks are performed to showcase the robot control accuracy. During these tasks, external forces and changing tip loading were intentionally applied on the robot to validate the adaptability of the proposed online learning controller.

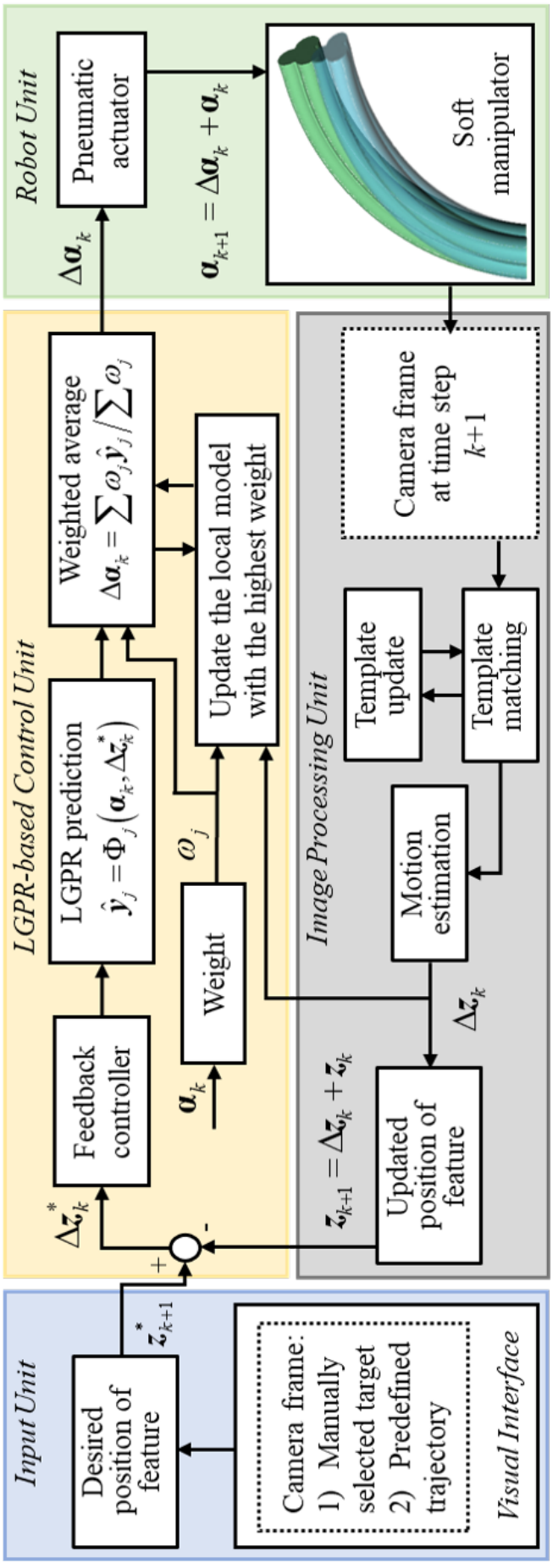


Fig. 3.3: Framework of the proposed online learning control method. Variables $\boldsymbol{\alpha}$ and $\boldsymbol{z}$ represent the actuation input and the position of tracked feature, respectively. The input unit defines the control task in the camera view, either tracking a manually selected target or a predefined trajectory. The LGPR-based control unit calculates the actuation command $\Delta\boldsymbol{\alpha}_k$, given the desired movements $\Delta\boldsymbol{z}_k^*$ and current actuator input $\boldsymbol{\alpha}_k$. 2D camera motion $\Delta\boldsymbol{z}_k$ is online estimated and fedback to update the LGPR models and close the control loop.

3.4.1 Experimental Setup

The experimental setup is shown in **Fig. 3.4**, where a soft robotic manipulator is fixed at the base, with its tip facing towards a LEGO® scene. The robot body was molded with room-temperature-vulcanization (RTV) silicone (Ecoflex 0050; Smooth-On, Inc.), which features low stiffness and high elasticity [167]. With a compact size in an outer diameter of $\varnothing$13 mm and a length of 67 mm, the soft robot can be used for endoscopic application. The robot integrates three individual cylindrical chambers that can be inflated by pneumatics or hydraulics. On the outer surface of each chamber, a layer of Kevlar string (1-mm pitch) was helically wound to limit its radial deformation. Thus, the chambers will only elongate or shorten axially, providing bending motion with an angle of $> 90^\circ$. In the experiments, the inflation of each chamber is driven by a pneumatic cylinder, which is coupled to a stepper motor through lead screw transmission. By adjusting the inflation volumes of three chambers, omni-directional aiming of the robot tip can be achieved. The robot tip is mounted with an endoscopic camera (Depth of view: 8-150 mm). Images with 400×400 pixels can be captured for processing. Given that the diagonal field of view is 90°, an image pixel can be translated to 0.16° camera motion. An LED module is also equipped for adjusting the brightness in the camera view when the operating environment is dark, since this lighting will move along with the robot end-effector and function the same role in each frame. All LEDs on the module should be in the same white color and brightness, thus avoiding additional disturbances to the original features.

Upon pneumatic actuation, the soft robot made of hyper-elastic material will present a nonlinear bending response, which is difficult to be accurately characterized for precise robot control. Alongside bending, inflating chambers would also generate twisting of the continuum robot body, which gives rise to the unexpected rotation of the camera view.

3.4.2 Pre-training Inverse Kinematics using Local GPR

The robot is first commanded to explore its bending workspace to collect a pre-training dataset $\boldsymbol{D}' = \{\boldsymbol{\alpha}_i, \Delta\boldsymbol{z}_i, \Delta\boldsymbol{\alpha}_i\}|_{i=1}^{N}$ for initialization of the LGPR models. The robot working scene used for pre-training can be different from the actual one. The actuation input $\boldsymbol{\alpha}_i$ satisfies a uniform distribution and was executed sequentially. At each actuation step, after reaching equilibrium, a camera image is captured, with the 3D position of the robot tip recorded by an EM tracking coil (NDI Aurora®).

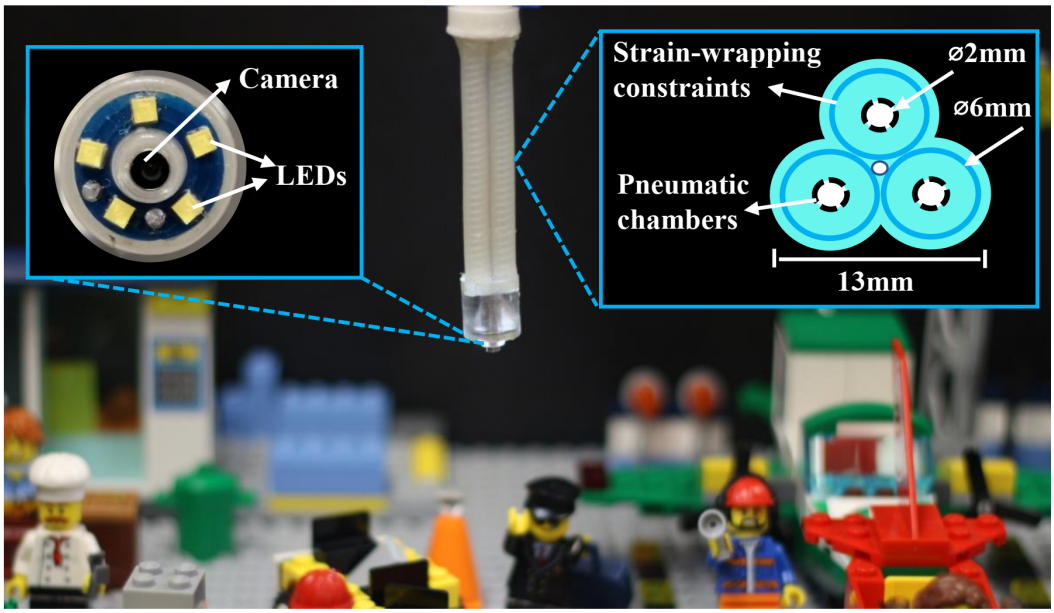


Fig. 3.4: Soft robot manipulated in a LEGO® scene. The soft manipulator is fabricated using silicone rubber, comprising three string-constrained pneumatic chambers for actuation. The robot tip is equipped with an endoscopic camera and five LEDs.

The next step is to pair the collected data. Between two actuation inputs $\boldsymbol{\alpha}_j$ and $\boldsymbol{\alpha}_i$, we can obtain the change of actuation as $\Delta\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_j - \boldsymbol{\alpha}_i$. To calculate the corresponding 2D motion $\Delta\boldsymbol{z}_i$ in camera view, a visual feature is selected at the center ($\boldsymbol{z}_c$) of the image captured under input $\boldsymbol{\alpha}_i$, and matched from the new image ($\boldsymbol{z}_{new}$) captured under input $\boldsymbol{\alpha}_j$. The motion vector $\Delta\boldsymbol{z}_i$ can be obtained by subtracting the positions of feature in two image frames, $\Delta\boldsymbol{z}_i = \boldsymbol{z}_{new} - \boldsymbol{z}_c$.

Note that the dimension of actuation space ($\Delta\boldsymbol{\alpha}_i$) is larger than the task space ($\Delta\boldsymbol{z}_i$). Such redundant actuation introduces nonconvexity to the inverse mapping, $\Delta\boldsymbol{\alpha}_i = \Phi(\boldsymbol{\alpha}_i, \Delta\boldsymbol{z}_i^*)$, in which a data pair $(\boldsymbol{\alpha}_i, \Delta\boldsymbol{z}_i)$ may map to more than one $\Delta\boldsymbol{\alpha}_i$. To learn a model with high accuracy and generality, the nonconvex dataset $\boldsymbol{D}'$ is pre-processed by removing the data samples $(\boldsymbol{\alpha}_i, \Delta\boldsymbol{z}_i, \Delta\boldsymbol{\alpha}_i)$ in a condition that the elements in $\boldsymbol{\alpha}_i = [\alpha_{1i} \;\; \alpha_{2i} \;\; \alpha_{3i}]^{\mathrm{T}}$ are all nonzero. Thus, we obtain a new dataset:

$$\boldsymbol{D} = \{\boldsymbol{\alpha}_i, \Delta\boldsymbol{z}_i, \Delta\boldsymbol{\alpha}_i \mid \alpha_{1i} \cdot \alpha_{2i} \cdot \alpha_{3i} = 0, \forall i = 1, \ldots, N\}. \tag{3.16}$$

Actuation input fulfilling the constraints in **Eq. 3.16** comprises at least one zero elements, which means at least one chamber is not inflated. Thus, every actuation command will lead to a unique robot pose in the workspace, indicating that the actuation space in dataset D is non-redundant. However, the actuation command predicted by the learned model may violate the constraint due to approximation error. To maintain the convexity of the training

dataset, the constraint in **Eq. 3.16** is also applied for new data pairs that were used to update the controller.

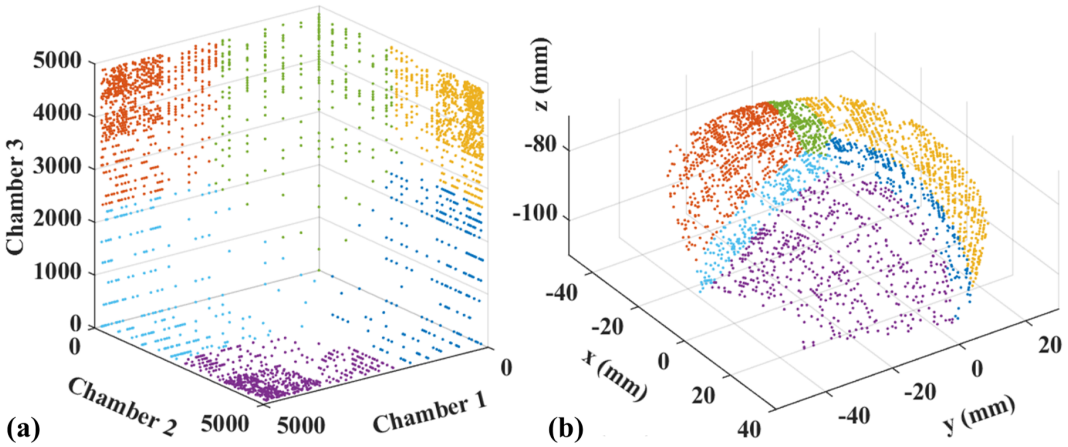


Fig. 3.5: Three thousand sample pairs of robot actuation and tip position collected for initialization of the inverse model. **(a)** Training set partitioned into six (colored) clusters based on their actuation inputs with the k-means algorithm. **(b)** Corresponding tip position distribution of six clusters is also shown.

To train multiple local GPR models that characterize the whole workspace, the training dataset $\boldsymbol{D}$ (randomly selected 1,000 samples) was partitioned into 6 groups using the k-means clustering algorithm. A Gaussian kernel was utilized to measure the distance between points. **Fig. 3.5** shows the 6 local clusters in 6 colors, with a data size of $[202, 257, 82, 294, 106, 59]$ respectively. The ribbon-like clusters have one soft chamber primarily actuated, while the other three clusters have two soft chambers actuated. The data sizes of all local models are limited to be <300.

3.4.3 Experiments and Results

With the experimental platform, two kinds of tasks are conducted to validate the proposed online learning controller. The first one is to track a manually selected target by locating it at the image plane center. The tracking error is defined as the distance between the current position of the target and the camera view center. The other task is to track the target along a predefined trajectory in the image plane, with the corresponding tracking error measured as the shortest distance between the current target position and the trajectory. Both tasks are performed in free space and under external disturbances, such as interactions and tip loading. Four experiments have been conducted as follows:

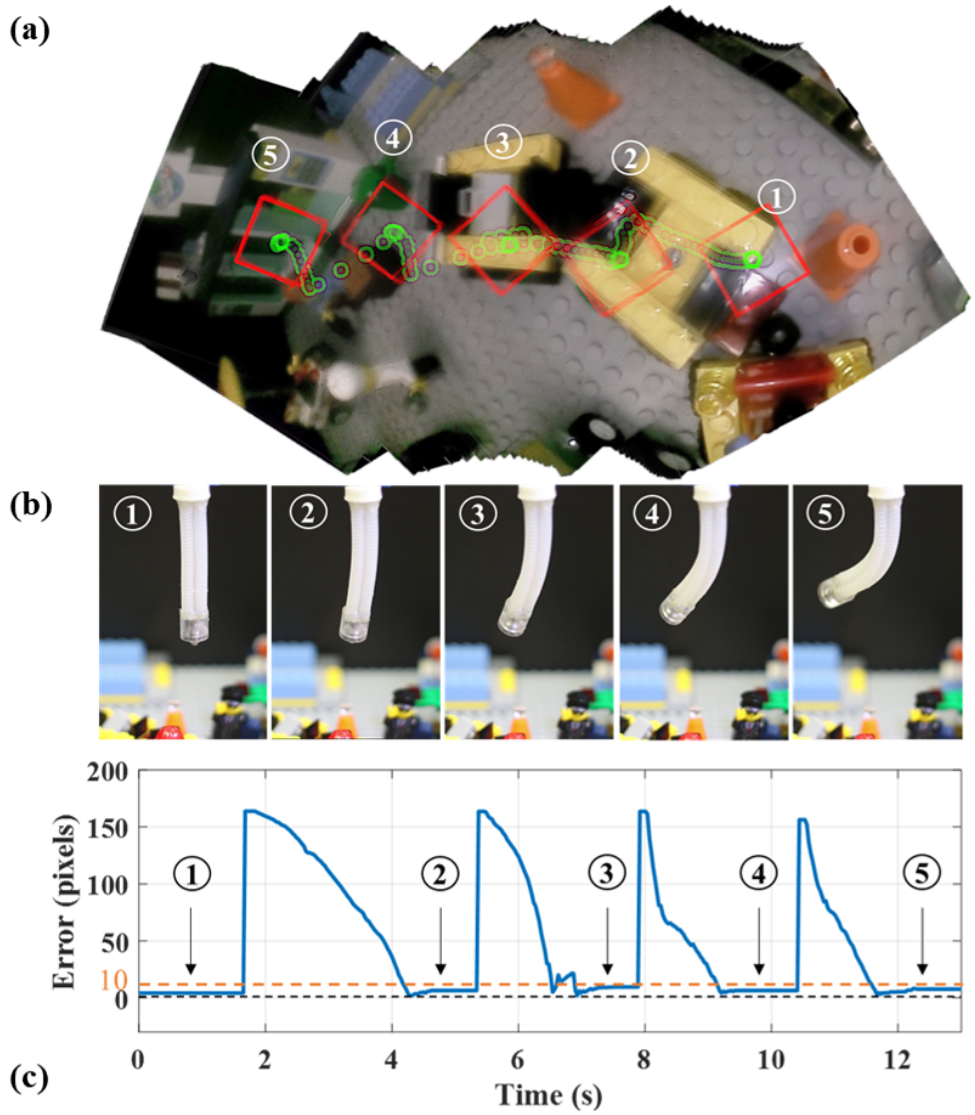


Fig. 3.6: Tracking of five target points, which are the center of red block, manually selected in the camera view. **(a)** Panorama image obtained by stitching the image sequences. The red box represents the template pattern centered at each selected target. The green circles denote the camera centers at each step, showing the footprints of matched block feature throughout the journey. **(b)** Corresponding robot configurations with the five targets traced at image center. **(c)** Tracking errors in unit of pixels.

Point-to-point Tracking

The aim of this task is to position any target way-points at the image center by manipulating the robot. This function could be useful in endoscopic navigation. As shown in **Fig. 3.6**, five targets (labels 1 to 5) are picked and tracked sequentially. Once a target point is selected, a 100×100 block (red box) is created with the target as its center. The block of image intensities serves as the template pattern to be matched in the image processing unit. By stitching the series of camera images throughout the tracking journey, a panorama image can be obtained, as shown in **Fig. 3.6a**, with the green circles denoting the image centers.

Fig. 3.6b shows the five snapshots of robot configurations captured with five targets tracked at the image center, respectively. A tolerance of 10-pixels was set for the tracking accuracy to ensure fast convergence. As depicted in **Fig. 3.6c**, the five target points are accurately aimed with an error $<$10 pixels, demonstrating that the proposed controller enables accurate soft robot control. It proves that LGPR models can provide a precise approximation of the robot inverse kinematics, compensating the disorientation between actuation and task space.

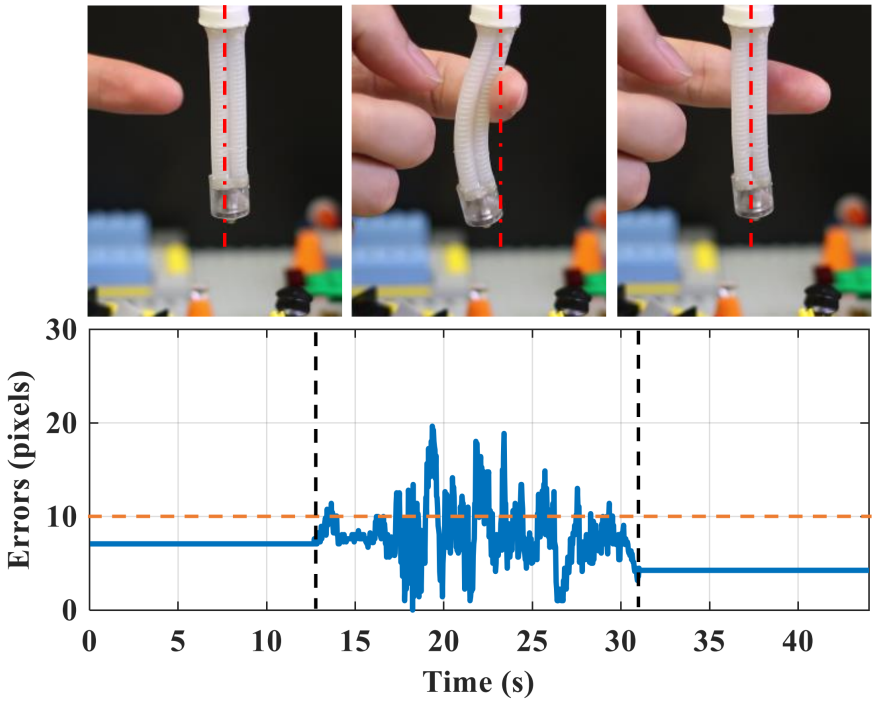
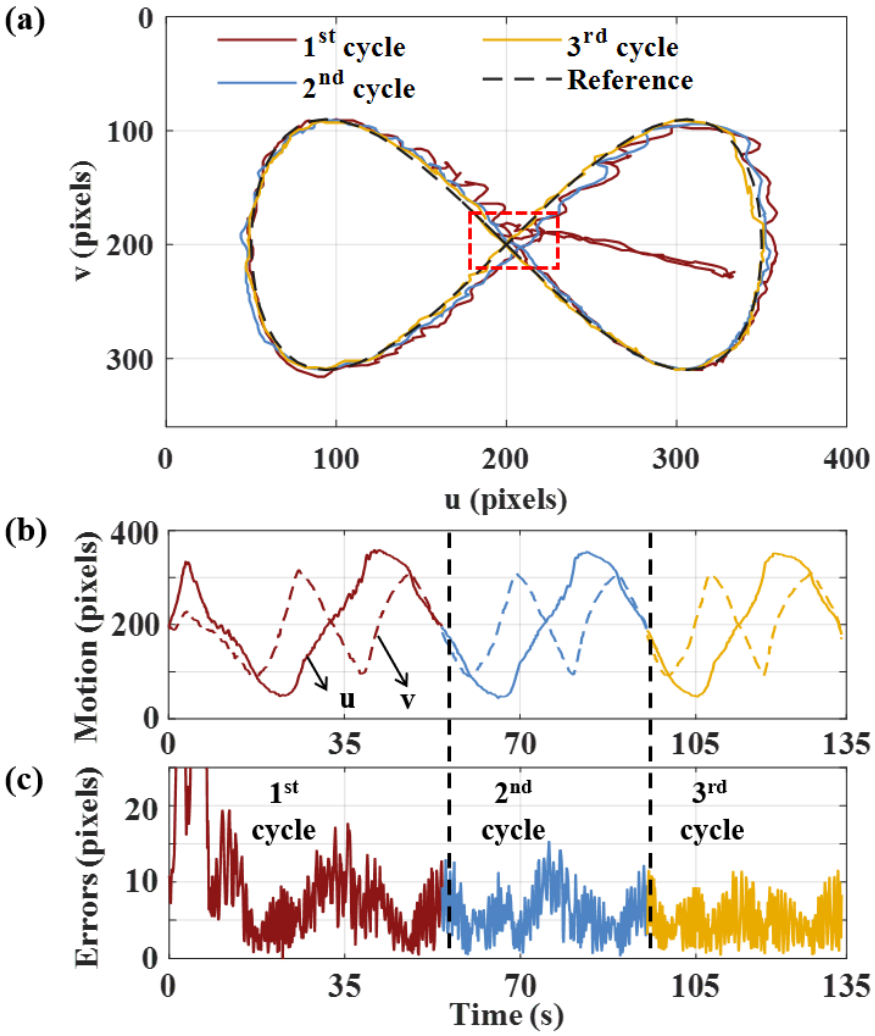


Fig. 3.7: Target tracking with external forces applied. The robot is pulled away from its initial straight configuration and then released. The bottom figure shows the tracking errors at three phases of loading. The orange dash line represents the 10-pixels tolerance.

Target Tracking under External Forces

The target tracking task has been further conducted with external disturbances applied on the robot body, for the purpose of validating the feedback controller's robustness. A target point was initially selected at the image center for robot tracking. External forces were intentionally introduced to deviate the robot from its initial configuration, thus causing tracking errors. The controller was commanded to minimize the error, thus fixating the target at the image center. In this task, online model updating is not involved, with the camera image only used for feedback control. **Fig. 3.7** shows the robot configurations and the tracking errors, which are maintained as less than 20 pixels, even with the robot affected by external forces. The force was applied continuously by a finger within a certain time window (12.7$\sim$31.0 sec). The results demonstrate that the robot could react against the

external loading with feedback control. Although the effective compliance of the robot is reduced, the inherent mechanical limitation of its soft body can still prevent excessive force output.



(d)

Cycle No.	1	2	3
RMSE (pixels)	16.8	6.6	5.4
Max. error (pixels)	64.5	15.3	11.5

Fig. 3.8: Path following test on a predefined "∞" trajectory. **(a)** Trajectory of the tracked targets in the camera view. A target (red-dashed block) is initially selected at the intersection of the "∞" trajectory for robot tracking. **(b)** u and v coordinates of the tracked block throughout the journey. **(c)** Tracking errors in three cycles. **(d)** Summary of the error statistics.

Path Following with Insufficient Pre-training

To evaluate the efficiency of the proposed online learning scheme, a path following task has been performed, with the LGPR model initialized from a small dataset of 300 samples. Due to the insufficient pre-training data, the obtained model may be inaccurate at some

places in the workspace. Therefore, the LGPR model must be online refined with the newly collected data during execution. The reference path is given as a sequence of positions in the 400×400 pixels camera view. **Fig. 3.8a** presents the tracked trajectory. **Fig. 3.8b** shows the coordinates of tracked targets along u and v axes in the image frame. The corresponding tracking errors in pixels are depicted in **Fig. 3.8c**. In the 1^{st} cycle, only a rough following of the reference path is achieved, with a root-mean-square error (RMSE) of 16.8 pixels and a maximum error of 64.5 pixels (**Fig. 3.8d**). In the following two cycles, with the updating of the controller, tracking accuracy is enhanced, with the RMSE and maximum error reduced to 5.4 pixels and 11.5 pixels, respectively. Although the controller was initialized with a poorly pre-trained model, the proposed online learning scheme enabled efficient refinement of the controller, thus achieving precise path following.

Path Following under Varying Load

A path following task has been further conducted with dynamic loading applied on the robot tip. Pre-training data was collected with the robot vertically positioned, and there was no additional weight attached to the robot. However, in the path following task, a variable payload was introduced by wrapping a balloon at the robot tip. The balloon can be inflated by pumping in water through a silicone tube. An additional weight of up to 21 grams can be applied, including 15 grams of water and 6 grams of the balloon cap. Note that the maximum load (21 g) reaches 105% of the robot original weight (20 g). In addition, the robot was placed at a horizontal pose, thus maximizing the effects of both the tip load and robot gravity, which has to be compensated by the proposed online learning controller.

Fig. 3.9a and **Fig. 3.9b** depict the tracking trajectory and corresponding tracking errors, respectively, with the results in three cycles in different colors. In the 1^{st} cycle, the balloon was not inflated, but the 6 grams balloon cap and robot gravity still could adversely affect the tracking performance. An error peak was observed at time $t = 22$ s. However, it was eliminated in the following cycles, demonstrating that the controller adapted to the additional tip load. In the 2^{nd} cycle, 12.5 grams of water (62.5% of robot weight) was quickly pumped into the balloon at $t = 89$ s, finished at $t = 90.5$ s, which gives rise to an error peak in one second (**Fig. 3.9c**). Snapshots of the robot configuration and the inflated balloon are presented **Fig. 3.9d**. The controller achieved fast convergence of the tracking error during $t = 90.5 \sim 93.5$ s. More water was injected into the balloon, resulting in a tip load up to 15 grams at $t = 94.8$ s (**Fig. 3.9d**, right). The additional load was increased at a relatively low speed, which can be compensated by the online learning controller, thus maintaining the tracking error as $<$12 pixels. At $t = 98$ s, the water in the balloon

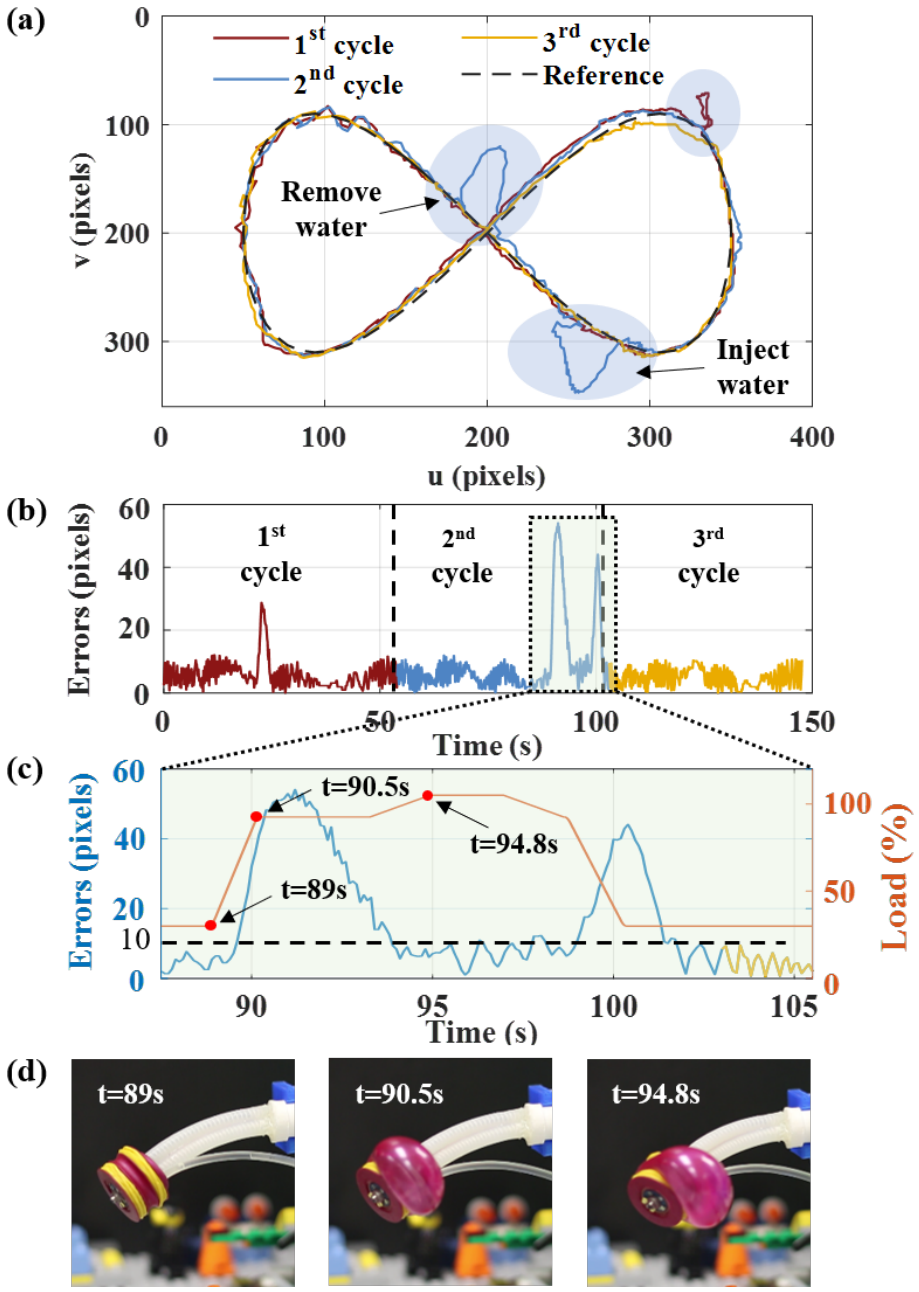


Fig. 3.9: Performance of path following under varying payload. A (6-gram) balloon is wrapped around the robot tip. Water is pumped in-and-out of the balloon to introduce a changeable tip load (6∼21 g). **(a)** Tracked trajectory in three successive cycles. Deviations in regard to the pre-training error (1st cycle in red), injecting and removing water (2nd cycle in blue) were observed. **(b)** Corresponding tracking errors throughout the journey. **(c)** Tracking error *vs* additional payload. The load is presented in % with respect to the robot original mass, 20 g. **(d)** Snapshots of the robot and balloon at three time steps.

was removed at a fast speed. The tracking trajectory was deviated from the reference path due to the sudden change of tip load. But the controller quickly diminished the deviation, converging the tracking error to a small level in a 1-second period. In the 3^{rd} cycle, precise path following was maintained with the tracking errors $<$12 pixels. From this experiment, it

can be analyzed that the closed-loop online learning controller can guarantee a stable path following under a variation (increment or decrement) of external loading within a 9.6%/s changing rate, where the percentage is referred to the added loading in the robot original weight. When the changing rate is in the range of 9.6∼42.7%/s, the robot may not keep converging to the path. However, once the rate reduces less than 9.6%/s, the robot tracking error will converge to a steady range again, with an error reducing rate of ∼14 pixel/s. The proposed controller demonstrates accurate path following even under varying tip load, indicating that the pre-trained LGPR models can be effectively updated online.

3.5 FBG-enhanced Soft Robotic Visual Servoing

In the previous validation of soft robotic visual servoing (**Section 3.3**), 2D motion estimation relying only on image plane feedback is easily affected by vision occlusion, blurring, or poor lighting. For the commonly-used template matching method, tracking performance greatly depends on the image quality. Fiber Bragg gratings (FBGs), a type of high-frequency flexible strain sensor, can be used as an assistant device for soft robot control. A method is proposed to enhance motion estimation in soft robotic visual servoing by fusing the results from template matching and FBG wavelength shifts to achieve more accurate tracking in applications such as minimally invasive surgery. Path following performance is validated in a simulated laparoscopic scene, demonstrating significant improvement to feature tracking and robot motion.

3.5.1 Motion Estimation Combined with FBG

Template matching, as illustrated in **Section 3.3.2**, is a commonly-used method for camera motion estimation. However, features in the surrounding scene and the lighting condition would greatly affect the estimation accuracy. These inaccuracies introduced by the image quality can not be addressed by solely improving the image processing method. FBGs can provide measurement on strains reflecting the robot configuration, which can be combined with the camera feedback to enable a more robust estimation of camera motion. Two major steps are introduced as bellow:

Learning-based motion estimation with FBG

Optical fiber containing FBGs has been attached along the robot body to reflect its deformation. To ensure effective sensing, we should pay attention to several considerations

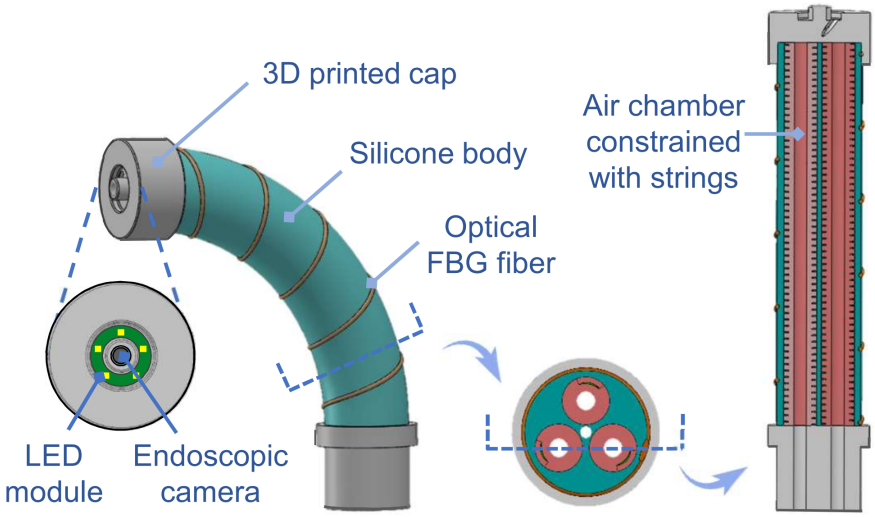


Fig. 3.10: Pneumatic-driven 3-chamber robot is used. An optical fiber with multiplexing 16 FBGs is helically wrapped around the manipulator for sensing feedback of robot configurations in real-time. A monocular endoscopic camera and a LED module are fixed on the tip cap of robot. Cross-section and axial views of the robot show the silicone chambers constrained individually with helical Kevlar strings.

when placing the fiber: 1) Locate the gratings at places involving large strain changes; 2) Distribute the FBGs so that the measured strain changes altogether can uniquely map to the robot configuration; 3) The effect of optical fiber on the mechanical properties of the soft robot should be minimized; 4) Contact on the robot surface should introduce negligible or small strain changes in FBGs. These considerations inspire us to wrap the fiber helically along the robot body and seal it with silicone adhesive. Using the L FBGs integrated in the optical fiber, wavelength shifts from 1 to L can be obtained and represented by $\boldsymbol{\lambda} = [\lambda_1 \lambda_2 \cdots \lambda_L]^T \in \mathbb{R}^L$. Wavelength shift $\boldsymbol{\lambda}(k)$ at time step k is the difference between current wavelength vector and the original wavelength vector $\boldsymbol{\lambda}_0$ (corresponds to initial robot configuration) (**Fig. 3.10**). For each two adjacent time steps k and $(k+1)$, we could obtain the 2D motion vectors in the camera frame $\Delta \boldsymbol{z}(k+1) \in \mathbb{R}^2$ (**Section 3.3.1**), and the wavelength shift $\boldsymbol{\lambda}(k)$ and $\boldsymbol{\lambda}(k+1)$. Their wavelength difference is represented as $\Delta\boldsymbol{\lambda}(k+1) = \boldsymbol{\lambda}(k+1) - \boldsymbol{\lambda}(k)$. In the camera view with abundant features, the estimation results from template matching during slow and smooth manipulation could be regarded as a ground truth, which is used to train the FBG-related estimation model.

Training: The pre-training procedure collects the captured endoscopic image at each time step, as well as the wavelength shift vector $\boldsymbol{\lambda}$. The actuation sequence $Q = [\ \boldsymbol{\alpha}(1) \quad \boldsymbol{\alpha}(2) \quad \cdots \quad \boldsymbol{\alpha}(N)\ ]$ (N is the sampling number) to the robot for data exploration is predefined to cover the whole workspace. The wavelength difference sequence can therefore

be obtained as

$$\Delta\boldsymbol{\Lambda} = \begin{bmatrix} \Delta\boldsymbol{\lambda}(1) & \Delta\boldsymbol{\lambda}(2) & \cdots & \Delta\boldsymbol{\lambda}(N) \end{bmatrix} \tag{3.17}$$

where $\Delta\boldsymbol{\lambda}(i) = \boldsymbol{\lambda}(i) - \boldsymbol{\lambda}(i-1)$, $i = 1,2,\ldots,N$. The images captured on each step are saved to calculate the 2D motion $\Delta\boldsymbol{z}(i) = \boldsymbol{z}(i) - \boldsymbol{z}(i-1)$ offline, forming the motion sequence

$$\Delta\boldsymbol{Z} = \begin{bmatrix} \Delta\boldsymbol{z}(1) & \Delta\boldsymbol{z}(2) & \cdots & \Delta\boldsymbol{z}(N) \end{bmatrix} \tag{3.18}$$

Using the feedforward neural network in the deep learning toolbox of MATLAB, with $\begin{bmatrix} \Delta\boldsymbol{\Lambda} & \boldsymbol{\Lambda} \end{bmatrix}^T$ as input and $\Delta\boldsymbol{Z}$ as output, we can train a motion estimation model represented as

$$\Delta\boldsymbol{z}(i) = M(\Delta\boldsymbol{\lambda}(i), \boldsymbol{\lambda}(i)), \ i = 1,2,\ldots,N. \tag{3.19}$$

Prediction: To increase the processing speed, the model generated in MATLAB is converted to a dynamic link library (DLL) that can be called in the C++ environment. The wavelength receiving via User Datagram Protocol (UDP) and the motion estimation are both accomplished in Qt Creator, with the motion estimation at the k^{th} step obtained by

$$\begin{aligned} \Delta\boldsymbol{z}_w(k) &= M(\Delta\boldsymbol{\lambda}(k), \boldsymbol{\lambda}(k)) \\ &= M((\boldsymbol{\lambda}(k) - \boldsymbol{\lambda}(k-1)), \boldsymbol{\lambda}(k)), \ k = 1,2,\ldots \end{aligned} \tag{3.20}$$

As this model is not related to real-time image processing once finished the training procedure, it can be regarded as an independent 2D motion estimator using FBG with a frequency of at least 30 Hz.

Visual-strain-combined Motion Estimation

The template matching method (**Section 3.3.2**) is utilized to estimate the motion on the image plane. During the searching of the sliding block that matches the template pattern, the corresponding similarity is calculated using **Eq. 3.6** and denoted by a variable γ. The matched block has the largest similarity to the template, with their center positions in camera view denoted by $\boldsymbol{z}(k)$ and $\boldsymbol{z}(k-1)$, respectively. By subtracting the positions, the 2D motion $\Delta\boldsymbol{z}_c(k)$ is obtained, with the corresponding coherence $\gamma(k)$. Based on the trained model from FBGs sensing (**Eq. 3.20**), another motion can be predicted as $\Delta\boldsymbol{z}_w(k)$. The accuracy of image-based estimation is affected by the quality of image features, while the FBGs sensing may be noised by temperature variation and other small interactions. To achieve a stable motion estimation but avoiding both their disadvantages, the combined

motion can be calculated as a weighted average:

$$\Delta z_p(k) = w\gamma(k) \cdot \Delta z_c(k) + (1 - w\gamma(k)) \cdot \Delta z_w(k) \tag{3.21}$$

where the constant w is used to adjust the ratio of coherence, functioning as a scaling factor that normalized the two motion estimations. It is predefined and tuned to the best value before robot execution. In the experiment, it was initialized as 0.7. Note that high tracking accuracy can be maintained with the parameter varied in a small range ($0.6 \sim 0.8$).

3.5.2 Experiments of FBG-enhanced Visual Servoing

Experiments have been conducted to validate the FBG-enhanced visual servoing in scenarios with insufficient features, by comparing with the pure camera-based method [181]. A single-core optical fiber with 16 FBGs (8 mm length for each FBG) was helically wrapped and adhered (by Sil-Poxy^{TM}, Smooth-on Inc.) along the robot body (**Fig. 3.10**). The bottom of the robot has been fixed to a 3D printed base that remains stationary. The experiment setup and results are given as follows:

Experimental Setup

As illustrated in **Fig. 3.11**, the soft manipulator is fixed downward, viewing the workspace scene built from swine viscera to simulate the endoscopic scene in laparoscopy. The depth of the scene surface relative to the camera front end is about 3 cm.

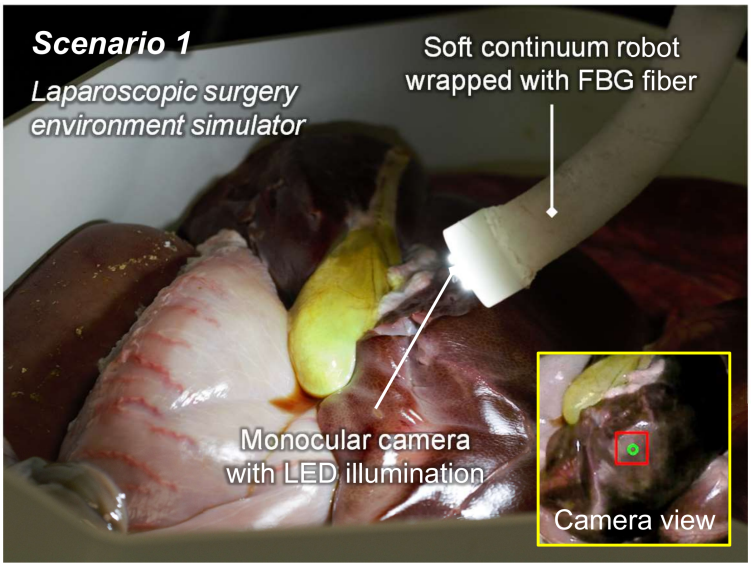


Fig. 3.11: Simulated scene mimicking the laparoscopic surgery environment. Pre-bending of the robot is applied to simulate the laparoscope setup. EM tracking coils are equipped to measure the instantaneous poses of robot tip.

The overall visual servoing accuracy is determined by both the motion estimation and robot controller. To evaluate their independent effects, two kinds of tracking errors are defined. The first one is measured only from the image frames, assuming the motion estimation to be accurate. Thus, this tracking error only depends on the controller performance (i.e., **Fig. 3.12a** and **Fig. 3.13a**). The other kinds of error takes account of the ground-truth measurement from EM tracking coils, thus including the error of estimated motion (**Fig. 3.12c** and **Fig. 3.13c**). With the aim of enhancing motion estimation accuracy in this section, the second kinds of errors in the path following tasks are compared, with versus without FBGs sensing combined. A pair of 6-dimensional (6D) EM tracking coils (NDI Aurora®) are fixed laterally on the robot tip to measure its ground-truth poses, based on which 3D motion vectors between successive steps can be calculated. Given that the normal direction of the image plane always aligns with that of the robot tip, the 3D motion can be projected to 2D motion on the image plane. A series of projected 2D motions constitutes the actual tracking path in the image frame.

To calculate the 3D motion vectors, the 3D positions and 4D quaternions of two EM tracking probes are denoted as $\boldsymbol{p}_i(k)$, and $\boldsymbol{h}_i(k)$, respectively, $i = 1,2$. The center position of the robot tip is measured as

$$\boldsymbol{p}(k) = (\boldsymbol{p}_1(k) + \boldsymbol{p}_2(k))/2 \tag{3.22}$$

The quaternions $\boldsymbol{h}_1(k)$ or $\boldsymbol{h}_2(k)$ can be transformed to rotation matrix $\boldsymbol{R}(k)$. The initial local coordinate at robot tip is represented by three orthogonal unit vector, $\boldsymbol{u}_0 = [\ 1 \quad 0 \quad 0\ ]^T$, $\boldsymbol{v}_0 = [\ 0 \quad 1 \quad 0\ ]^T$ and $\boldsymbol{n}_0 = [\ 0 \quad 0 \quad 1\ ]^T$. At time step k during robot operation, the basis vectors are transformed to $\boldsymbol{u}(k) = \boldsymbol{R}(k) \cdot \boldsymbol{u}_0$, $\boldsymbol{v}(k) = \boldsymbol{R}(k) \cdot \boldsymbol{v}_0$ and $\boldsymbol{n}(k) = \boldsymbol{R}(k) \cdot \boldsymbol{n}_0$. Given the 3D motion vector $\Delta\boldsymbol{p}(k) = \boldsymbol{p}(k) - \boldsymbol{p}(k-1)$, and the current camera frame, $\boldsymbol{u}(k)\boldsymbol{v}(k)$, the projected motion vector is obtained as:

$$\Delta\hat{\boldsymbol{p}}(k) = \Delta\boldsymbol{p}(k) - \frac{\Delta\boldsymbol{p}^T(k) \cdot \boldsymbol{n}(k)}{|\boldsymbol{n}(k)|^2} \cdot \boldsymbol{n}(k) \tag{3.23}$$

w.r.t. the coordinate of EM generator. The 2D motion vector w.r.t. the image coordinate $\boldsymbol{u}(k)\boldsymbol{v}(k)$ can be given as:

$$\Delta\boldsymbol{z}(k) = \left[\begin{array}{cc} \dfrac{\Delta\hat{\boldsymbol{p}}^T(k) \cdot \boldsymbol{u}(k)}{|\boldsymbol{u}(k)|} & \dfrac{\Delta\hat{\boldsymbol{p}}^T(k) \cdot \boldsymbol{v}(k)}{|\boldsymbol{v}(k)|} \end{array} \right]^T \tag{3.24}$$

The local camera frame may be disorientated from the EM coordinate by a twisting angle

θ, which needs to be calibrated so as to match the 2D motion estimated from EM coils, $\Delta \boldsymbol{z}(k)$, with that measured through image processing, $\Delta \boldsymbol{z}_c(k)$. As the angle is constant, it can be obtained as the mean angle difference calculated from a sampling data set. Then the angle bias can be corrected to find the aligned motion vector:

$$\Delta \boldsymbol{z}_s(k) = \boldsymbol{T} \cdot \Delta \boldsymbol{z}(k) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \cdot \Delta \boldsymbol{z}(k) \tag{3.25}$$

The motion vector $\Delta \boldsymbol{z}_s(k)$ is used in the evaluation of motion estimation accuracy with and without FBG enhancement.

Visual Servo in Feature-deficient Scenes

The improved motion estimation method has been validated via path following tasks. In **Section 3.4.3**, a desired path in the "∞" shape [181] is tested. The smooth path is defined by 72 points on the 400×400 pixels camera frame (with the mean interval of 15.7 pixels), generating a closed curve with a continuous gradient. To increase the tracking difficulty, the path designed in this section is the outline of the Batman logo with several acute angles, consisting of 127 points with a mean interval of 7.9 pixels. The robot is instructed to track the same template pattern in the laparoscopic scene (**Fig. 3.11**) along the Batman path for 3 cycles, using template matching alone (**Fig. 3.12**) and the enhanced method proposed in this section (**Fig. 3.13**), respectively.

In the camera view, the tracking trajectory of the target block roughly accords with the Batman path when the 2D motion estimation is achieved by both template matching and FBG-enhanced method (**Fig. 3.12a**, RMSE: 8.6 pixels, **Fig. 3.13a**, RMSE: 8.6 pixels), benefitting from the well-performed closed-loop controller. However, the accuracy of motion estimation can be reflected in the actual projected path of the end-effector (**Fig. 3.12c**, **Fig. 3.13c**), which is supposed to be similar to **Fig. 3.12a** and **Fig. 3.13a**. However, due to the accumulation of estimation errors, the ground-truth tracking path deviates from the reference path. The tracking target was intentionally defined on the surface of the liver lacking features, and was finally lost from tracking (**Fig. 3.12b**). The initial and final camera views with the tracking block were captured and shown in **Fig. 3.12b** and **Fig. 3.13b**, where the black dotted block denotes the initial target.

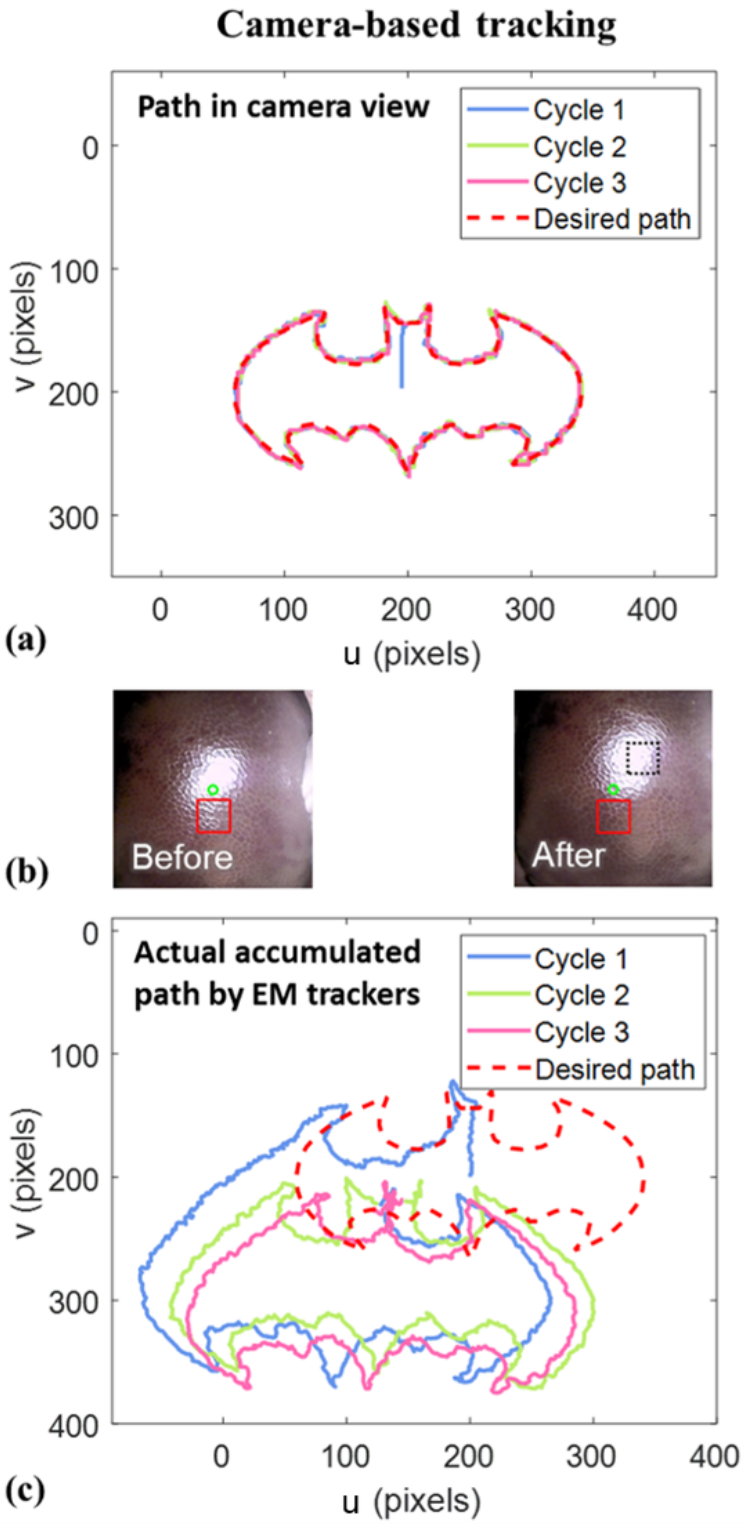


Fig. 3.12: Robot following of a predefined "Batman" path in the scene of **Fig. 3.11**. **(a)** The motion estimated by image processing of the endoscopic view *alone*. **(b)** Bright reflective spot displaced along the tissue due to the robot motion. Features in the red block are selected by the user *before* the motion displacement (Left). *After*, such a block is expected to keep matching/tracking at the same square of features (Right. Black dotted block indicates the position of red block selected at the previous time step), acting as a static reference for robot to "draw" the path. **(c)** Actual path of end-effector recorded by EM tracking coils, which project on the same $u-v$ coordinates. The recorded deviation along those 3 cycles is caused by the error of such a red block matching/tracking.

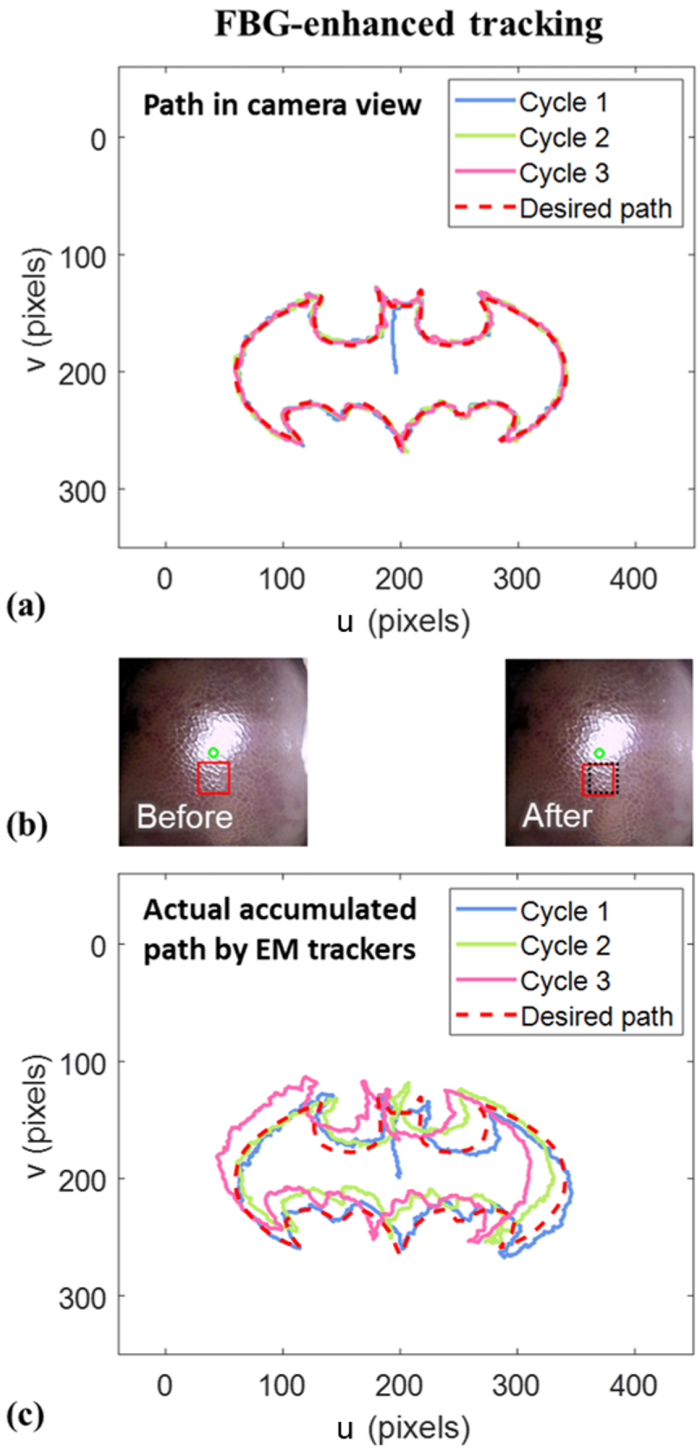


Fig. 3.13: Tracking performance of the same "Batman" path in the scene of **Fig. 3.11**, with the motion estimated by FBG-enhanced method. **(a)** Estimated motion of tracked feature in the endoscopic view. **(b)** After 3 cycles, the offset of tracked feature from the red block to the black dotted block is obviously reduced, resulting in more accurate tracking in **(c)** recorded by the EM tracking probe.

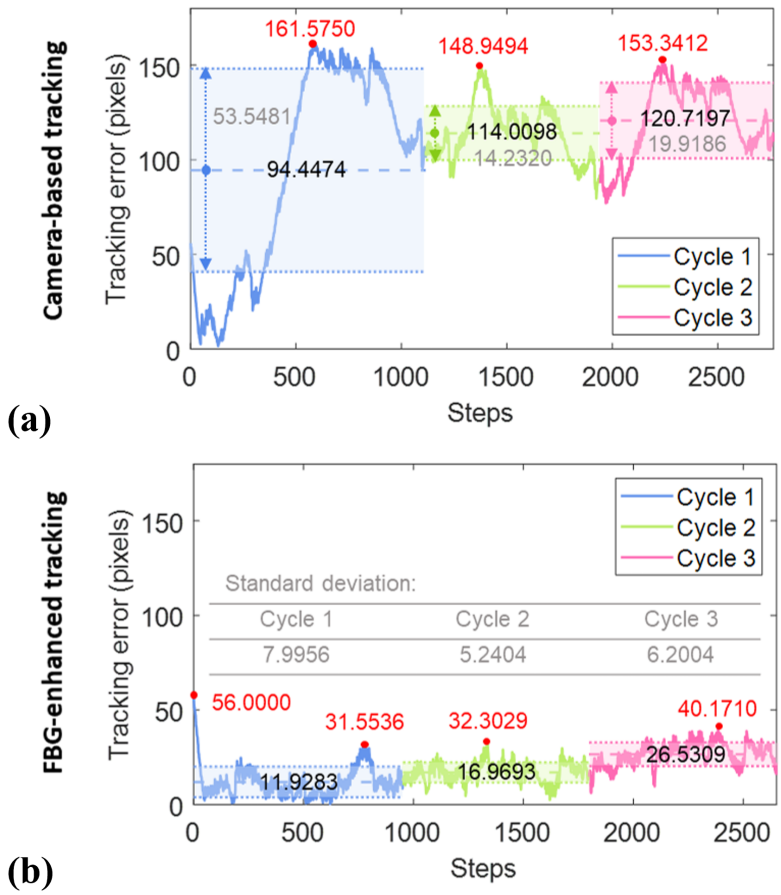


Fig. 3.14: Errors of "Batman" path following with a simulated laparoscopic setup (**Fig. 3.11**). **(a)** Visual servoing with camera-based motion estimation, inducing a mean error of 94.45 (SD: 53.55) in the 1^{st} cycle. **(b)** Visual servoing with FBG-enhanced motion estimation. The tracking errors correspond to the trajectories in **Fig. 3.12c** and **Fig. 3.13c**, with red dots denoting the maximum error in each cycle.

The actual tracking errors measured with EM coils are depicted in **Fig. 3.14**. The RMSE for each cycle gradually increases over three successive cycles, which are 108.6, 114.9, and 122.4 pixels when using solely-image-based estimation, and 14.4, 17.8, and 27.2 pixels when integrating FBGs sensing. With the FBGs enhancement, an 82.3% reduction of the visual servoing error has been achieved.

As the template matching method depends on the features in the camera view, it could be inaccurate in scenes lacking features or with a uniform pattern, like the liver surface test, thus giving rise to errors in motion estimation and visual seroving control. However, FBGs sensing will not be affected by any disturbances in the camera view, standing to be a promising modality that can be combined with the camera to enable enhanced motion estimation and soft robotic visual servoing.

3.6 Augmented Reality-enhanced Visual Servoing

Eye-in-hand visual servoing controllers (**Section 3.3** and **Section 3.5**) can be used for soft robotic navigation solely based on the camera view, which may not cover all the regions of interest. Therefore, a further improvement of camera feedback is proposed by incorporating the information of medical imaging. Augmented reality (AR) is a potential technique that could enhance the camera image guidance, which is beneficial for robotic interventions. This section presents the integration of visual servoing and AR for the control of a flexible robotic endoscope. Pre-operatively segmented anatomy can be visualized in the endoscopic view via AR to improve awareness of critical structures and provide guidance towards regions of interest (ROIs). With AR, the camera view can also be virtually extended beyond the normal field-of-view, enhancing positional awareness relative to out-of-view ROIs. Combined with visual servoing control of a flexible endoscope, the surgeon can directly prescribe targets in the image frame for automated navigation. These concepts is applied experimentally with an anatomical skull phantom. A virtual tumor target located in the extended view was successfully tracked. The key steps of our experimental validation are introduced as bellow:

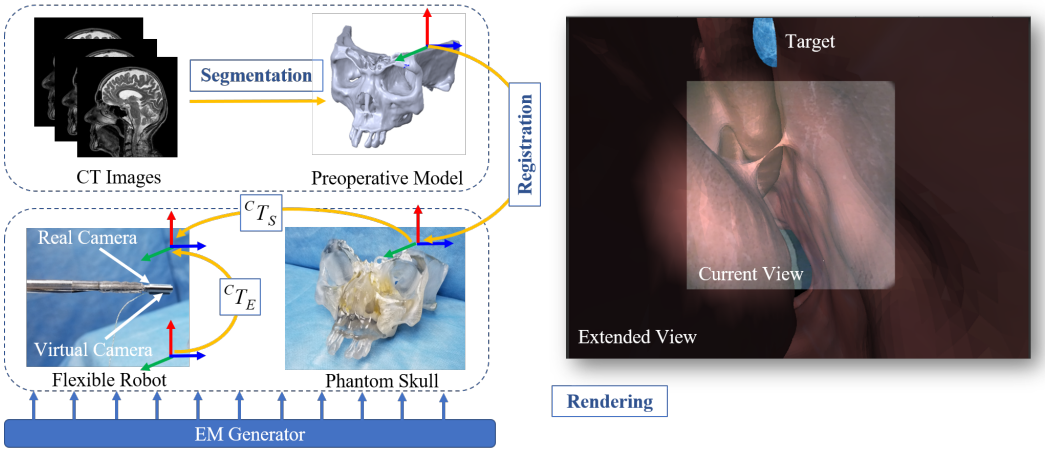


Fig. 3.15: Flowchart of the coordinate registration loop of the proposed augmented reality system. The pre-op model is registered to the phantom skull. EM tracking is implemented to register the virtual camera which renders an extended virtual view overlaid on the real camera view. The extended view only contains virtual objects, whereas the current view contains both virtual and real camera images.

Coordinate Registration

The coordinate registration loop of the proposed AR system is shown in **Fig. 3.15**. The pre-op skull model can be segmented from pre-op 3D images, e.g., CT images. Six feature points on the phantom skull can be selected and pinpointed with an EM tracking probe

for the registration process. Point-based rigid registration is performed to locate the skull model in the EM coordinates. An EM tracking marker is used to track the camera position and orientation. The accurate and stable registration of the camera pose with respect to the pre-op model is the key step for achieving AR augmentation. The pre-op skull model can be rendered according to the EM tracking position, which constructs the virtual camera. By aligning the virtual and real camera view, an augmented view is generated. The virtual view extends beyond the real camera view, offering increased FOV and greater visualization of surrounding anatomy and segmented ROIs. The transformation process can be represented as:

$$^{C}\boldsymbol{T}_{S} = {}^{C}\boldsymbol{T}_{E}.{}^{E}\boldsymbol{T}_{S} \tag{3.26}$$

where the transformations $^{C}\boldsymbol{T}_{S}$, $^{C}\boldsymbol{T}_{E}$, $^{E}\boldsymbol{T}_{S} \in \mathbb{R}^{4\times4}$ are respectively measured from the monocular camera reference to the world coordinate (skull) frame, camera frame to the EM tracker frame, and the EM tracker frame to the skull frame. The transform $^{C}\boldsymbol{T}_{E}$ between EM tracking and camera coordinates is estimated using a checkboard-based calibration method [182].

Experimental Setup for AR-Guided Visual Servoing

The experimental setup is shown in **Fig. 3.16**. A phantom skull model was 3D-printed based on segmented bony landmarks within CT images. To fit within the confined space of the phantom's nasal cavity, a compact tendon-driven soft robot with 2 bending DoFs was fabricated with an outer diameter of 4 mm. The robot is integrated with a monocular camera with a resolution of 400×400 pixels. A 3D-printed mounting was used to hold the soft robot body. The robot tip was attached with a 6-DoFs EM tracking probe. The EM field generator was placed beneath the soft robot and phantom skull model. The phantom skull is fixed relative to the EM generator.

Augmented Reality Overlay

The augmented reality interface is comprised of the virtual extended view, current camera view, and target, as shown in **Fig. 3.17**. A virtual tumor target (highlighted in blue) is fixed to the virtual skull model and is not present in the real skull. The experiment was conducted with the tip of the robot inserted into the nasal cavity at a depth of approximately 1 cm. The skull anatomy is aligned with respect to the real camera view, with a fiducial registration error smaller than 1.3 mm. The registration accuracy is comparable to commercial navigation systems for surgery (e.g., StealthStation™, Medtronic) [183].

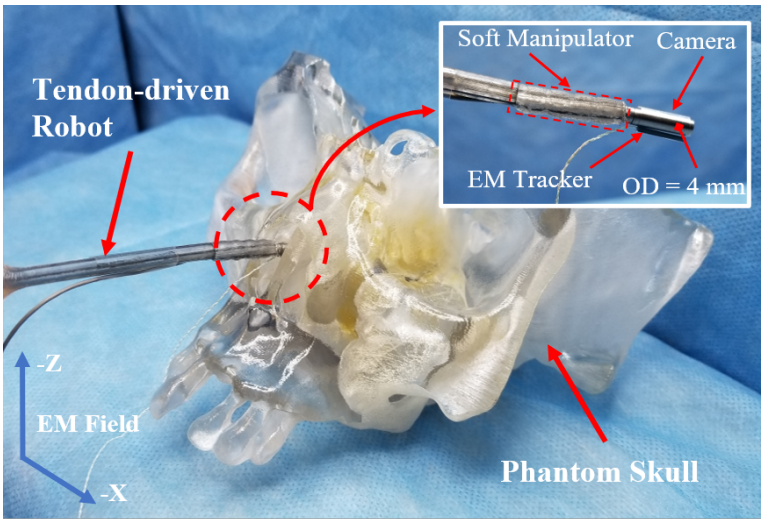


Fig. 3.16: Experimental setup for augmented reality guided visual servoing. A compact tendon-driven robot (outer diameter 4 mm) is integrated with a monocular camera. An EM generator is placed beneath the 3D printed skull phantom and tendon-driven robot.

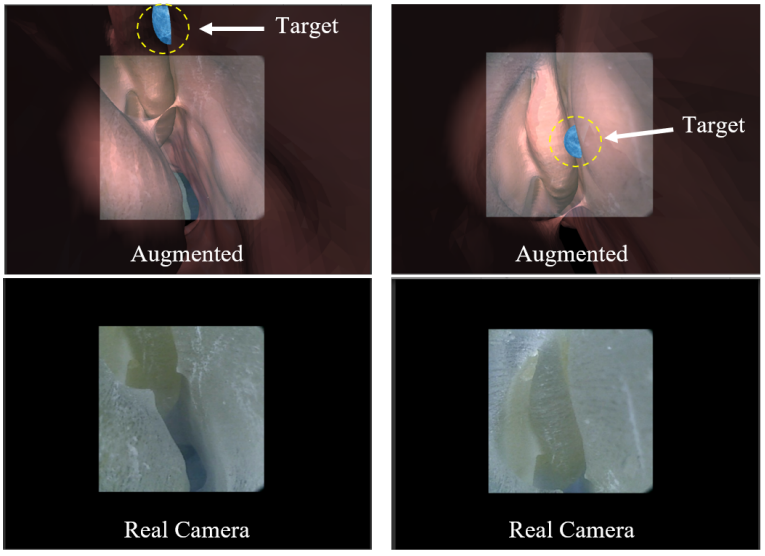


Fig. 3.17: Augmented reality interface. The augmented view is generated by aligning the virtual camera with the real endoscopic camera through hand-eye calibration. **(a)** Target initially located in extended view (above), with corresponding real camera view shown below. **(b)** Augmented view after visual servoing of robotic endoscope towards the target (above) and corresponding real camera view (below).

Visual Servoing Combined with AR

In this experiment, the robot is controlled to aim the camera center at a target outside the initial camera view. AR is utilized to display the target in an extended view, which creates visual guidance for robot tracking. As shown in **Fig. 3.17a**, the target tumor is initially located in the extended view and is selected by the user. After visual servoing of the robotic endoscope for target tracking (**Section 3.3**), the virtual tumor is located in the center of

the camera view, as shown in **Fig. 3.17b**. Four intermediate waypoints (labels 2 to 5) were manually selected from the image plane. The robot traced the target waypoints successively at its image plane center. **Fig. 3.18a** shows the mosaic image obtained throughout the tracking journey. A series of yellow circles denote the image plane center at each time step. **Fig. 3.18b** presents the tracking errors, showing that the robot achieves accurate tracing of each desired target with an error <10 pixels. The extended time required to converge to each waypoint target (error plateaus indicated by pink arrows) is induced by the backlash of the tendon-driven robot actuation.

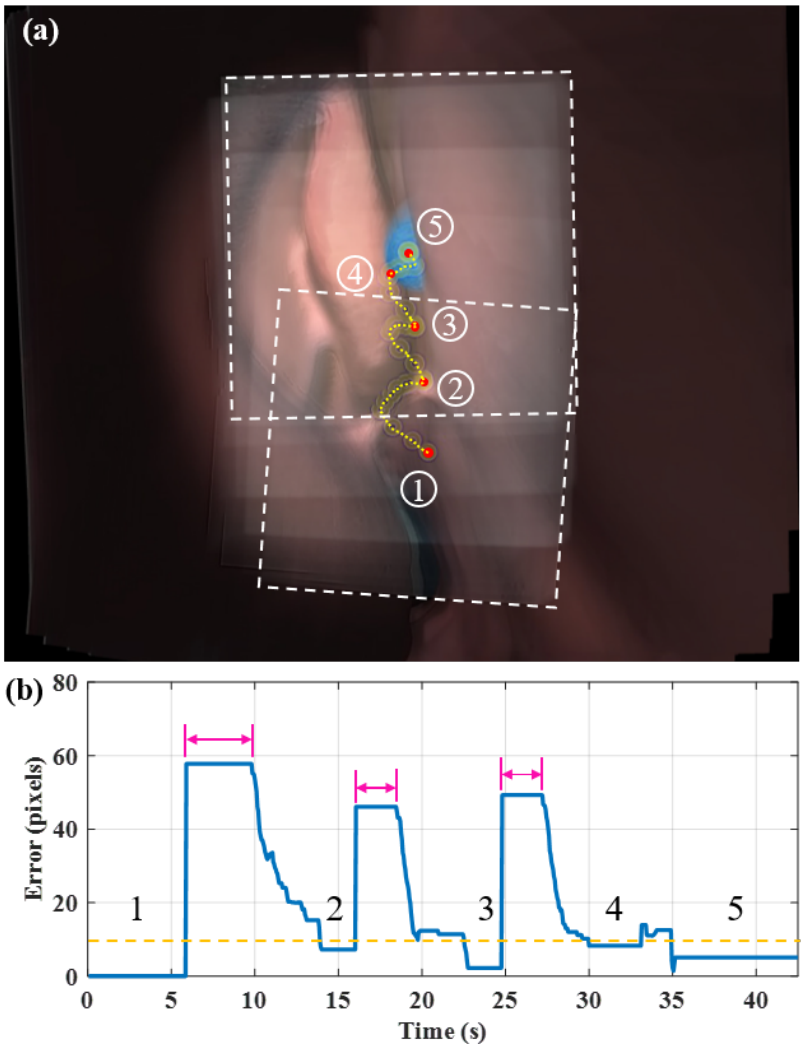


Fig. 3.18: Tracking of a target located in the extended view. Four intermediate way-points are manually selected in the camera view for robot tracking as in **Fig. 3.6**. **(a)** Mosaic image obtained during the tracking journey. **(b)** Absolute tracking errors in pixels. The plateaus of error (last for 2-4 s) (indicated by pink arrows) are due to mechanical backlash of the tendon-driven robot.

3.7 Conclusion

This chapter proposes a nonparametric online learning control framework, enabling precise eye-in-hand visual servoing of a pneumatic-driven soft robot made from hyper-elastic materials. The controller is initialized by learning the inverse transition mapping solely from collected sensing data, thus alleviating the need for analytically modeling the robot kinematics or calibrating camera intrinsic/extrinsic. Promising accuracy in the learning of inverse mapping is assured without having to tune the hyperparameters by the nonparametric learning approach. Localized GPR models enable fast online updates to accommodate new input data that reflect the latest robot status. Compared to existing model-based visual servoing methods, which did not consider unknown mechanical disturbances, the proposed learning-based method first resolves unknown variable external loading. It has the potential to be used in minimally invasive surgical applications, e.g., endoscopy and laparoscopy [152, 184].

The online learning controller achieved precise path following for a fluid-driven soft robot with varying payload applied on its tip (up to 105% of the robot weight). Stable tracking was maintained even when the load was quickly increased ($\sim$ 62.5% of robot weight over 1.5 s), while a slowly-varied payload can be effectively compensated. After removing the loading, path following accuracy further increased with an RMSE of 5.6 pixels and a maximum of 10.8 pixels. In future work, the prior knowledge of robot or camera parameters will be incorporated with the learning framework. For example, the LGPR models can be initialized with virtual data simulated from finite element analysis [167], simplifying the process of data collection for pre-training. More dynamic tasks will be conducted to evaluate the proposed online learning control. Visual seroving for soft robots with multiple segments, involving high redundancy, will be investigated.

Strain sensing by FBGs is combined to enhance the visual servo performance in feature-deficient scenes. AR can be integrated to virtually extending the camera view beyond the normal FOV, thus offering enhanced guidance for operators.

Chapter 4

Eye-to-hand Vision-based Control for Soft Robotic Laser Steering

4.1 Introduction

Soft robotic manipulators integrating flexible lasers bring new opportunities for minimally invasive surgery. These robots could be fabricated in a compact size and equipped as an endoscope tip. Furthermore, made of non-ferromagnetic materials, they allow further applications to MRI-guided laser surgery, which could enhance safety by providing intra-op monitoring of laser ablation progress. However, the fluid actuation via long water pipes designed for MRI safety further complicates the soft robot kinematics. To achieve precise laser projection ($<$1 mm, [107]) in a confined workspace and avoid undesirable tissue damage, a controller with high accuracy and stability is required. An effective way is to close the control loop. Cameras are used normally to provide visual feedback to surgeons. Using computer vision techniques to process camera images, we can obtain the laser spot position on the surgical site.

Eye-in-hand setup with the camera moving with robot tip (**Chapter 3**) is normally used for guiding robotic navigation to a targeted ROI. In contrast, laser spot steering is conducted with the camera immobile and viewing a fixed ROI, which is an eye-to-hand setup. Vision-guided laser steering control was investigated to automate laser spot projection [25, 185]. Epipolar or trifocal visual servoing demonstrated excellent control accuracy in laser spot steering. However, the geometric model of actuated mirrors was utilized to construct the controller, which is not suitable for use in soft laser manipulators. In contrast

to actuated mirrors with well-formulated kinematics, soft robotic laser manipulator has nonlinear kinematics and a time-varying origin of the laser beam, resulting in that the robot kinematics is coupled with the epipolar geometry.

In this chapter, a geometric model that combines the epipolar constraint with the soft robot kinematics is proposed to achieve precise control of laser spot projection. Through this approach, the inverse transition mapping from image space to actuation space can be obtained without limitation to the projection tissue geometry. A learning-based modeling approach is also proposed to eliminate the need for any prior information about the robot system, including camera calibration. Based on the inverse transition mapping, feedback controllers could be formed to ensure accurate laser spot targeting in the camera view. The control performance is validated on an MR-safe robotic laser manipulating platform integrated with a fiber-optic camera. The key contributions of this work are listed as follows:

1. Visual servoing of soft laser manipulator based on epipolar geometry modeling, which enables precise laser spot control without prior knowledge of the tissue surface geometry;
2. Data-driven controller enabling accurate visual servoing of the laser spot along a path prescribed *in-situ* on the camera view;
3. Experimental validation of the proposed laser spot visual servoing on target tracking and path following tasks.

4.2 Schematics of the Laser Steering System

To control the laser beam projection on oral tissue, a desired path or region for laser ablation is defined on a visualization interface. Using this path, the controller can automatically steer the laser spot on the 2D image plane. As shown in **Fig. 4.1**, a hydraulic-driven soft robot is used to manipulate the laser fiber collimator, which is housed in the central channel of the robot. An external camera is equipped on the outer shell to provide visual feedback of the laser spot on the surgical site. The actuation space of soft manipulator is represented by the lengths of soft chambers, $\mathbf{L}$; adjusting the chamber lengths varies the configurations of the soft manipulator, which drives the collimator towards different directions; the vectors $\tilde{\mathbf{p}}_0$ and $\tilde{\mathbf{p}}_1$ are the laser spot images in the camera view; the linear displacements Λu and Λv are

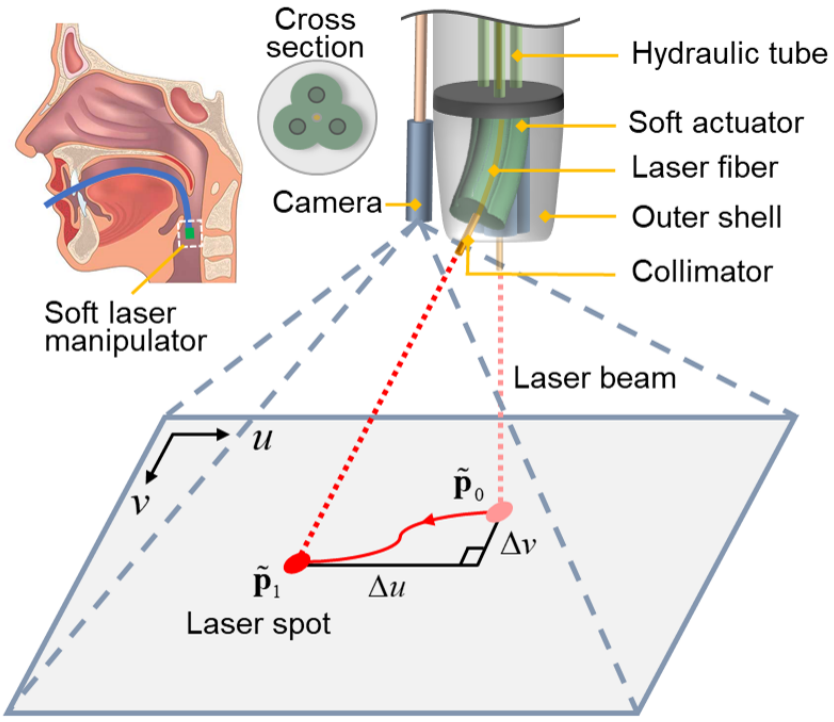


Fig. 4.1: Schematics of the laser steering system. A soft manipulator is actuated by adjusting the fluid volume of soft chambers, which steers the laser spot from $\tilde{\mathbf{p}}_0$ to $\tilde{\mathbf{p}}_1$ within the camera view.

along u and v directions respectively. The nonlinear inverse transition mapping from image space to robot actuation is modeled based on constant curvature (CC) robot kinematics and epipolar geometry. Referring to the desired laser spot position and current robot state, actuation commands can be generated through the inverse transition mapping.

4.3 Laser Spot Detection

Precise detection of the actual laser spot position on the projection scene is the prerequisite of closed-loop control, thus allowing automatic correction of the targeting errors. Due to the monochrome and high intensity of the laser spot, the laser aiming region can be extracted from the live camera video based on intensity thresholding [186]. The processing steps are described below: 1) Thresholding: The captured red-green-blue (RGB) image is converted to a grayscale image, or we can extract a specific color channel according to the color of the aiming light. The image is binarized by setting the intensity values higher than a threshold as 255, with others to zero. The threshold value can be experimentally obtained using the corresponding laser projection setup. This step produces a binary output image. 2) Filtering: An erosion filter is applied to remove small particles that remain in the binary image. This

process clears the background noise and only keeps the laser spot blob in the output image. 3) Centroid calculation: By averaging the coordinates of all remaining white pixels, the centroid position of the laser spot in the image frame can be obtained.

Due to the scattering effect of light on tissue, the laser spot seen in the camera image could be irregular, non-circular, or even split into multiple laser spots. This introduces noise into the centroid calculation and may cause sudden jumps of the measured laser spot position. To ensure smooth tracking of the laser spot, a linear motion model is integrated to predict its position, which is combined with the measured centroid position through Kalman filtering. The linear motion model is defined as:

$$\begin{bmatrix} \tilde{\mathbf{p}}^*_{t+dt} \\ \dot{\tilde{\mathbf{p}}}^*_{t+dt} \end{bmatrix} = \Lambda \begin{bmatrix} \tilde{\mathbf{p}}_t \\ \dot{\tilde{\mathbf{p}}}_t \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \mathbf{I}_{2\times 2} & dt\,\mathbf{I}_{2\times 2} \\ \mathbf{0}_{2\times 2} & \mathbf{I}_{2\times 2} \end{bmatrix} \tag{4.1}$$

where Λ is the state transition matrix; subscript "t" and "t+dt" denote the time step t and t+dt, respectively; the sampling period of camera images is denoted by dt; the combined position and velocity are represented by $\tilde{\mathbf{p}}_t$ and $\dot{\tilde{\mathbf{p}}}_t$ respectively; the predicted laser spot position and velocity are denoted by $\tilde{\mathbf{p}}^*_{t+dt}$ and $\dot{\tilde{\mathbf{p}}}^*_{t+dt}$ respectively. The observation equation is given as:

$$\tilde{\mathbf{p}}_{obs} = \mathbf{H} \begin{bmatrix} \tilde{\mathbf{p}}^*_{t+dt} \\ \dot{\tilde{\mathbf{p}}}^*_{t+dt} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{I}_{2\times 2} & \mathbf{0}_{2\times 2} \end{bmatrix} \tag{4.2}$$

where $\mathbf{H}$ is the observation matrix, and $\tilde{\mathbf{p}}_{obs}$ is the laser spot position observed from the model. The combined laser spot position $\tilde{\mathbf{p}}_{t+dt}$ and velocity $\dot{\tilde{\mathbf{p}}}_{t+dt}$ are calculated as:

$$\begin{bmatrix} \tilde{\mathbf{p}}_{t+dt} \\ \dot{\tilde{\mathbf{p}}}_{t+dt} \end{bmatrix} = \Lambda \begin{bmatrix} \tilde{\mathbf{p}}_t \\ \dot{\tilde{\mathbf{p}}}_t \end{bmatrix} + \mathbf{G}_t \left(\tilde{\mathbf{p}}_{mea} - \tilde{\mathbf{p}}_{obs} \right) \tag{4.3}$$

where $\tilde{\mathbf{p}}_{mea}$ is the measured laser spot position from camera images; the Kalman gain at time step t is denoted by $\mathbf{G}_t$, which can be updated as follows:

$$\begin{aligned} \hat{\mathbf{X}}_{t+dt} &= \Lambda\, \mathbf{X}_t\, \Lambda^T + \mathbf{Q}; \\ \mathbf{G}_{t+dt} &= \hat{\mathbf{X}}_{t+dt}\, \mathbf{H}^T \big/ \big(\mathbf{H}\, \hat{\mathbf{X}}_{t+dt}\, \mathbf{H}^T + \mathbf{Y} \big); \\ \mathbf{X}_{t+dt} &= \left(\mathbf{I}_{4\times 4} - \mathbf{G}_t\, \mathbf{H} \right) \hat{\mathbf{X}}_{t+dt} \end{aligned} \tag{4.4}$$

where $\hat{\mathbf{X}}_{t+dt}$ and $\mathbf{X}_{t+dt}$ denote the predicted and updated state covariance matrix at time step t+dt, respectively; the covariance matrix of process noise is denoted by $\mathbf{Q}$; the covariance matrix of measurement noise is represented by $\mathbf{Y}$.

4.4 Epipolar Geometry Modelling for Soft Robotic Laser Steering Control

In this section, epipolar geometry is investigated to relate the soft robot configuration and laser spot image, thus establishing the inverse transition mapping from image space to robot actuation. With this inverse mapping, a feedback control law is derived without prior information of tissue geometry. Experimental validation demonstrates that the controller enables accurate path following for the soft laser manipulator.

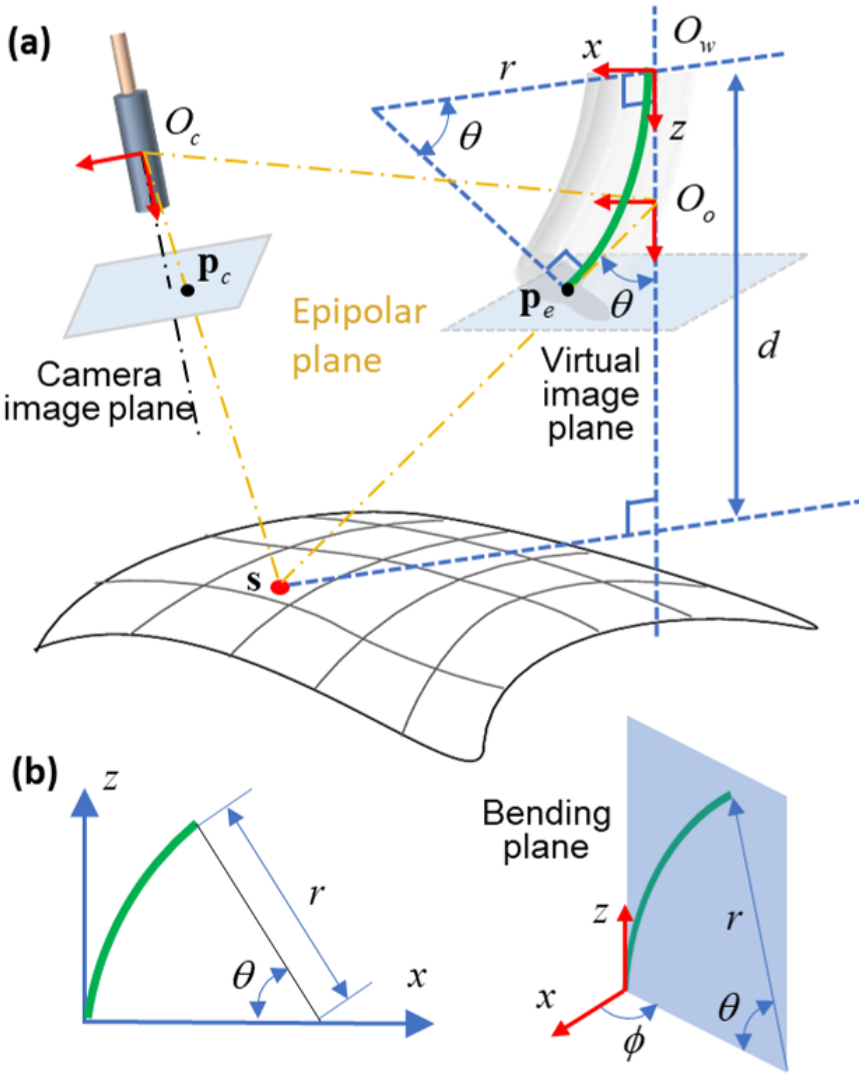


Fig. 4.2: **(a)** Illustration of the epipolar geometry-based modeling. The 3-D coordinates O_c and O_o denote the camera center and the robot center, respectively. Along with the laser projection spot s, these three points form the epipolar plane. **(b)** Three main parameters characterize the robot configuration, namely the bending radius r, plane angle ϕ and bending angle θ.

4.4.1 Soft Robots Modelling

As shown in **Fig. 4.2**, the laser manipulator is represented as a bending curve. Point O_o is the intersection of the tangent lines at robot base O_w and robot tip. Assuming the soft manipulator bends with constant curvature, we can obtain the translation vector of O_o relative to the base coordinate O_w:

$$ {}^{w}_{o}\mathbf{t} = \begin{bmatrix} 0 & 0 & r\tan\frac{\theta}{2} \end{bmatrix}^T \tag{4.5} $$

Through hand-eye calibration, we can obtain the homogeneous transformation ${}^{c}_{w}\mathbf{T}$, the rotation matrix ${}^{c}_{w}\mathbf{R}$, and the translation vector ${}^{c}_{w}\mathbf{t}$, from robot base coordinate O_w relative to camera coordinate O_c. Coordinate O_o is defined in parallel to the base coordinate O_w; thus, the transformation matrix ${}^{w}_{o}\mathbf{T}$ from coordinate O_o relative to O_w can be determined only by the translation vector ${}^{w}_{o}\mathbf{t}$. The parallelism is consistently satisfied since the coordinate O_o is defined referring to the CC kinematic model's coordinate. The translation distance/vector can be estimated based on the geometric relationship in the model. Then, we have the homogeneous transformation between coordinate O_o and O_c:

$$ {}^{c}_{o}\mathbf{T} = {}^{c}_{w}\mathbf{T}\,{}^{w}_{o}\mathbf{T} = \begin{bmatrix} {}^{c}_{w}\mathbf{R} & {}^{c}_{w}\mathbf{R}\,{}^{w}_{o}\mathbf{t} + {}^{c}_{w}\mathbf{t} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{4.6} $$

Virtual camera and epipolar constraint:

Let us denote $\mathbf{s}$ as the spatial projection spot of the laser beam, $\mathbf{p}_c$ as the position of the laser spot in the camera coordinate, $\mathbf{p}_e$ as the position of robot tip in the reference coordinate O_o, which can be assumed as a virtual camera. The laser beam is considered as the optical axis, and the robot tip is regarded as the virtual image of the laser spot. Therefore, points $\mathbf{p}_c$ and $\mathbf{p}_e$ can be regarded as the images of the same spatial point $\mathbf{s}$ in two camera frames. In the frame of the virtual camera, we can concisely form the laser projection vector based on the robot configuration, i.e., **Eq. 4.8**. The epipolar constraint [187] is accordingly applied to link the two frames:

$$ \mathbf{p}_e^T\,\mathbf{F}\,\mathbf{p}_c = 0 \tag{4.7} $$

where $\mathbf{p}_e$ can be expressed as:

$$ \mathbf{p}_e = \begin{bmatrix} \|\mathbf{p}_e\|\sin\theta\cos\phi & \|\mathbf{p}_e\|\sin\theta\cdot\sin\phi & \|\mathbf{p}_e\|\cos\theta \end{bmatrix}^T \tag{4.8} $$

where $\|\mathbf{p}_e\| = r\cdot\tan\left(\theta/2\right)$. The bending radius of soft manipulator is denoted by r; the plane

angle is represented by ϕ; and the bending angle is given by θ, as defined in **Fig. 4.2b**. The fundamental matrix of the two camera frames, **F**, can be given as [187]:

$$\mathbf{F} = {}^{c}_{o}\mathbf{R}^{T}\ ({}^{c}_{o}\mathbf{t})^{\wedge} = {}^{c}_{w}\mathbf{R}^{T}\ ({}^{c}_{w}\mathbf{R}\,{}^{w}_{o}\mathbf{t} + {}^{c}_{w}\mathbf{t})^{\wedge} \tag{4.9}$$

where $(\cdot)^{\wedge}$ maps a vector to an rotation matrix, based on the Lie algebra of rotation group SO(3) [188], as bellow:

$$\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \right)^{\wedge} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{4.10}$$

As shown in **Fig. 4.2a**, the spatial point **s** and its image in the camera and robot (virtual camera) frames, namely $\mathbf{p}_c$ and $\mathbf{p}_e$, and the origins of the two frames, O_o and O_c, form the epipolar plane. The epipolar constraint in **Eq. 4.7** can be decomposed as the dot product of two orthogonal vectors in the robot frame, $\{O_o\}$, and the camera frame, $\{O_c\}$, respectively:

$$\mathbf{p}_e^{T} \cdot (\mathbf{F}\,\mathbf{p}_c) = 0 \ , \ \ {\mathbf{p}_c}^{T} \cdot \left(\mathbf{F}^{T}\,\mathbf{p}_e\right) = 0 \tag{4.11}$$

where $\mathbf{F}\,\mathbf{p}_c$ and $\mathbf{F}^{T}\,\mathbf{p}_e$ are the normal vectors to the epipolar plane, which are expressed in the frame of robot and camera, respectively [25].

Laser spot position, $\mathbf{p}_c$, is normalized in the form of $[p_x \quad p_y \quad 1\ ]^{T}$, and linked with its homogeneous pixel coordinate $\tilde{\mathbf{p}}$, namely $[u \quad v \quad 1\ ]^{T}$, by multiplying camera intrinsic matrix **K**:

$$\tilde{\mathbf{p}} = \mathbf{K}\,\mathbf{p}_c \tag{4.12}$$

By differentiating the epipolar constraint in **Eq. 4.7**, we get:

$$\left(\mathbf{F}^{T}\,\mathbf{p}_e\right)^{T}\dot{\mathbf{p}}_c + (\mathbf{F}\,\mathbf{p}_c)^{T}\,\dot{\mathbf{p}}_e + \mathbf{p}_e^{T}\,\dot{\mathbf{F}}\,\mathbf{p}_c = 0 \tag{4.13}$$

where $\dot{\mathbf{p}}_c$ is the velocity of the laser spot. The time derivative of robot tip position $\dot{\mathbf{p}}_e$ (**Eq. 4.8**) can be linearized with a Jacobian matrix $\mathbf{J}_1$, namely:

$$\dot{\mathbf{p}}_e = \mathbf{J}_1 \begin{bmatrix} \dot{r} & \dot{\theta} & \dot{\phi} \end{bmatrix}^{T} \tag{4.14}$$

The Jacobian matrix $\mathbf{J}_1$ can be derived by differentiating **Eq. 4.8** relative to the robot configurations, $\begin{bmatrix} r & \theta & \phi \end{bmatrix}$:

$$\mathbf{J}_1 = \begin{bmatrix} \frac{\partial \mathbf{p}_e}{\partial r} & \frac{\partial \mathbf{p}_e}{\partial \theta} & \frac{\partial \mathbf{p}_e}{\partial \phi} \end{bmatrix} \tag{4.15}$$

where each term is calculated as:

$$\frac{\partial \mathbf{p}_e}{\partial r} = \begin{bmatrix} \tan\frac{\theta}{2}\cdot\sin\theta\cdot\cos\phi & \tan\frac{\theta}{2}\cdot\sin\theta\cdot\sin\phi & \tan\frac{\theta}{2}\cdot\cos\theta \end{bmatrix}^T \tag{4.16}$$

$$\frac{\partial \mathbf{p}_e}{\partial \theta} = \begin{bmatrix} r\cdot\cos\phi\cdot\left(\tan\left(\theta/2\right)\cdot\cos\theta + \sin\theta\cdot\frac{1}{2\cos^2\left(\theta/2\right)}\right) \\ r\cdot\sin\phi\cdot\left(\tan\left(\theta/2\right)\cdot\cos\theta + \sin\theta\cdot\frac{1}{2\cos^2\left(\theta/2\right)}\right) \\ r\left(-\tan\left(\theta/2\right)\cdot\sin\theta + \cos\theta\frac{1}{2\cos^2\left(\theta/2\right)}\right) \end{bmatrix} \tag{4.17}$$

$$\frac{\partial \mathbf{p}_e}{\partial \phi} = \begin{bmatrix} -r\cdot\tan\frac{\theta}{2}\cdot\sin\theta\cdot\sin\phi & r\cdot\tan\frac{\theta}{2}\cdot\sin\theta\cdot\cos\phi & 0 \end{bmatrix}^T \tag{4.18}$$

The time derivative of the fundamental matrix $\dot{\mathbf{F}}$ only depends on a non-constant term in ${}^{w}_{o}\mathbf{t}$, i.e., $r\tan\left(\theta/2\right)$. Notice that it would not be zero since the virtual camera has a time-varying translation from the robot base coordinate. Then we can obtain:

$$\dot{\mathbf{F}} = {}^{c}_{w}\mathbf{R}^T \left({}^{c}_{w}\mathbf{R}\,{}^{w}_{o}\dot{\mathbf{t}}\right)^{\wedge} = \eta\,{}^{c}_{w}\mathbf{R}\left({}^{c}_{w}\mathbf{R}^T\,{}^{w}_{o}\mathbf{t}\right)^{\wedge} \tag{4.19}$$

where the multiplier η can be given as:

$$\begin{aligned} \eta &= \frac{d\left(r\tan\left(\theta/2\right)\right)/dt}{r\tan\left(\theta/2\right)} = \mathbf{J}_2\begin{bmatrix} \dot{r} & \dot{\theta} & \dot{\phi} \end{bmatrix}^T \\ &= \mathbf{J}_2\,\mathbf{J}_1^{-1}\,\mathbf{J}_1\begin{bmatrix} \dot{r} & \dot{\theta} & \dot{\phi} \end{bmatrix}^T = \mathbf{J}_2\,\mathbf{J}_1^{-1}\,\dot{\mathbf{p}}_e \end{aligned} \tag{4.20}$$

where $d(\cdot)/dt$ denotes the derivative of time, and $\mathbf{J}_2$ is a Jacobian matrix containing first-order partial derivatives that can be obtained as:

$$\mathbf{J}_2 = \begin{bmatrix} \tan\frac{\theta}{2} & \frac{r}{2\cos^2\left(\theta/2\right)} & 0 \end{bmatrix} \tag{4.21}$$

Thus, the last term in **Eq. 4.13** can be related to other two terms through the velocity of robot tip $\dot{\mathbf{p}}_e$, namely:

$$\begin{aligned} \mathbf{p}_e^T\,\dot{\mathbf{F}}\,\mathbf{p}_c &= \mathbf{p}_e^T\left[{}^{c}_{w}\mathbf{R}^T\left({}^{c}_{w}\mathbf{R}\,{}^{w}_{o}\mathbf{t}\right)^{\wedge}\right]\,\mathbf{p}_c\,\eta \\ &= k_1\,\eta = k_1\,\mathbf{J}_2\,\mathbf{J}_1^{-1}\,\dot{\mathbf{p}}_e \end{aligned} \tag{4.22}$$

where the scalar k_1 is obtained by multiplying the terms before the multiplier η given by **Eq. 4.20**.

4.4.2 Controller Design

Inverse mapping-based controller I:

By decomposing $\dot{\mathbf{p}}_e$ along the normal vector of epipolar plane and two orthogonal vectors inside the epipolar plane [25], we obtain:

$$\dot{\mathbf{p}}_e = \gamma\, {}^o\mathbf{h} + \beta_1 \frac{\mathbf{p}_e}{\|\mathbf{p}_e\|^2} + \beta_2 \frac{(\mathbf{p}_e \times {}^o\mathbf{h})}{\|\mathbf{p}_e \times {}^o\mathbf{h}\|^2} \tag{4.23}$$

where ${}^o\mathbf{h} = (\mathbf{F}\,\mathbf{p}_c)/\|\mathbf{F}\,\mathbf{p}_c\|$. $\|\cdot\|$ denotes the 2-norm of a vector. The decomposing coefficients along each vector are denoted by γ, β_1 and β_2. Substituting **Eq. 4.23** into the middle term of epipolar constraint in **Eq. 4.13**, and applying the orthogonal property in **Eq. 4.11**, we derive the coefficient γ as:

$$\gamma = -\frac{\left(\mathbf{F}^T\mathbf{p}_e\right)^T \dot{\mathbf{p}}_c}{\|\mathbf{F}\,\mathbf{p}_c\|} - \frac{\mathbf{p}_e^T \dot{\mathbf{F}}\,\mathbf{p}_c}{\|\mathbf{F}\,\mathbf{p}_c\|} \tag{4.24}$$

The first term of γ can be regarded as the projection of laser spot velocity $\dot{\mathbf{p}}_c$ onto the vector, $\left(\mathbf{F}^T\mathbf{p}_e\right)/\|\mathbf{F}\,\mathbf{p}_c\|$, which is orthogonal to the epipolar plane. We can decompose $\dot{\mathbf{p}}_c$ into a component along this vector and a remaining component:

$$\dot{\mathbf{p}}_c = \xi\, {}^c\mathbf{h} + \left(\mathbf{I}_{3\times3} - {}^c\mathbf{h}\,{}^c\mathbf{h}^T\right)\dot{\mathbf{p}}_c \tag{4.25}$$

where ${}^c\mathbf{h} = \left(\mathbf{F}^T\mathbf{p}_e\right)/\left\|\mathbf{F}^T\mathbf{p}_e\right\|$. By projecting the remaining component of $\dot{\mathbf{p}}_c$ onto the second and third term in **Eq. 4.23** [25], we can approximately get β_1 and β_2:

$$\beta_1 = \frac{\left\|\mathbf{F}^T\mathbf{p}_e\right\|}{\|\mathbf{F}\,\mathbf{p}_c\|}\,\mathbf{D}_1\,\dot{\mathbf{p}}_c\,, \quad \mathbf{D}_1 = \mathbf{p}_e^T\,{}^o_c\mathbf{R}\,\left(\mathbf{I}_3 - {}^c\mathbf{h}\,{}^c\mathbf{h}^T\right) \tag{4.26}$$

$$\beta_2 = \frac{\left\|\mathbf{F}^T\mathbf{p}_e\right\|}{\|\mathbf{F}\,\mathbf{p}_c\|}\,\mathbf{D}_2\,\dot{\mathbf{p}}_c\,, \quad \mathbf{D}_2 = (\mathbf{p}_e \times {}^o\mathbf{h})^T\,{}^o_c\mathbf{R}\,\left(\mathbf{I}_3 - {}^c\mathbf{h}\,{}^c\mathbf{h}^T\right) \tag{4.27}$$

As a consequence, we have:

$$\dot{\mathbf{p}}_e = \mathbf{A}\,\dot{\mathbf{p}}_c - {}^o\mathbf{h}\,\frac{\mathbf{p}_e^T\dot{\mathbf{F}}\,\mathbf{p}_c}{\|\mathbf{F}\,\mathbf{p}_c\|} = \mathbf{A}\,\dot{\mathbf{p}}_c - {}^o\mathbf{h}\frac{k_1\mathbf{J}_2\mathbf{J}_1^{-1}}{\|\mathbf{F}\,\mathbf{p}_c\|}\dot{\mathbf{p}}_e \tag{4.28}$$

where the matrix $\mathbf{A}$ can be given as:

$$\mathbf{A} = \left[\mathbf{p}_e\,\mathbf{D}_1 + (\mathbf{p}_e \times {}^o\mathbf{h})\,\mathbf{D}_2 - {}^o\mathbf{h}\,{}^c\mathbf{h}^T\right]\,\frac{\left\|\mathbf{F}^T\mathbf{p}_e\right\|}{\|\mathbf{F}\,\mathbf{p}_c\|} \tag{4.29}$$

Thus, we have obtained an inverse mapping from the image velocity $\dot{\tilde{\mathbf{p}}}$ to the robot tip velocity $\dot{\mathbf{p}}_e$:

$$\left(\mathbf{I}_{3\times3} + {}^{o}\mathbf{h}\,\frac{k_1\mathbf{J}_2\mathbf{J}_1^{-1}}{\|\mathbf{F}\,\mathbf{p}_c\|}\right)\dot{\mathbf{p}}_e = \mathbf{A}\,\dot{\mathbf{p}}_c = \mathbf{A}\,\mathbf{K}^{-1}\,\dot{\tilde{\mathbf{p}}} \tag{4.30}$$

The robot tip velocity $\dot{\mathbf{p}}_e$ (**Eq. 4.9**), which is determined by the velocity of configuration parameters, $[\dot{r} \quad \dot{\theta} \quad \dot{\phi}]^T$, can be related to the variation velocity of chamber lengths, $\dot{\mathbf{L}}$, using a Jacobian matrix $\mathbf{J}_{act}$ obtained based on the CC model [47]:

$$\begin{bmatrix} \dot{r} \\ \dot{\theta} \\ \dot{\phi} \end{bmatrix} = \mathbf{J}_{act}\begin{bmatrix} \dot{L}_1 \\ \dot{L}_2 \\ \dot{L}_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial r}{\partial \mathbf{L}} \\ \frac{\partial \theta}{\partial \mathbf{L}} \\ \frac{\partial \phi}{\partial \mathbf{L}} \end{bmatrix}\begin{bmatrix} \dot{L}_1 \\ \dot{L}_2 \\ \dot{L}_3 \end{bmatrix} \tag{4.31}$$

$$\frac{\partial r}{\partial \mathbf{L}} = \frac{L_0}{2c_1}\begin{bmatrix} \sqrt{c_1} - \frac{\sum_{i=1}^{3} L_i}{2\sqrt{c_1}}\,(2L_1 - L_2 - L_3) \\ \sqrt{c_1} - \frac{\sum_{i=1}^{3} L_i}{2\sqrt{c_1}}\,(2L_2 - L_1 - L_3) \\ \sqrt{c_1} - \frac{\sum_{i=1}^{3} L_i}{2\sqrt{c_1}}\,(2L_3 - L_1 - L_2) \end{bmatrix}^T \tag{4.32}$$

where, $c_1 = \sum_{i=1}^{3} L_i^2 - L_1 \cdot L_2 - L_2 \cdot L_3 - L_1 \cdot L_3$.

$$\frac{\partial \theta}{\partial \mathbf{L}} = \frac{1}{r}\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} - \frac{1}{r^2}\frac{\partial r}{\partial \mathbf{L}} \tag{4.33}$$

$$\frac{\partial \phi}{\partial \mathbf{L}} = c_2 \cdot \begin{bmatrix} \frac{-2\sqrt{3}}{3\,(L_2 - L_3)} \\ \frac{1}{2\sqrt{3}\,(L_1 - L_3)\,(L_2 - L_3)^2} \\ \frac{1}{2\sqrt{3}\,(L_2 - L_1)\,(L_2 - L_3)^2} \end{bmatrix}^T \tag{4.34}$$

where, $c_2 = \dfrac{3(L_2 - L_3)^2}{3(L_2 - L_3)^2 + \sqrt{3}\,(L_3 + L_2 - 2L_1)}$.

The inverse transition mapping from image space to actuation space is accordingly obtained as:

$$\dot{\mathbf{L}} = \mathbf{B}\,\dot{\tilde{\mathbf{p}}} \tag{4.35}$$

where,

$$\mathbf{B} = \left(\mathbf{I}_{3\times3}\mathbf{J}_1\mathbf{J}_{act} + {}^{o}\mathbf{h}\frac{k_1\mathbf{J}_2\mathbf{J}_{act}}{\|\mathbf{F}\mathbf{p}_c\|}\right)^{-1}\mathbf{A}\mathbf{K}^{-1} \tag{4.36}$$

Thus, the actuation command of each chamber can be generated according to the desired laser spot displacement.

A proportional feedback control law can be constructed based on the time-varying error in pixels between the current and target positions of the laser spot:

$$\dot{\tilde{\mathbf{p}}} = -\lambda \left(\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^*\right) \tag{4.37}$$

where $\tilde{\mathbf{p}}^*$ is the desired laser spot position, and the proportional gain λ is a positive constant. Note that the control law in **Eq. 4.37** will ensure exponential convergence of the laser spot tracking errors in the camera view.

Forward mapping-based controller II:

Following the similar approach as in controller I, we can form a forward mapping between the image velocity $\dot{\tilde{\mathbf{p}}}$ and robot tip velocity $\dot{\mathbf{p}}_e$. The forward mapping will also be used for constructing a robot state observer (e.g., **Eq. 4.49**), as described in the next section. By decomposing $\dot{\mathbf{p}}_c$ along the normal vector of the epipolar plane and two orthogonal vectors inside the epipolar plane, we obtain:

$$\dot{\mathbf{p}}_c = \delta\, {}^c\mathbf{h} + \varepsilon_1 \frac{\mathbf{p}_c}{\|\mathbf{p}_c\|^2} + \varepsilon_2 \frac{(\mathbf{p}_c \times {}^c\mathbf{h})}{\|\mathbf{p}_c \times {}^c\mathbf{h}\|^2} \tag{4.38}$$

where the coefficient δ can be derived by substituting **Eq. 4.38** into **Eq. 4.13**, given as follows:

$$\delta = -\frac{(\mathbf{F}\,\mathbf{p}_c)^T\,\dot{\mathbf{p}}_e}{\|\mathbf{F}^T\,\mathbf{p}_e\|} - \frac{\mathbf{p}_e^T\,\dot{\mathbf{F}}\,\mathbf{p}_c}{\|\mathbf{F}^T\,\mathbf{p}_e\|} \tag{4.39}$$

Similar to **Eq. 4.25**, the robot tip velocity $\dot{\mathbf{p}}_e$ can be decomposed as a component along the normal vector of epipolar plane, and a remaining vector:

$$\dot{\mathbf{p}}_e = \zeta\, {}^o\mathbf{h} + \left(\mathbf{I}_{3\times 3} - {}^o\mathbf{h}\,{}^o\mathbf{h}^T\right)\,\dot{\mathbf{p}}_e \tag{4.40}$$

By projecting the remaining component of $\dot{\mathbf{p}}_e$ onto the second and third term in **Eq. 4.38**, we can accordingly get the coefficients ε_1 and ε_2:

$$\varepsilon_1 = \frac{\|\mathbf{F}\,\mathbf{p}_c\|}{\|\mathbf{F}^T\,\mathbf{p}_e\|}\,\mathbf{D}_1\,\dot{\mathbf{p}}_c\,, \quad \mathbf{D}_1 = \mathbf{p}_c^T\,{}^c_o\mathbf{R}\,\left(\mathbf{I}_3 - {}^o\mathbf{h}\,{}^o\mathbf{h}^T\right) \tag{4.41}$$

$$\varepsilon_2 = \frac{\|\mathbf{F}\,\mathbf{p}_c\|}{\|\mathbf{F}^T\,\mathbf{p}_e\|}\,\mathbf{D}_2\,\dot{\mathbf{p}}_c\,, \quad \mathbf{D}_2 = (\mathbf{p}_c \times {}^c\mathbf{h})^T\,{}^c_o\mathbf{R}\,\left(\mathbf{I}_3 - {}^o\mathbf{h}\,{}^o\mathbf{h}^T\right) \tag{4.42}$$

Substituting these coefficients into **Eq. 4.38**, the forward mapping can be formed as:

$$\dot{\mathbf{p}}_c = \mathbf{M}\,\dot{\mathbf{p}}_e \tag{4.43}$$

where the transition matrix $\mathbf{M}$ has similar form as **Eq. 4.29**:

$$\mathbf{M} = \left[\mathbf{p}_c\,\mathbf{D}_1 + (\mathbf{p}_c \times {}^c\mathbf{h})\,\mathbf{D}_2 - {}^c\mathbf{h}\,{}^o\mathbf{h}^T\right]\,\frac{\|\mathbf{F}\,\mathbf{p}_c\|}{\|\mathbf{F}^T\,\mathbf{p}_e\|} - {}^c\mathbf{h}\frac{k_1\,\mathbf{J}_2\,\mathbf{J}_1^{-1}}{\|\mathbf{F}^T\,\mathbf{p}_e\|} \tag{4.44}$$

Note that the third element of $\dot{\mathbf{p}}_c$ and $\dot{\tilde{\mathbf{p}}}$ are zeros. Thus they can be related by $\dot{\tilde{\mathbf{p}}} = \tilde{\mathbf{K}}\,\dot{\mathbf{p}}_c$, where $\tilde{\mathbf{K}}$ keeps the diagonal elements of camera intrinsic matrix $\mathbf{K}$ with other elements set as zero. Thus, the transition between the laser spot velocity in pixel coordinates and the chamber length variation rate can be obtained as:

$$\dot{\tilde{\mathbf{p}}} = \tilde{\mathbf{K}}\,\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act}\,\dot{\mathbf{L}} \tag{4.45}$$

A feedback controller is designed as:

$$\dot{\mathbf{L}} = -(\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act})^T\,\tilde{\mathbf{K}}^{-1}\,\lambda\,\Delta\tilde{\mathbf{p}} \tag{4.46}$$

where the proportional gain λ is a positive constant.

To prove the stability of the controller in **Eq. 4.46**, a Lyapunov function is designed as:

$$V = \frac{1}{2}\Delta\tilde{\mathbf{p}}^T\,\tilde{\mathbf{K}}^{-1}\,\tilde{\mathbf{K}}^{-1}\,\lambda\,\Delta\tilde{\mathbf{p}} \tag{4.47}$$

By differentiating **Eq. 4.47**, we get:

$$\begin{aligned}
\dot{V} &= \Delta\tilde{\mathbf{p}}^T\,\lambda\,\tilde{\mathbf{K}}^{-1}\,\tilde{\mathbf{K}}^{-1}\dot{\tilde{\mathbf{p}}} \\
&= \Delta\tilde{\mathbf{p}}^T\,\lambda\,\tilde{\mathbf{K}}^{-1}\,\tilde{\mathbf{K}}^{-1}\underbrace{\tilde{\mathbf{K}}\,\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act}\,\dot{\mathbf{L}}}_{\dot{\tilde{\mathbf{p}}}} \\
&= -\Delta\tilde{\mathbf{p}}^T\,\lambda\,\tilde{\mathbf{K}}^{-1}\,\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act}\,\underbrace{(\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act})^T\,\tilde{\mathbf{K}}^{-1}\,\lambda\,\Delta\tilde{\mathbf{p}}}_{\dot{\mathbf{L}}} \\
&= -\left[(\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act})^T\,\tilde{\mathbf{K}}^{-1}\,\lambda\,\Delta\tilde{\mathbf{p}}\right]^T\left[(\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act})^T\,\tilde{\mathbf{K}}^{-1}\,\lambda\,\Delta\tilde{\mathbf{p}}\right] \\
&= -\left\|(\mathbf{M}\,\mathbf{J}_1\,\mathbf{J}_{act})^T\,\tilde{\mathbf{K}}^{-1}\lambda\,\Delta\tilde{\mathbf{p}}\right\|^2 \leq 0
\end{aligned} \tag{4.48}$$

Therefore, the proposed feedback controller in **Eq. 4.46** is asymptotically convergent. The epipolar constraint is degenerate when the laser spot is located at the line between the robot frame centre and camera frame centre. However, this case is avoided by the restricted bending workspace of the robot.

4.4.3 Robot State Observer

The soft robot configurations are determined by the lengths of fluid chambers, which thus can be used to represent the robot state. The accurate estimation of chamber lengths is significant on the calculation of transition mappings, namely **Eq. 4.35** and **Eq. 4.45**, thus the control commands. However, it is always difficult to measure the chamber lengths directly. To this end, the chamber lengths are estimated by combining the information from the model prediction and that from the image feedback. To achieve a desired laser spot displacement, the required change of chamber lengths can be approximately calculated using the proposed controllers, **Eqs. 4.35** and **4.37**, or **Eq. 4.46**. Then the laser spot position at the next time step can be predicted based on the forward mapping as in **Eq. 4.45**.

As given in **Eq. 4.49** and **Eq. 4.50**, a Kalman filter is constructed as the state observer combining the predicted and measured laser spot positions. The transition matrixes in the forward mapping (**Eq. 4.45**) are approximated as the observation matrix $\mathbf{H}_k$, $\mathbf{H}_k = \tilde{\mathbf{K}}\mathbf{M}\mathbf{J}_1\mathbf{J}_{act}$, which is calculated based on robot state at time step k.

$$\begin{cases} \mathbf{L}(k) = \mathbf{L}(k-1) + \mathbf{B}_k \Delta\tilde{\mathbf{p}}^* + \mathbf{w}_k \\ \tilde{\mathbf{p}}(k) \;\; = \tilde{\mathbf{p}}(k-1) + \mathbf{H}_k \left[\mathbf{L}(k) - \mathbf{L}(k-1)\right] + \mathbf{v}_k \end{cases} \tag{4.49}$$

The combined robot state, $\hat{\mathbf{L}}(k)$, is given as:

$$\begin{aligned} \hat{\mathbf{L}}(k) = \mathbf{L}(k-1) &+ \mathbf{B}_k \Delta\tilde{\mathbf{p}}^* \\ &+ \mathbf{G}_k \left[\tilde{\mathbf{p}}_{mea}(k) - \tilde{\mathbf{p}}(k-1) - \mathbf{H}_k \mathbf{B}_k \Delta\tilde{\mathbf{p}}^*\right] \end{aligned} \tag{4.50}$$

where $\tilde{\mathbf{p}}_{mea}(k)$ is the measured laser spot position at time step k. The Kalman gain is denoted by $\mathbf{G}_k$, which can be updated based on the same rule as **Eq. 4.4**. Increasing the covariance of process noise will reduce the weight of model prediction. Increasing the covariance of measurement noise will enhance the smoothing of noise but decrease the rate of convergence.

The architecture of the proposed control method is shown in **Fig. 4.3**, which combines the components of laser spot detection, the epipolar-geometry based controller, and the robot state observer. The control target defined in the camera view can be a reference trajectory or even a path drawn by an operator using a mouse cursor. As the formulation of the proposed controllers does not include any information on the laser projection depth, they will allow stable laser spot visual servoing even on an irregular projection surface.

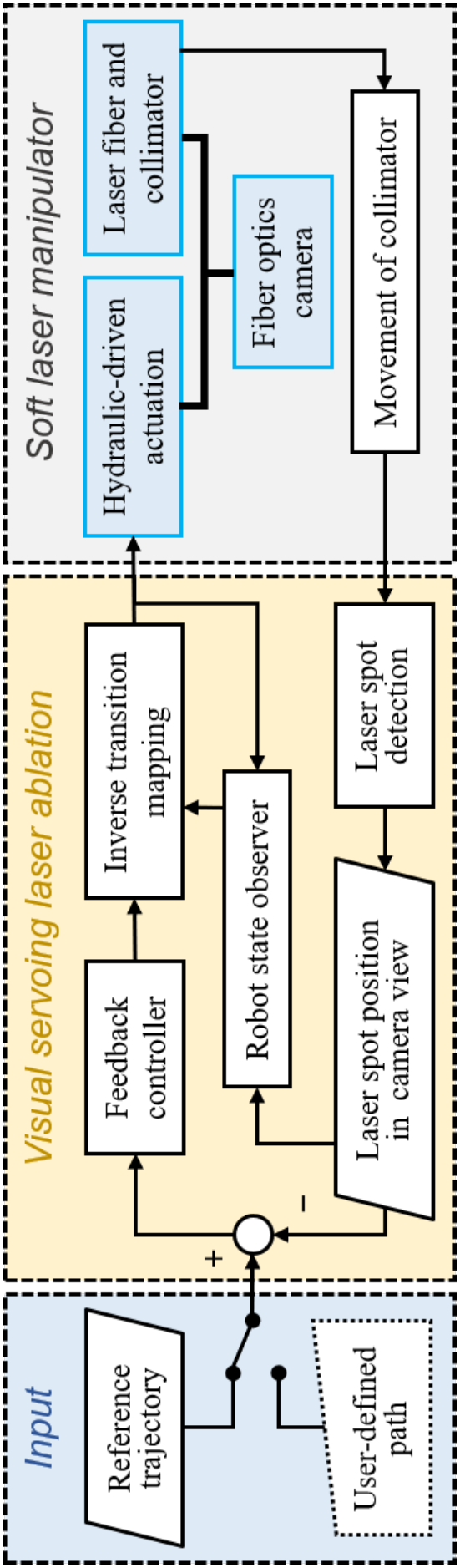


Fig. 4.3: Control diagram of the proposed laser spot visual servoing. The inverse transition mapping is estimated with constant curvature and epipolar geometry model. The robot states, namely lengths of chambers, are observed through Kalman filter fusing the actuation command with the image feedback.

4.4.4 Simulation and Results

To validate the performance of proposed controllers, simulation studies are firstly conducted. CC model is employed to represent the kinematics of soft manipulator, which projects a laser spot $\mathbf{s}$ on a planar surface. Thus a corresponding image $\tilde{\mathbf{p}}$ appears in the camera view:

$$\tilde{\mathbf{p}}' = \mathbf{K}\,\mathbf{T}_1\,{}^c_o\mathbf{T}\begin{bmatrix}\mathbf{s} & 1\end{bmatrix}^T, \quad \tilde{\mathbf{p}} = \tilde{\mathbf{p}}'/\tilde{\mathbf{p}}'_z \tag{4.51}$$

where $\mathbf{T}_1 = \begin{bmatrix}\mathbf{I}_{3\times3} & \mathbf{0}_{1\times3}\end{bmatrix}$. The laser spot position in the robot coordinate O_o can be calculated as:

$$\begin{cases} \mathbf{s}_z = d - r\,\tan\left(\theta/2\right) \\ \mathbf{s}_x = \mathbf{s}_z\,\tan\theta\,\cos\phi \\ \mathbf{s}_y = \mathbf{s}_z\,\tan\theta\,\sin\phi \end{cases} \tag{4.52}$$

where d is the depth of laser spot with respect to the robot base coordinate O_w.

The robot and camera parameters are set to match the experimental platform. The natural length of the hydraulic chambers is 9 mm, and the radial distance between the chamber and the central channel is 3 mm. The camera is housed in parallel to the long axis of the soft manipulator, with an axial distance of 7 mm. Referring to the camera pixel size and focal length, the camera intrinsic matrix is defined as:

$$\mathbf{K} = \begin{bmatrix} 1{,}000 & 0 & 640 \\ 0 & 1{,}000 & 480 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.53}$$

The soft manipulator with the laser collimator is placed 20 mm above a planar projection surface. Three simulation studies are conducted as below:

Controller I *vs* Controller II

The performance of Controller I and II are compared via a target-tracking task. As shown in **Fig. 4.4a**, both two controllers enabled accurate tracing of a goal point in the camera view. **Fig. 4.4b** shows the tracking errors of two controllers in u and v directions, respectively. Using the same proportional gain, Controller I converges faster than controller II due to the smaller step size of controller II, which can be increased by using a higher gain. Both controllers achieved equal convergent speed along two directions. In the next simulation, the camera was intentionally rotated about its principal axis by 20°, which was unknown to the controllers. This means that the camera view is rotated with respect to the laser manipulator,

resulting in a rotated motion vector of the laser spot in the camera view. **Fig. 4.5** shows the corresponding tracking results. Controller II achieved a straighter path than controller I, while ensuring a similar convergence rate. This indicates that controller II is more robust to the uncertainties in the camera parameters. In addition, controller II does not involve frequent calculation of matrix inversion. Therefore, Controller II will be utilized in the following simulations and experiments.

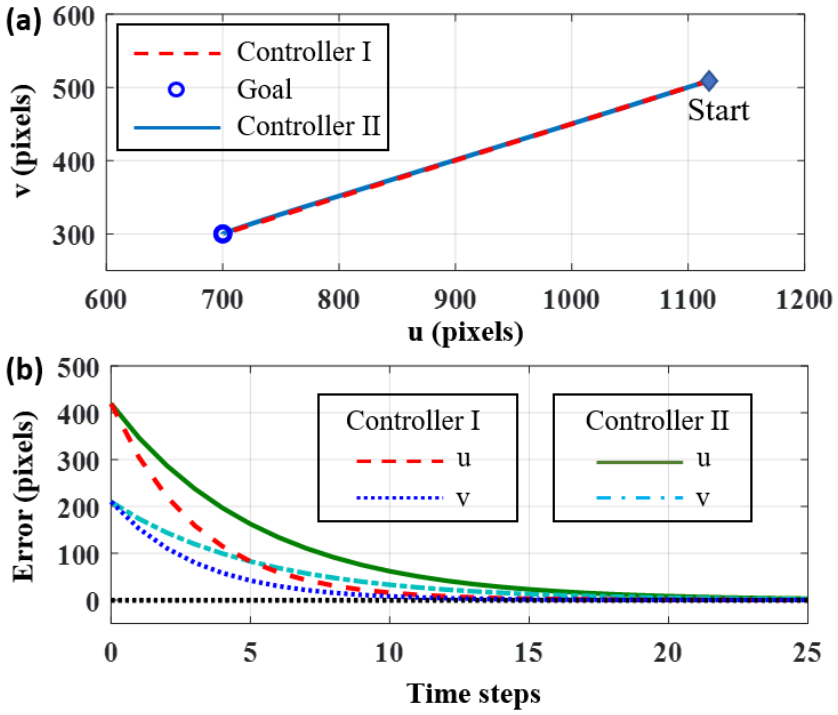


Fig. 4.4: Simulation of target tracking using Controller I and Controller II. **(a)** Tracking trajectories; **(b)** Tracking errors.

With *vs* Without observer

Fig. 4.6 shows the simulation results on following a spiral path with and without using the robot state observer. This result demonstrates that the state observer can compensate for the Gaussian noise and model uncertainty, by taking the measured laser spot position into account. This would be advantageous for the hydraulic actuation system, as the desired change of chamber lengths may not be fully completed due to the possible air bubbles existing in the hydraulic tubes.

Parametric Analysis in Simulation

To implement the proposed visual servoing controller, the camera intrinsic and extrinsic parameters, as well as the robot geometry parameters, need to be initialized. Obtaining the exact value of those parameters is challenging, and as such, the controller should be robust

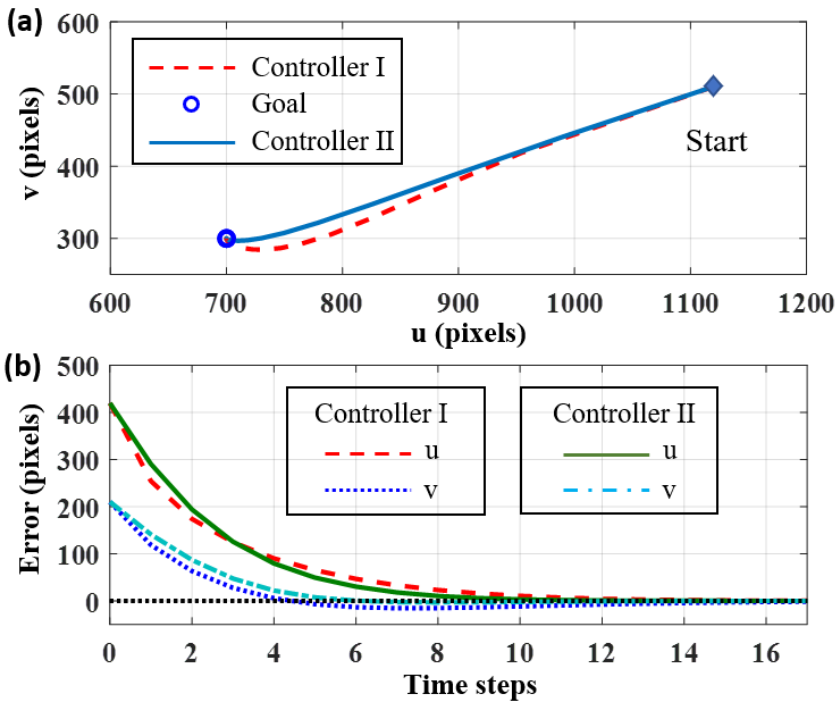


Fig. 4.5: Simulation of target tracking using Controller I and Controller II. The camera is intentionally rotated about its principal axis by 20°, which is unknown to the controllers. **(a)** Tracking trajectories; **(b)** Tracking errors.

against error that arises from them. To this end, simulations are first conducted to study the robustness of the proposed controller by introducing additional noise to the modeling parameters. The simulated control task is to follow a circle (center: (400, 400), radius: 50 pixels) in the camera view. With exact robot and camera parameters, the controller achieved a root-mean-square (RMS) tracking error of 0.08 pixels (∼1.2 μm). All simulations were conducted to ensure the same path following speed, one cycle in 1,000 steps.

1) Camera intrinsic parameters: They are found to have a limited influence on the control accuracy. For example, the second diagonal element of intrinsic matrix **K** is multiplied by 2, and the principle coordinate along the u axis is divided by 1.2. The controller maintains a stable path following with an RMSE of 0.31 pixel (∼4.5 μm).

2) Camera extrinsic parameters: By introducing 20% error (∼2 mm) to the translation parameters, and rotating the camera 30 degrees about the x, y, and z axes respectively, corresponding RMSEs of 1.51, 0.22, and 0.31 pixels are obtained, i.e., around 21.9, 3.2 and 4.5 μm. It infers that the proposed controller can be applied in endoscopic applications, of which the assembly accuracy of the camera is around 5 deg and 1 mm [185].

3) Robot geometry parameters: The controller is found to be robust to a relatively large

change of the robot structural parameters. A 20% error is added to the natural chamber length, as well as the radial distance between the robot central channel and fluid chambers. The RMSE of following a circle is only about 0.37 pixels (∼5.4 μm).

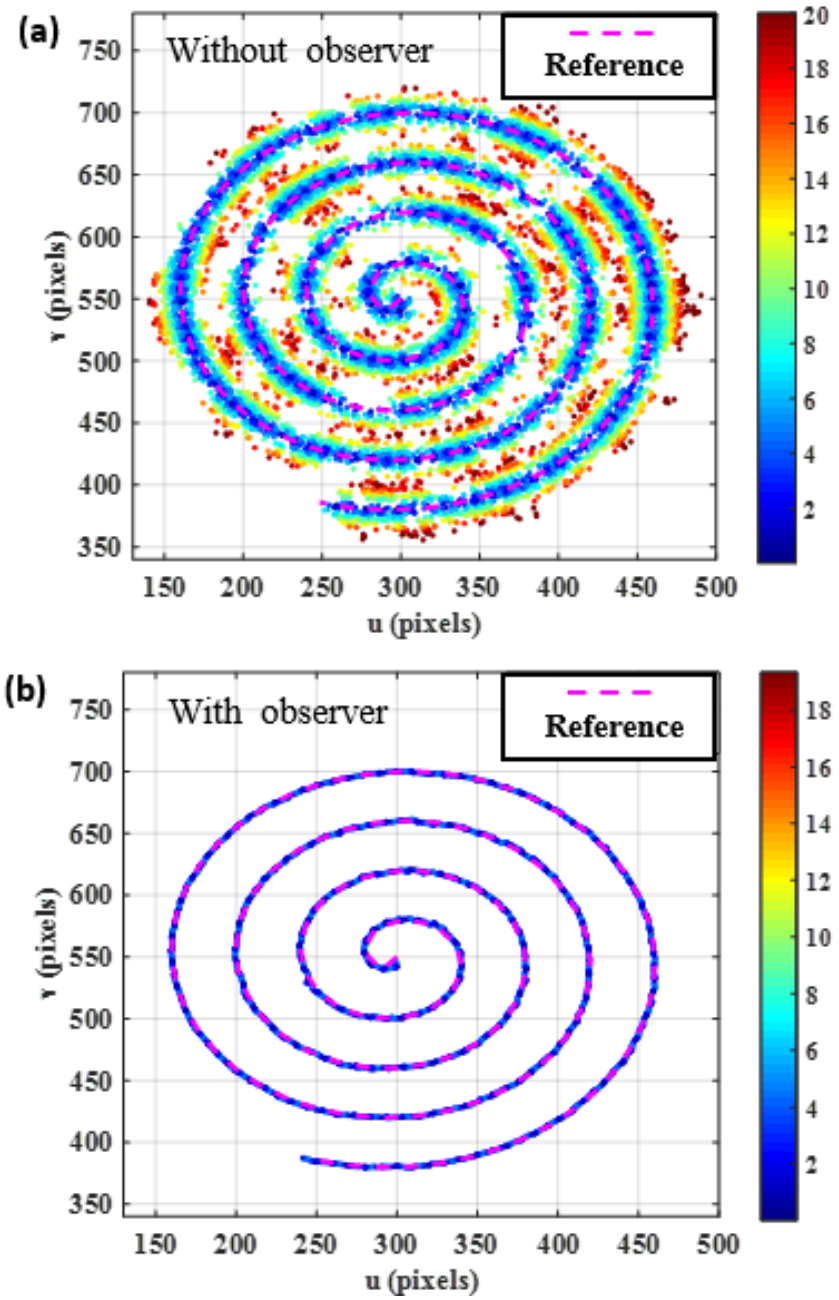


Fig. 4.6: Simulation of a spiral path following. Gaussian random noise is induced to the length estimates of chambers at each time step. Actual footprint of laser spot compared with the reference path are plotted. The larger deviation of laser spot from reference path, the warmer color of its footprint. **(a)** Without robot state observer; **(b)** With observer.

4.4.5 Experiments and Results

In this section, experiments are conducted to evaluate the performance of the proposed controller on an MR-safe soft robotic laser manipulator. A fiber optics camera has been integrated to provide image feedback of the laser spot. Point tracking and path following tasks are performed. The experiment setup and results are introduced as follows:

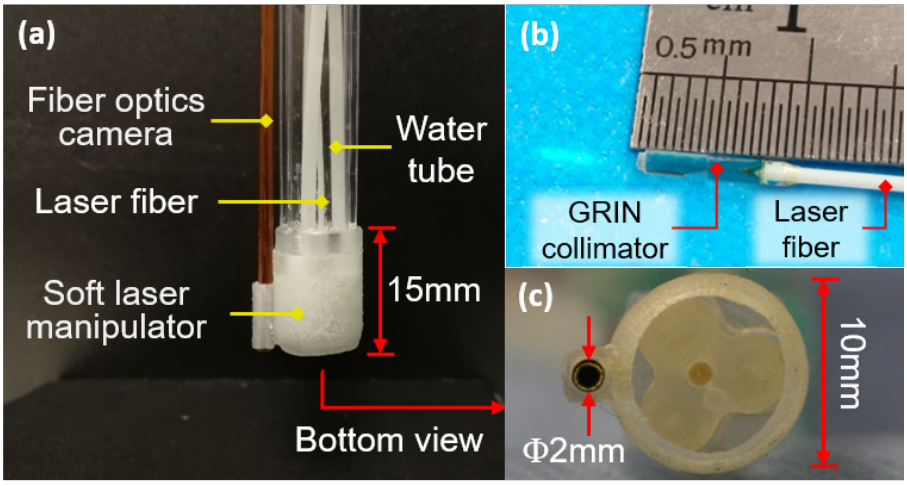


Fig. 4.7: **(a)** Soft laser manipulator with a fiberscope equipped to provide image feedback. **(b)** GRIN-lens collimated laser fiber. **(c)** Bottom view of the soft manipulator.

Experimental Platform

Experimental validation has been conducted using the platform, as shown in **Fig. 4.7**. The laser steering module comprises three cylindrical elastomeric chambers separated by 120°, which are inflated and deflated by hydraulic cylinders. The soft manipulator is 3D printed with AgilusClear30^{TM} material. Each cylinder is connected to a stepper motor, providing precise control of the cylinder volume. The hydraulic transmission is achieved through 10-meter-long pipelines (OD=4 mm). By pumping in and out water, each elastic chamber can be elongated or shortened individually. Regulating the inflation volumes of three chambers offers omni-directional bending (+/- 25°) of the optical laser fiber housed in the center channel. The fiber end is assembled with a collimator to produce a small laser spot projection at a longer distance compared to a free-end fiber. The collimator selected is a gradient-index (GRIN) rod lens with parameters: 0.23 pitch, face angle = 8°, outer diameter = 1.0 mm, length = 3 mm, anti-reflection coating at 1,550 nm. It is packaged with the optical fiber ferrule using a glass tube. The distance between the fiber end and the lens' front end is near zero. The size of the laser spot on the target is about 1mm, which could be further reduced by optimizing the lens parameters, e.g., pitch value, length. The laser manipulator

provides a scanning area of 20×20 mm^2 at a projection distance of ∼20 mm. The executable space is ensured by imposing maximum limits on each motor. With this setting, laser energy can be projected on the tissue surface in a small spot, avoiding potential overheating caused by fiber tip sticking.

An additional outer cover is proposed to capsule the soft fluid chambers. Steering of the laser lens can be performed inside the cover, thus avoiding external disturbances, such as tissue contact in the confined ONP. This design enhances the robustness of laser beam control by isolating the soft manipulator from its surroundings, therefore allowing laser ablation to be performed with higher accuracy. A fiber optics camera with 30,000 pixels in a diameter of 2 mm is utilized to provide visual feedback. Images are captured through a bundle of optical fibers (10-meters long) and coupled to a high-resolution RGB camera of 1,280×960 pixels. The control algorithm is implemented in Matlab (Intel Core i7-6700HQ CPU@3.0GHz with 16-GB RAM).

Point-to-point Tracking

A point tracking experiment has been conducted to validate the performance of the proposed controllers. The homogenous transformation between robot and camera coordinate is initialized based on robot design parameters. The rotational angle of the camera view is estimated from laser spot movement in the camera image while inflating one soft chamber. As shown in **Fig. 4.8**, the robot is controlled to track four points one by one. Once the laser spot enters a region within 3 pixels (∼105 μm) of a target point, the desired target will be switched to the next one. Similar to the simulation result (**Fig. 4.4**), both controllers achieved accurate tracking of the four targets. However, the laser spot trajectory with Controller II is straighter than Controller I, particularly from target 2 to 3, which implies that Controller II has a more precise approximation of the inverse transition mapping.

Path Following

Laser spot path following tasks have been conducted to evaluate the accuracy of the proposed visual servoing controller (II). In the 1,280×960 pixels image frame, a reference path is predefined as a list of target positions to be tracked by the laser spot. The tracking error in the unit of pixels is defined as the shortest distance between the current position of the laser spot and the desired trajectory.

Two successive cycles are selected from the continuous path following journey, with the actual trajectory of the laser spot and the corresponding tracking errors presented in **Fig. 4.9a** and **Fig. 4.9b**, respectively. Over the 1^{st} cycle, the RMS tracking error is 2.37

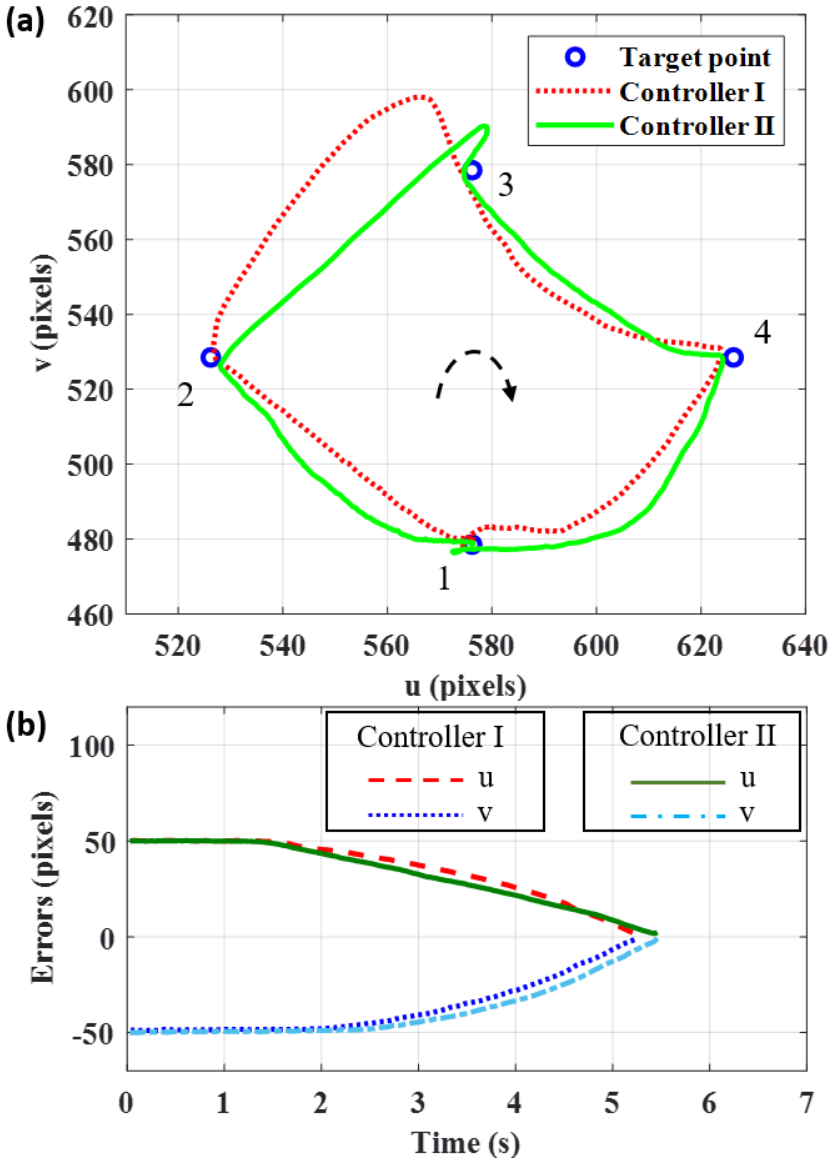


Fig. 4.8: Performances on tracking multiple target points using Controller I compared with using Controller II. The robot is controlled to aim the laser spot at four target points one by one (Targets 1 to 4). **(a)** Tracking trajectories; **(b)** Tracking errors from Target 1 to Target 2.

pixels (∼83 μm), and the maximum tracking error is 7.47 pixels (∼261.5 μm). The accurate path following is maintained in the 2^{nd} cycle, during which the RMSE and maximum error are 1.82 pixels (∼63.7 μm) and 5.92 pixels (∼207.2 μm), respectively. The projection distance is about 20 mm (**Fig. 4.9c**), resulting in a projected "Mickey Mouse" boundary with a width of about 10.5 mm (300 pixels) on the target surface. **Fig. 4.9d** depicts the current laser spot image along with its trajectory throughout the tracking journey. Thus, the overall accuracy of laser spot targeting can be concluded with a mean tracking error $<$105 μm (3 pixels). The maximum path following speed is tested as about 58 pixels/s (∼2 mm/s), corresponding to an RMSE of 6.13 pixels (∼214.6 μm).

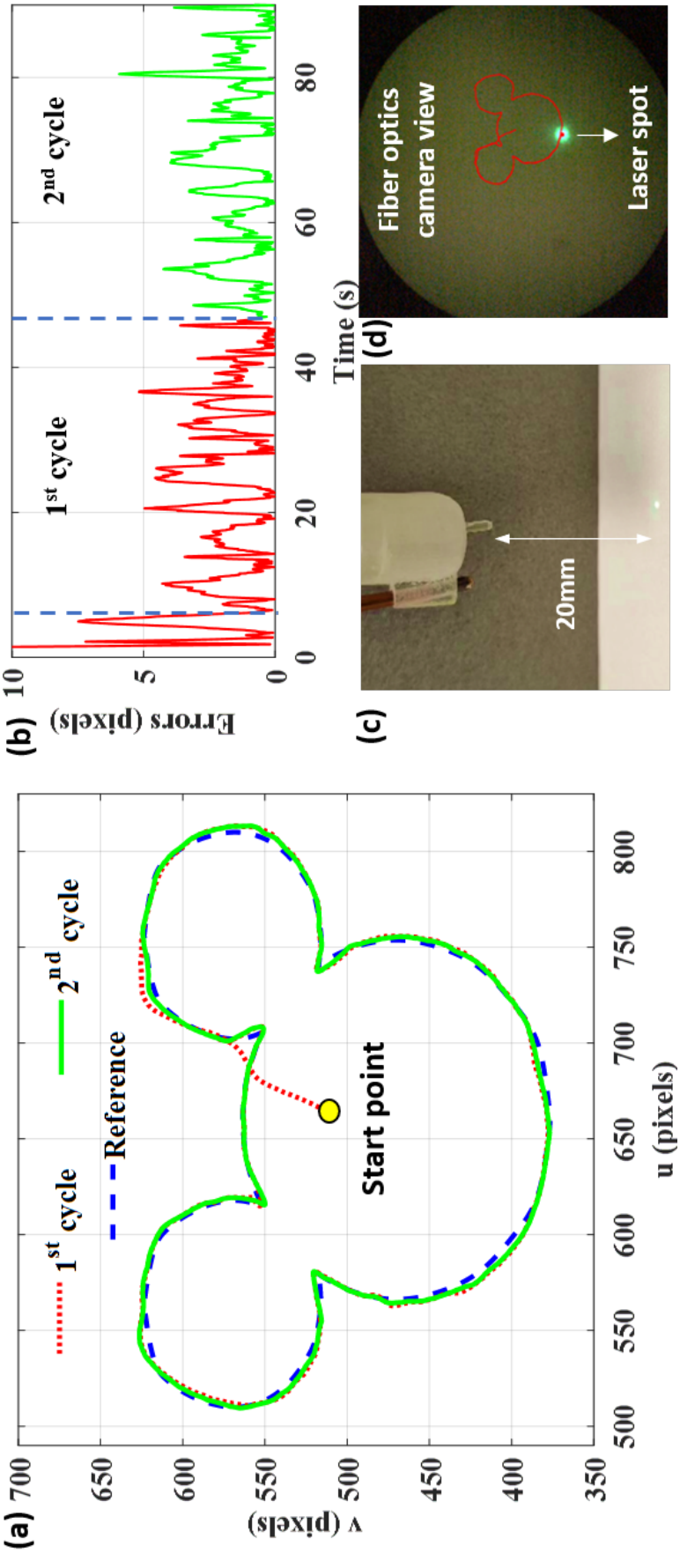


Fig. 4.9: Performance of laser spot visual seroving on following a predefined "Mickey Mouse" path. **(a)** Actual tracking trajectory of the laser spot in the image frame. **(b)** Tracking errors during the 90-second servoing period. **(c)** Setup of the soft laser manipulator. **(d)** Current image of laser spot superimposed with its footprint.

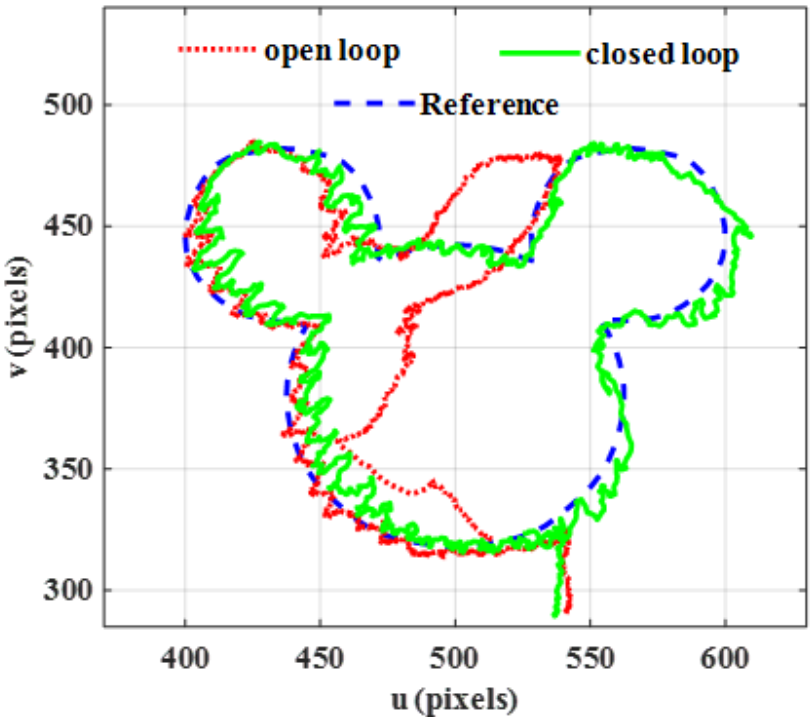


Fig. 4.10: Performances on path following using respectively open-loop and closed-loop control. Water leakage is intentionally caused by puncturing one water pipe to a soft chamber, thereby actuation disturbance. The closed-loop controller can still compensate for it and maintain precise path following.

Open-loop vs closed-loop: To validate the robustness of the closed-loop controller, one water pipe connected to a soft chamber was intentionally punctured, inducing water leakage during the robot operation. **Fig. 4.10** shows the tracking results using open-loop control compared with closed-loop control. The laser spot cannot follow the desired path with open-loop actuation. However, the closed-loop controller can compensate for the actuation disturbance, thus ensuring accurate path following.

Repeatability: **Fig. 4.11** shows the repeatability of Controller II when performing path following. The controller was set to track a "Mickey Mouse" path with 70 loops. The RMS tracking error remains small within 4 pixels ($\sim$140 μm) while the maximum error maintains about 10 pixels ($\sim$350 μm). It demonstrates that the laser manipulation system, along with the proposed visual servo controller, is stable and reliable.

With vs without state observer: To validate the effect of state observer, a path following test has also been conducted with water leakage intentionally introduced by puncturing one water pipe to a soft chamber. **Fig. 4.12** shows the tracking results using state observer compared with omitting state observer. It demonstrates that the state observer reduced the

RMS tracking error from 4.26 pixels (∼149.1 μm) to 2.96 pixels (∼103.6 μm).

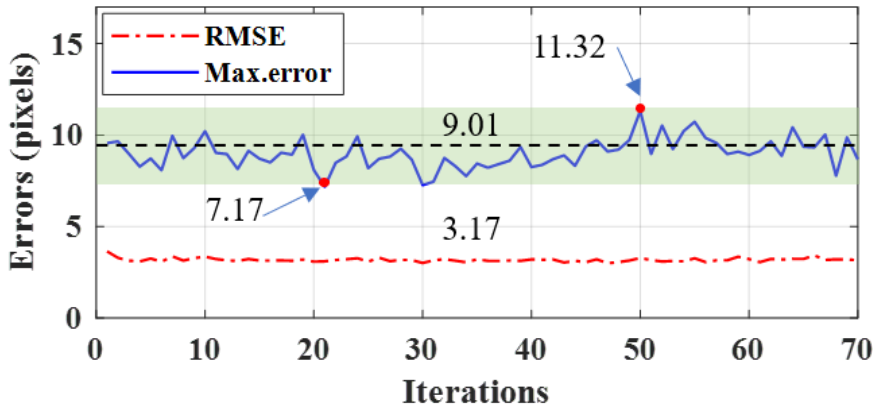


Fig. 4.11: Repeatability of the control performances in following the "Mickey Mouse" path as in **Fig. 4.9** (70 cycles in total).

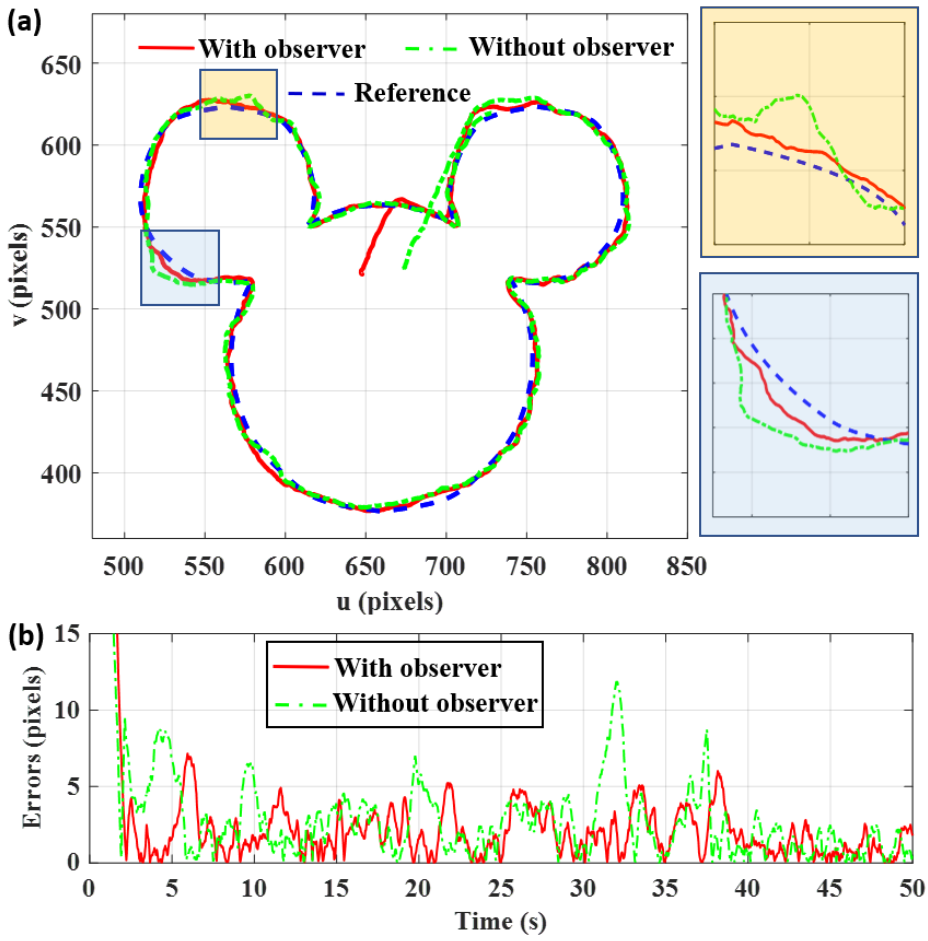


Fig. 4.12: Performance on path following with and without state observer. Water leakage is intentionally introduced by puncturing one water pipe to a soft chamber. **(a)** Tracking trajectories; **(b)** Tracking errors.

Hand-drawing Following

The performance of visual servoing Controller II has been validated via a hand-drawing following task. The positions of the mouse cursor in the camera view are detected as the desired target of the laser spot. The drawing speed is estimated to be about 10 pixels/second. The laser spot is controlled to instantly follow the mouse cursor manipulated by an operator. As shown in **Fig. 4.13**, the laser spot can accurately follow the mouse cursor, with an RMS tracking error of 3.97 pixels (∼180.5 μm) and a maximum error of 12.26 pixels (∼557.3 μm).

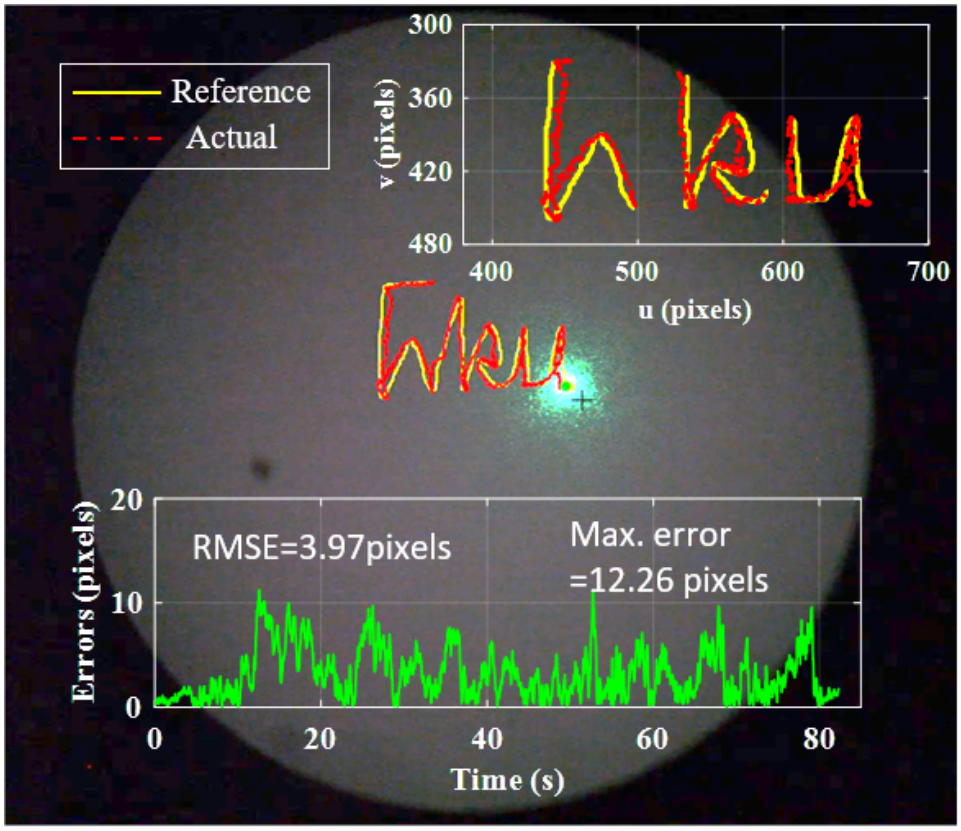


Fig. 4.13: User-defined path "drawn" by a mouse cursor. Both the laser spot footprint and the drawn path are superimposed in the camera view, so as to evaluate the tracking errors (RMSE and Max.).

4.5 Learning-based Laser Steering Control

In this section, a learning-based visual servoing controller is implemented to achieve automatic targeting of the laser spot that could be projected along any trajectory that is prescribed *in-situ* on the endoscopic view. The prerequisite is to form an adaptive kinematics model approximating the nonlinear inverse mapping from image space to

robot actuation space. Conventional kinematics methods require accurate characterization of robot structural parameters and actuation response, which is a well-known challenge for the intrinsic nonlinear hyper-elastic property of elastomeric materials. In contrast, data-driven methods [167] such as machine learning approaches could accommodate modeling uncertainties from the actual operation data, as well as eliminating the need for hand-eye calibration.

4.5.1 Inverse Transition Mapping

Similar to **Section 3.3.1**, the actuation input (in equilibrium) at time step k is denoted by $\boldsymbol{\alpha}(k) \in \mathbb{U}^3$, where $\mathbb{U}^3$ represents the 3D actuation space. The configuration of the collimator, including end position and orientation normal, is denoted by a collective variable $\boldsymbol{\theta}(k)$, which is related to the robot configuration. During quasi-static movement, there exists a unique mapping between robot configuration space and actuation space. Therefore, the actuation command can be used to reflect the robot configuration, which is nearly unmeasurable for soft robots. The task space is defined in the 2D camera view (**Fig. 4.7**), with the laser spot position denoted as $\tilde{\mathbf{p}} \in \mathbb{R}^2$. Direct inversion of the forward transition mapping is not possible because the manipulator is redundant, and the solution set of all possible solutions do not forms a convex set. Driven by the insight that the solution set of the differential IK problem form a convex set locally, the inverse transition model from task space to actuation space is represented by spatially localizing the actuation variable, i.e.:

$$\Delta\boldsymbol{\alpha}(k) = g\left(\boldsymbol{\alpha}(k), \Delta\tilde{\mathbf{p}}(\mathbf{k})\right) \tag{4.54}$$

Neural network is utilized for training the inverse mapping rather than GPR. This is because the focus is on accuracy and consistency, which are assured by a large number of data samples to characterize the whole workspace. The computation cost of GPR model prediction is higher in the case of a large training data size. The laser steering module is also protected from external disturbances by an extra outer constraint, thus releasing the need for efficient online learning of the controller. The actuation command at previous time step $\boldsymbol{\alpha}(k)$ is incorporated as an input for generating the actuation at next time step $\boldsymbol{\alpha}(k+1) = \boldsymbol{\alpha}(k) + \Delta\boldsymbol{\alpha}(k)$. This achieves localizing $\boldsymbol{\alpha}$ from the sample data instead of the learning architecture. Localization can be formed within a bounded range of $\Delta\boldsymbol{\alpha}(k)$. Note that it will inherently bound the model output spatially too.

4.5.2 Learning-based Controller

The learning-based controller is initialized by training a neural network with time-varying 2-D laser spot displacements, $(\Delta u, \Delta v)$, and soft actuation status, (p_i, i=1,2,3), as its input, and the corresponding actuation changes, (Δp_i, i=1,2,3), as output (**Fig. 4.14a**). The data was collected by commanding the laser manipulator to explore the workspace with a pre-defined scanning pattern (e.g., spiral). Relative movements rather than absolute positions of the laser spot are utilized to train a model so that it is robust to variation of viewing depth, which would scale the relative movements but completely shift the absolute positions. The actuation status represents the absolute inflation volumes (0∼0.004 mL) of three soft chambers and determines the robot configuration. It is also incorporated as an input to facilitate spatially localizing the inverse kinematics problem, thereby ensuring the convexity of the solution set. This is driven by the insight that the differential IK of a nonlinear system can be linearized locally.

The green laser spot with monochromatic color and high intensity is recognized by thresholding the captured RGB images. Five thousand data samples were collected and divided randomly into training/validation/testing sets in the ratio of 80:10:10, respectively. A multilayer perceptron with two hidden layers and one output layer, using tan-sigmoid and linear transfer function, respectively, was employed for learning. Levenberg-Marquardt backpropagation was utilized to train the network parameters. The sizes of hidden layers were selected (20 nodes in each layer) to fulfill a high goodness-of-fit with the correlation coefficient >0.99. The network was trained using the "*fitnet*" function in Matlab with Intel® Core™ i7-8750H CPU @2.20GHz. During each control loop, the desire laser spot movement, as well as the actuation status, will be inputted into the trained model. The trained NN model will always output a valid set of liquid volumes pumping in/out of the three soft chambers, which is inherently bounded by the training data. The trained NN model enables output prediction at high frequency (>100 Hz) for real-time robot control. Laser spot positions in camera view can be extracted online and fed back to close the control loop of soft actuation (**Fig. 4.14b**). Image acquisition and laser spot detection were implemented in a process written by C++, which was running in parallel with the main control loop. The feedback controller compensates for the tracking errors in real-time (>20 Hz) and further improves the control accuracy. A feedback delay of ∼30 ms was introduced by the camera image acquisition and processing, as well as communication with the main control loop. However, it is acceptable due to the limited bandwidth of soft actuation, resulting in a

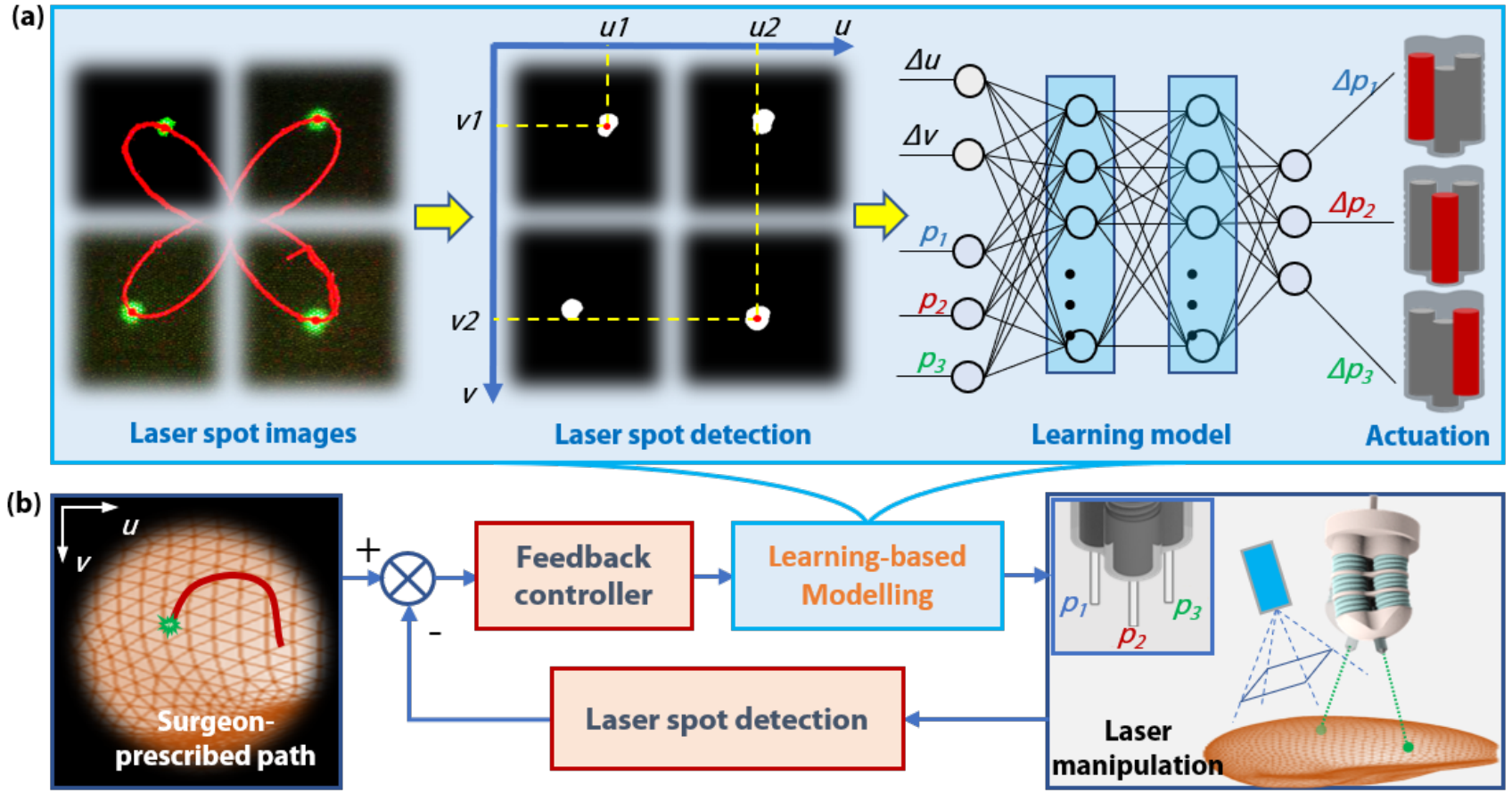


Fig. 4.14: Architecture of the proposed learning-based controller. **(a)** Overview of the learning-based model. The laser spot position in the camera view is varied by the laser manipulator. 2-D displacements of the laser spot, along with the soft actuation states, act as inputs of a multilayer perceptron which map to the actuation changes of each chamber. **(b)** The feedback control loop. With this learned inverse mapping, the proposed eye-to-hand visual servoing controller allows laser steering along paths prescribed *in-situ* by the operator.

maximum laser steering speed of 2 mm/s.

4.5.3 Experiments of Learning-based Laser Steering Control

Laser steering control has been tested on a soft robotic laser manipulator using a series of exercises that map to pre-defined patterns. The laser manipulator was configured to scan the laser over an area of 15×15 mm^2 on a projection plane located 15-mm below the robot tip, with a maximum bending angle of 26º. The scanning area is sufficient for coverage of an average oral tumor, which was reported to have a mean length of 4.04 mm and width of 3.08 mm when identified using MRI [189]. Precise steering of the laser spot within a small and local area using the hydraulic-driven soft laser manipulator could raise challenges due to the actuation nonlinearity and fabrication uncertainty. To this end, a learning-based modeling method was implemented to characterize the robot kinematics, which establishes the mapping from laser spot movements on the projection plane to robot actuation. The laser spot position can be computed from robot tip orientation which was measured using an EM-tracking system (**Fig. 4.15a**), or by directly extracting from camera images (**Fig. 4.15b**). Based on the trained model, a feed-forward controller was constructed to predict the actuation commands of three chambers for the desired laser spot positions. Four different path following tests were performed under different laser control conditions. Open-loop laser spot tracking was used to follow spiral and pentagon paths with mean tracking errors of 150 μm and 189.4 μm, respectively (**Fig. 4.16a**). In the open-loop path following of the pentagon, the repeatability of laser beam manipulation was also calculated using the deviation of successive cycles from the first cycle. The tracking results of the pentagon path following demonstrated a mean deviation error of 44.6 ± 8.2 μm, which suggests that the robotic laser manipulator paired with the learning-based model could produce repeatable and accurate path following even with open-loop control.

To further improve the control accuracy, a closed-loop controller has also been investigated. For the purpose of lab-based validation, an EM tracking probe was incorporated to provide feedback. The robot achieved enhanced tracking accuracy even when following an irregular batman symbol path with the root-mean-square tracking error reduced to 121.5 μm. Considering the limitation of EM sensors in the MR environment, closed-loop visual servo control based on the MR-compatible fiberscope view is also investigated. Apart from providing feedback for control, the fiberscope could view the laser spot projected on the surgical site, thus offering an interface for both task definition and

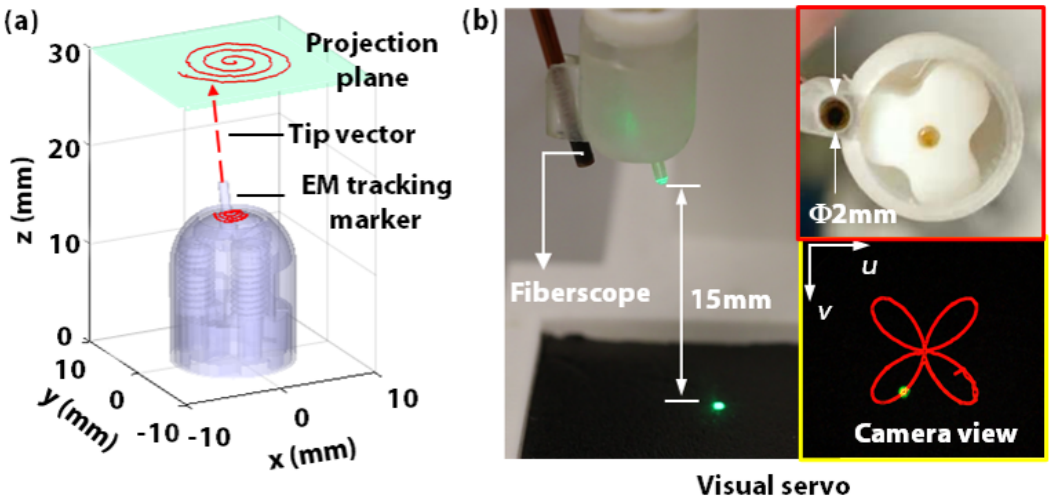


Fig. 4.15: Lab-based laser beam steering control. **(a)** Path following of the laser spot on a projection plane using the soft laser manipulator. The laser spot position can be determined based on the robot tip position and orientation measured with an EM tracking system. **(b)** Laser spot visual servoing is achieved with an MRI-compatible fiberscope.

visual monitoring. A 10-meter-long fiber optic camera (**Fig. 4.15b**) was utilized to meet the requirements of operation under MRI. As demonstrated in **Fig. 4.16b**, the laser spot could be accurately controlled to follow a four-leaf path defined in the camera view, with a mean tracking error of 1.04 pixels, corresponding to 69.3 μm on the projection plane. In ablation procedures, the laser spot footprint may need to fill a targeted area. Apart from the spiral pattern shown in **Fig. 4.16a**, the laser manipulator also maintains excellent accuracy when filling an area with complex boundary in a zig-zag pattern, as shown in **Fig. 4.17**.

The tracking error statistics of the path following experiments are shown in **Fig. 4.16b**, indicating that the laser manipulator is capable of tracing any basic and complex paths with high accuracy ($<$200 μm), which meet the clinical requirements on laser spot targeting ($<$1 mm) [107]. The spiral and pentagon pattern results demonstrate that the laser manipulator can accurately perform a circular and linear motion that commonly comprises other, more complex patterns. In addition, the robot maintains high accuracy even on following complicated trajectories, e.g., a batman symbol, which involves sharp turns and thus sudden changes of the robot trajectory. This is typically a challenging task for soft robots to perform precisely due to the need for sudden volume changes of the robot's soft actuation chambers, coupled with the compliance of the robot body.

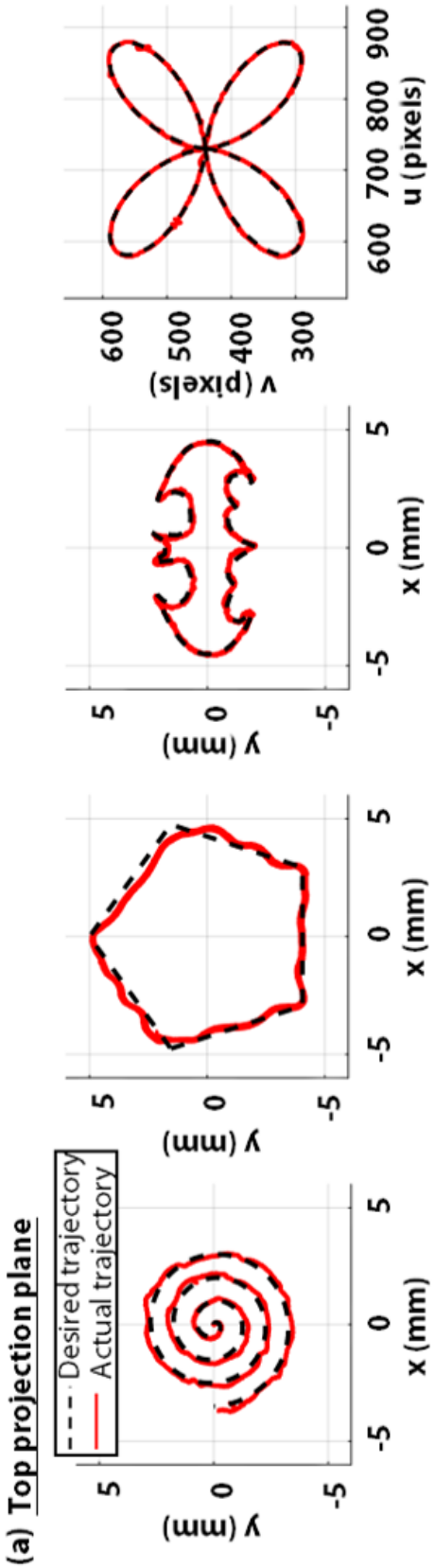


Fig. 4.16: **(a)** Laser spot trajectories tracked with open-loop control (column 1 and 2), EM-tracked closed-loop control (column 3), and laser spot visual servoing control (column 4). **(b)** Associated tracking errors.

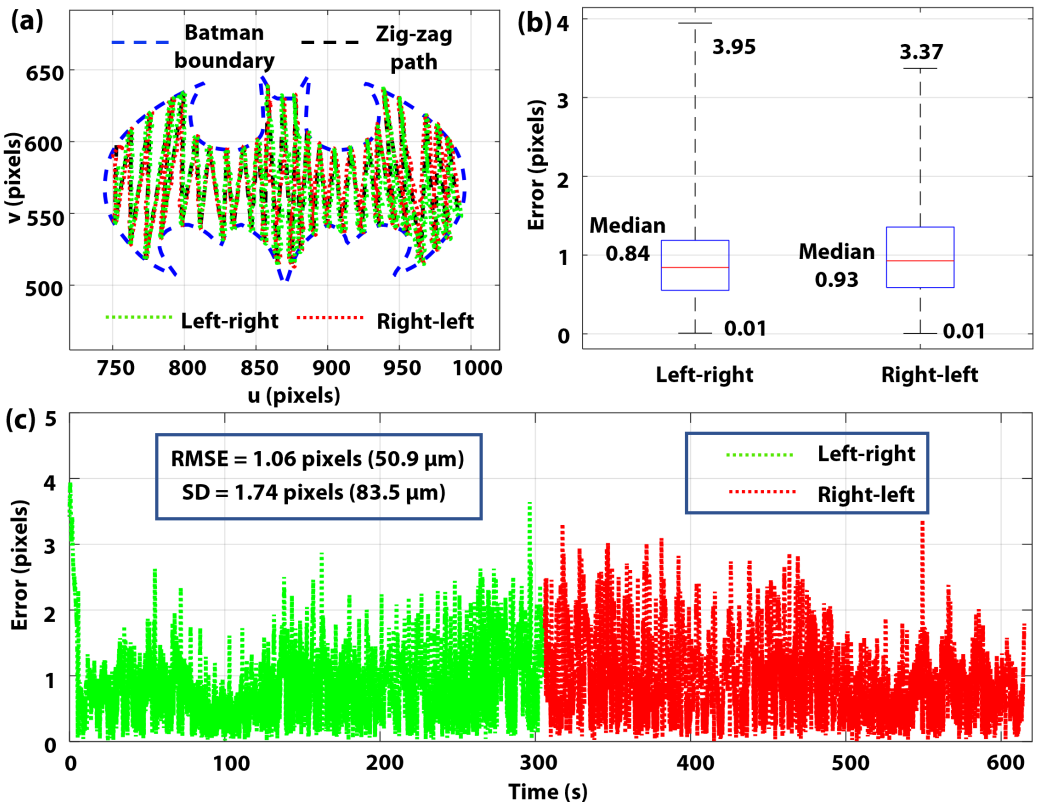


Fig. 4.17: Laser spot steering in a zig-zag pattern. **(a)** Laser spot trajectory tracked when filling the area of a batman shape with a zig-zag pattern. **(b)** Tracking error distributions in both the left-to-right and right-to-left paths. **(c)** Associated tracking errors over 10 minutes of tracking.

4.6 Conclusion

This chapter presents a visual servo controller for a soft robotic laser steering system. Epipolar geometry is first constructed to relate the camera frame with soft robot kinematics, thus modeling the inverse transition mapping. Based on this inverse mapping, a feedback controller is established, without having to incorporate prior knowledge of the tissue surface geometry. A state observer is also designed to provide a dynamic estimation of robot configurations, enabling accurate computation of the inverse transition mapping. The closed-loop controller provides robustness to actuation disturbances compared to open-loop control. It brings new opportunities to precise laser projection using a soft manipulator, which could be equipped on an endoscope tip for endoscopic laser surgery. Moreover, the MRI-compatibility of soft laser manipulator enables further application in MRI-guided laser treatment, which could enhance the intra-op monitoring of laser ablation outcomes (**Chapter 6**).

The proposed controller enabled precise targeting of the laser spot using a fluid-driven soft robotic manipulator. This is achieved by delicately adjusting the inflation volumes of three fluid chambers over a 10-meters long pipeline, which is a challenging task demanding for high-precision control. With the proposed controller, accurate laser spot path following was demonstrated, with an RMS tracking error of $<$4 pixels in camera view, corresponding to a physical error of laser spot projection $<$140 μm, which meets the clinical requirements on laser spot targeting ($<$1 mm) [107]. This is maintained even throughout 70 repeated cycles, with the maximum tracking error smaller than 11.32 pixels ($\sim$396.2 μm).

In general, both the model-based and the learning-based controller can achieve accurate laser spot steering control. However, the model-based method needs for calibration and initialization of the camera extrinsic parameters, which could be varied during the robot setup. The learning-based method can accommodate this through the use of training data acquired *in-situ*. But the initialization of the learning model may be time-consuming. In our future work, the synergetic use of model-based and learning-based methods will be of our research interest. The influence of MRI sequence parameters on the imaging of tissue being ablated will be investigated. Real-time tracking with MRI [144] will be incorporated to enable registration between MR image coordinate and robot coordinate. Visual servo based on MR-thermometry will also be of our research interest. Proper surgical laser sources, e.g. pulsed laser, will be investigated to achieve tissue cutting or ablation in cooperation with the laser beam manipulation. Another plan is to improve the laser spot control performance by decoupling the velocity profile and path following [190].

Chapter 5

Soft Robotic System for MRI-guided Transoral Laser Surgery

5.1 Introduction

Most of the existing MRI-guided laser surgery are based on contact-laser, without the need to frequently manipulate the laser beam. Currently, there is no laser steering manipulator that meets the requirement of MR safety. Soft robotic manipulators possess a unique combination of properties that are not typically seen in surgical robotics. Driven by pressurized fluids, they have the potential to address these challenges through high conformability, dexterity, and the ability to be MR-safe. Researchers have sought to apply soft robotics to several clinical applications such as colonoscopy [151, 152] and keyhole surgery [154]. Here, they leveraged the safety and conformability of soft and flexible actuators, where precision is a secondary goal. Few studies have reported intra-op MRI-guided soft robotic systems, particularly at miniature scales, while maintaining sufficient accuracy for treatment. Submillimeter-scale laser manipulation is required in transoral laser microsurgery [107], while still ensuring the safety of delicate surrounding structures.

To this end, a unique fluid-driven soft robotic system that employs reinforced miniature soft actuators is proposed to provide precise laser delivery for TLM. A hydraulic actuation approach is taken for its intrinsic MR-safety and ability to miniaturize the robot design by allocating the larger actuator systems to outside of the MRI room [135]. When compared to pneumatic actuation [191], hydraulics does not suffer from the high compressibility of air,

which can cause mechanical transmission delay and associated control inaccuracies. The system generates zero EM interference, allowing the introduction of intra-op MRI guidance to evaluate the laser/tissue interaction process, which plays an important role in balancing adequate tumor resection and function preservation. The soft robotic system is compact (Ø12×100 mm) and lightweight (∼200 g), while providing 5 degrees-of-freedom motion. Integrating with a laser fiber, the robot enables flexible delivery and precise targeting of surgical lasers. A miniaturized laser manipulator is designed with optimized reinforcement to provide delicate laser beam steering (+/- 26°) in a confined area. The manipulator is actuated through micro-volume (<0.004 mL) fluidic inflation of elastomer chambers, which are regulated with a learning-based controller. The learning-based controller can accommodate any modeling uncertainties introduced during the robot setup. A fiberscope is used to monitor the laser spot footprint and provide closed-loop guidance for laser beam steering control. During laser ablation, MR thermal imaging can be conducted to evaluate the heat diffusion in tissue, thus offering real-time monitoring of the ablation margin. 3D MR scanning can be performed intra-operatively to enable *in-situ* assessment of the ablation completeness. The key contributions of this work can be differentiated as follows:

1. Development of an MR-safe endoscopic robot system allowing flexible access to lesions in the confined oral and pharyngeal cavities. The system provides a stable platform for laser beam delivery and navigation to the targeted lesions with MRI guidance. The robot is compact and lightweight, enabling mounting on the patient and operation with standard MR imaging coils;

2. Design of a reinforced miniature fluid-driven soft manipulator capable of delicate, precise, and repeatable laser spot steering on the mucosa. The MR-safe hydraulic actuation enables simultaneous robot operation and MR imaging while ensuring minimal image artifacts.

5.2 Design Requirements

Oncologic management of HNCs requires local-regional control of disease, as well as preservation of speech and swallowing functions [114]. There is often a dilemma between adequate tissue margin for clearance and functional destruction from over-resection. To enable agile delivery of laser energy to target regions, the robot should provide dexterous navigation through intraoral and oropharyngeal cavities, even down to the hypopharynx

and larynx. The robotic scope should ensure safe contact with the upper airway, and also offer a stable platform for precise and local control of laser beam projection on oral and pharyngeal lesions. The distal laser manipulator should satisfy two key clinical requirements: sufficient workspace to cover the average men and women vocal fold size (20×20 mm^2) with a projection distance of 20∼25 mm; laser spot targeting accuracy of $<$1 mm [107], aforementioned in **Chapter 4**.

Provided with intra-op MR anatomical and temperature images, tumor margins can be identified *in-situ*, and the ablation/incision depth can be monitored in real-time. This ensures tumors/lesions are marked and resected with a margin of 5-10 mm beyond the tumor [192]. To monitor minute physiological changes in the region of the lesion, zero image artifacts have to be ensured even during robot operation in close proximity to the imaging site. To this end, the robot should meet the rigorous requirement of zero EM interference, comprising of no conductive, metallic, or magnetic components, i.e., the MR safety defined in the ASTM F2503 standard [75]. In addition, the robot should be compact and lightweight, enabling mounting on the patient and operation with standard MR imaging coils.

5.3 System Overview

The proposed soft robotic platform for MRI-guided transoral laser microsurgery is shown in **Fig. 5.1a**. A soft robot integrated with a lens-pigtailed laser fiber enables safe endoscopic intervention through the dental anchorage in the oral cavity, and directs energy delivery to the lesion target deep to the larynx. Two segments of the soft robot, i.e., active bending section and distal laser manipulator, were employed to offer, respectively, coarse robot navigation and fine laser beam pan-tilting (**Fig. 5.1b**). Each segment comprises three elastomeric chambers with volume varied by fluidic inflation/deflation to provide omni-directional bending. Miniature spring reinforcement constraints are integrated into the laser manipulator to enable repeatable and accurate steering of laser beam even inside a protective outer shell (**Fig. 5.1c**). Together with insertion, the robot provides five DoFs in manipulating the laser lens. Insertion/retraction is achieved by using a two-cylinder hydraulic actuator, which is coupled with a passive bending section to transfer rotational displacement to linear advancement. Both the soft chambers and cylinders are hydraulically driven via 10-meters long pipelines, which are channeled through the MRI room waveguide and connected to robot consoles in the control room. The robotic scope is compact with a size of Ø12×100 mm and is lightweight at 200 g, thus allowing non-invasive mounting to

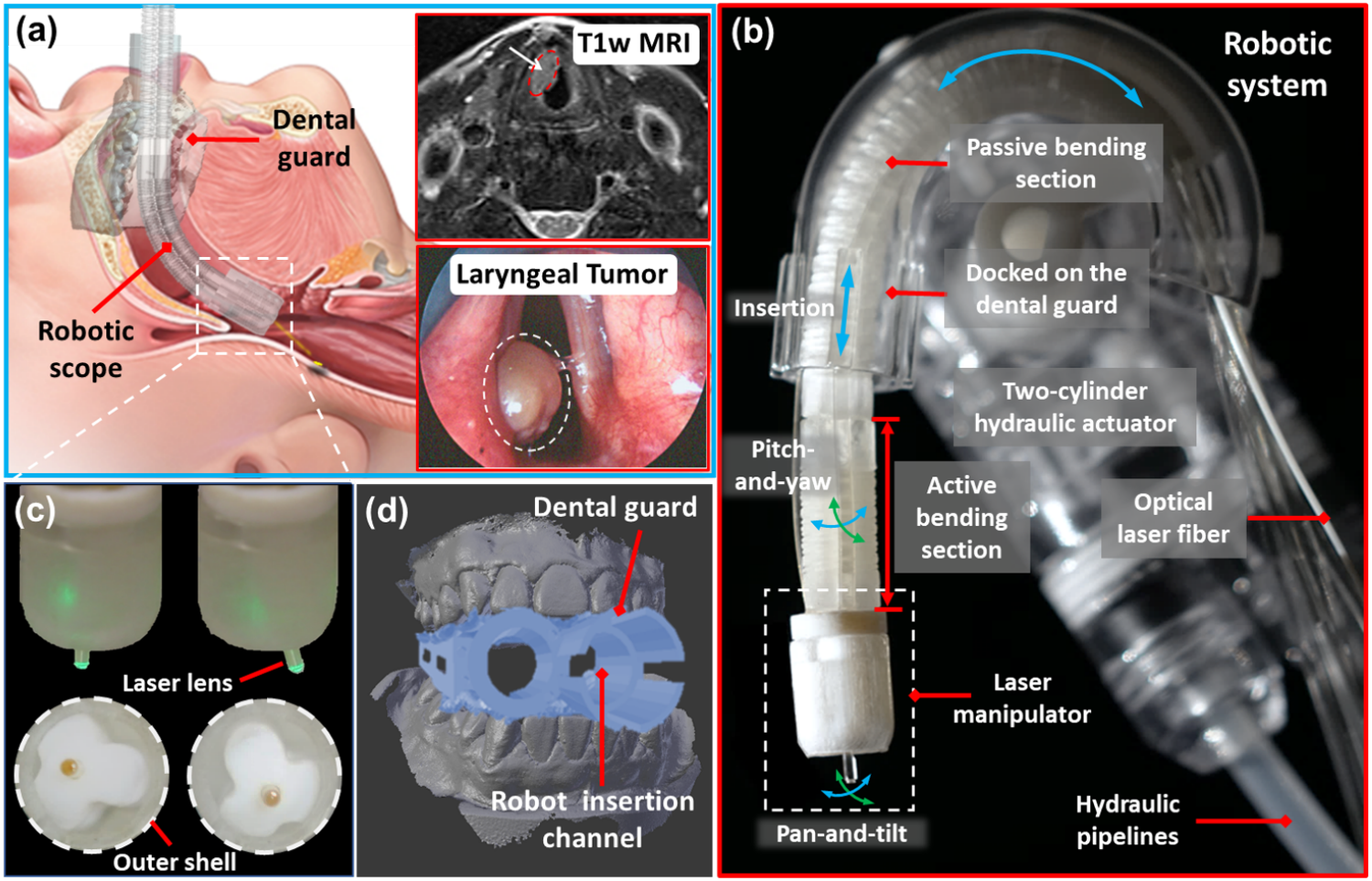


Fig. 5.1: Overview of the robotic system. **(a)** Illustration of the soft fluid-driven robot positioned in the larynx through the oral cavity (left) and aiming at a laryngeal tumor [51,52] (bottom right). Intra-op MRI can be introduced to provide the operator with fine monitoring of the laser ablation progress (top right). **(b)** Overview of the five DoFs endoscopic motion. Pan-and-tilt are provided by both the active bending section and the laser manipulator. The passive bending section facilitates insertion and retraction of the robot by transferring rotation of the hydraulic actuator to linear motion. **(c)** Side view (top) and below view (bottom) of the soft laser manipulator, which enables omni-directional laser beam steering inside a protective outer shell. **(d)** Patient-specific dental guard designed for robot anchorage and creating an open-jaw position of the patient.

the patient with a custom-made dental guard (**Fig. 5.1d**). MR-based tracking markers can be assembled with the dental guard to offer positional localization in MRI, thus registering the robot base with anatomical MRI coordinates and allowing the introduction of intra-op MRI guidance to monitor the laser ablation progress. Intra-op MRI can provide clinicians with enhanced perception of tissue incision width and depth [193], while providing the unique capability of measuring dynamic changes of tissue temperature in 3D, which is crucial in alerting excessive hyperthermia to the lesion surroundings, particularly in depth.

5.4 Soft Robotic Laser Targeting System

In this section, the design of a hydraulic-driven soft robotic system is presented. Apart from the laser scanning segment, a continuum segment is incorporated for adjusting the orientation of the camera view. Finite element analysis was conducted to optimize the robot design parameters. The mechanical performances of the robot were evaluated.

5.4.1 Soft Manipulator Design and Fabrication

To fulfill the rigorous requirement of MR safety, soft robots fabricated with elastomers, or polymer-based materials, are studied. These have gained increasing interests in medical applications, e.g., endoscopes, accredited to their flexibility and safe compliance with soft tissue surroundings. However, soft chambers made of pure elastomer material cannot sustain high pressure, limiting their loading capabilities and allowing uncontrolled expansion upon pressurization. Prior arts proposed to restrict soft chambers' radial expansion by paper-reinforced [194] or fiber-reinforced methods [195], which restricted their deformation to 1D elongation. However, such reinforcement cannot improve the soft robots' stiffness, thus the mechanical responsiveness. The discrete constraint would also give rise to non-uniform distributions of the chamber wall thickness upon pressurization, which may result in high local stresses and rupture.

To this end, a soft robotic steering mechanism with a spring reinforcement constraint (SRC) is proposed to enable omni-directional laser lens manipulation. The mechanism consists of three separate elastic chambers, located 120° apart in a triangular configuration. Reinforced springs fabricated from stiff materials are applied to the surface of each individual chamber during 3D printing. With the SRC, the soft chambers could bend with higher stiffness and lower hysteresis. The robotic scope is composed of two actuated

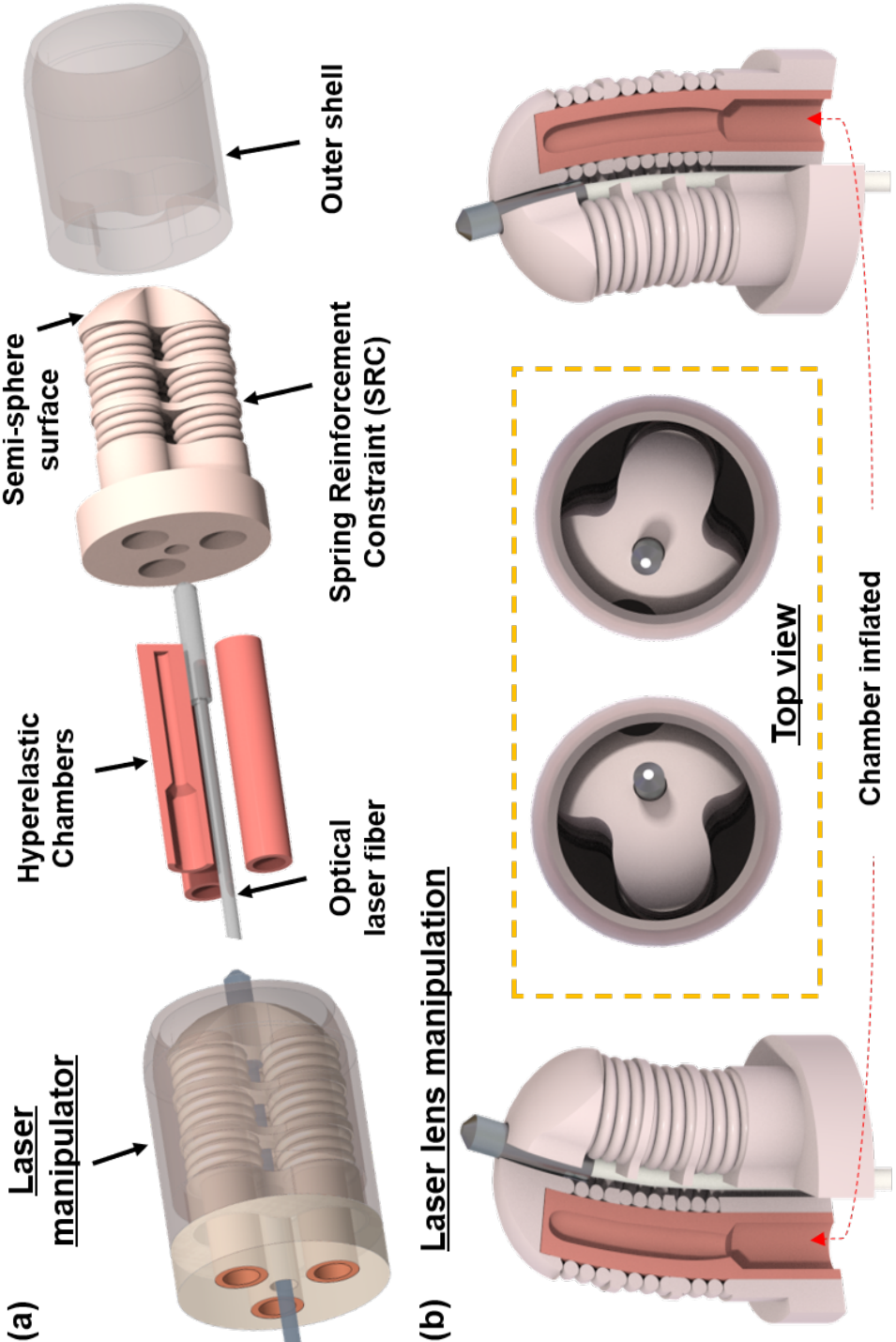


Fig. 5.2: Design of the soft robotic laser steering system. **(a)** Laser manipulator integrating three soft chambers with rigid spring reinforcement constraints. An outer shell protects the robot bending and laser steering from external disturbances. The hyper-elastic chambers are 3D-printed from flexible AgilusClear30™. The outer shell and SRCs are 3D-printed from VeroClear™. **(b)** The laser lens is housed in the center channel of the laser manipulator and can be steered omni-directionally.

segments. The distal segment, namely the laser manipulator (**Fig. 5.2a**), incorporates SRC and soft chambers at a total length of 10 mm, offering omni-directional pan/tilting of the laser lens in a local area (**Fig. 5.2b**). An extra outer shell (Ø12 mm) is designed to protect the laser beam maneuvering from external disturbances, e.g., contact with the oral and pharyngeal tissues. Note that the outer shell will not affect the bending of soft manipulator. This isolation design enhances the robustness of the laser manipulator and ensures laser beam steering with high accuracy. The second actuated segment, which is the active bending section (**Fig. 5.3**), comprises three soft reinforced chambers in a total length of 24 mm and an outer diameter of 9.2 mm, providing coarse navigation for lesion targeting. The soft continuum segments were fabricated through 3D printing (Stratasys, Connex 3 Objet 350), which enables hybrid printing of rigid and soft materials. This fabrication method offers the option of fast robot customization, e.g., chamber lengths. There is a trend towards 3D printing materials that are biocompatible and sterilizable. However, the presented robot can also be fabricated by molding with mass-producible materials, e.g., medical-grade elastomers. Additionally, the presented system is designed to be single-use, which can reduce sterilization and cross-contamination concerns.

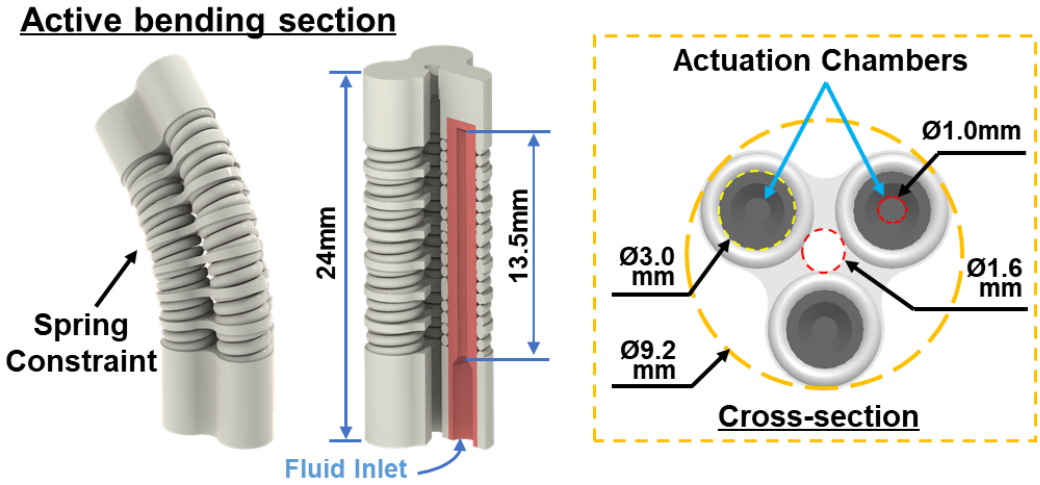


Fig. 5.3: Active bending section with length of 24 mm and outer diameter of 9.2 mm, comprising of three soft chambers with individual spring constraints.

5.4.2 Design Optimization based on Finite Element Analysis

FEA is conducted to simulate the robot characteristics upon pressurization (**Fig. 5.4a**), thus facilitating the design optimization. The specific strain-stress relationship of the selected elastomer (i.e., AgilusClear30^{TM}) was evaluated through a uniaxial tensile test and

approximated by the Ogden model in the simulation. The robot structure was tessellated with the quadratic tetrahedron element (e.g., Abaqus C3D10H). Strain hardening/stiffening effects were considered such that the nonlinear stress-strain behavior can be simulated using specific methods, such as arc-length.

Two model characteristics in terms of (i) bending angle and (ii) maximum von Mises stress are analyzed to evaluate the effect of partitions in the reinforced spring model. As shown in **Fig. 5.4b**, the effective bending angle decreases with more spring partitions, meaning that greater pressure is required to overcome the additional model rigidity. This implies that the bending stiffness of the soft manipulator is increased with spring partitions. However, the maximum von Mises stress over the whole model is also increased by adding partitions, as shown in **Fig. 5.4c**. The enhanced stiffness could improve the hysteresis and responsiveness of the soft manipulator, thus allowing accurate laser beam control. However, high local stress would introduce model fracture and reduce the robot's durability. Three is the optimized number of partitions for our design, balancing the trade-off among compactness, bending stiffness, and local material rupture. Another concern is the twisting of the soft actuator caused by the extension of each spring structure. FEA result indicates the average twisting angle is negligible ($<1°$) and reduced with the spring partition number. Each partition connects the three individual chambers/spring coils, which causes them to share the twisting loading, ultimately simplifying the bending kinematics and improving robot controllability.

5.4.3 Flexible Laser System

Tissue ablation is predominantly governed by achieving intensity levels beyond the specific ablation threshold. The tissue ablation threshold depends on both the laser wavelength and power density. A focused lens is optimal in terms of reducing laser spot size, thus increasing power density. But it introduces difficulty to the device assembly and miniaturization. The rod collimator compactly integrated with laser fiber can converge the laser beam to a small spot (<1 mm), thus ensuring high power density. The fiber-optic collimator can be customized in a wavelength range of 400$\sim$2,000 nm, among which some wavelengths have been applied in ENT (ear, neck, and throat) surgery with bare-tip fibers.

A fiberoptic medical laser enables flexible energy delivery towards target tumors located in the deep laryngeal region. An 808-nm diode laser with a power of 3 W was selected for the lab-based evaluation. The laser was guided by a 10-meters long optical fiber, with a

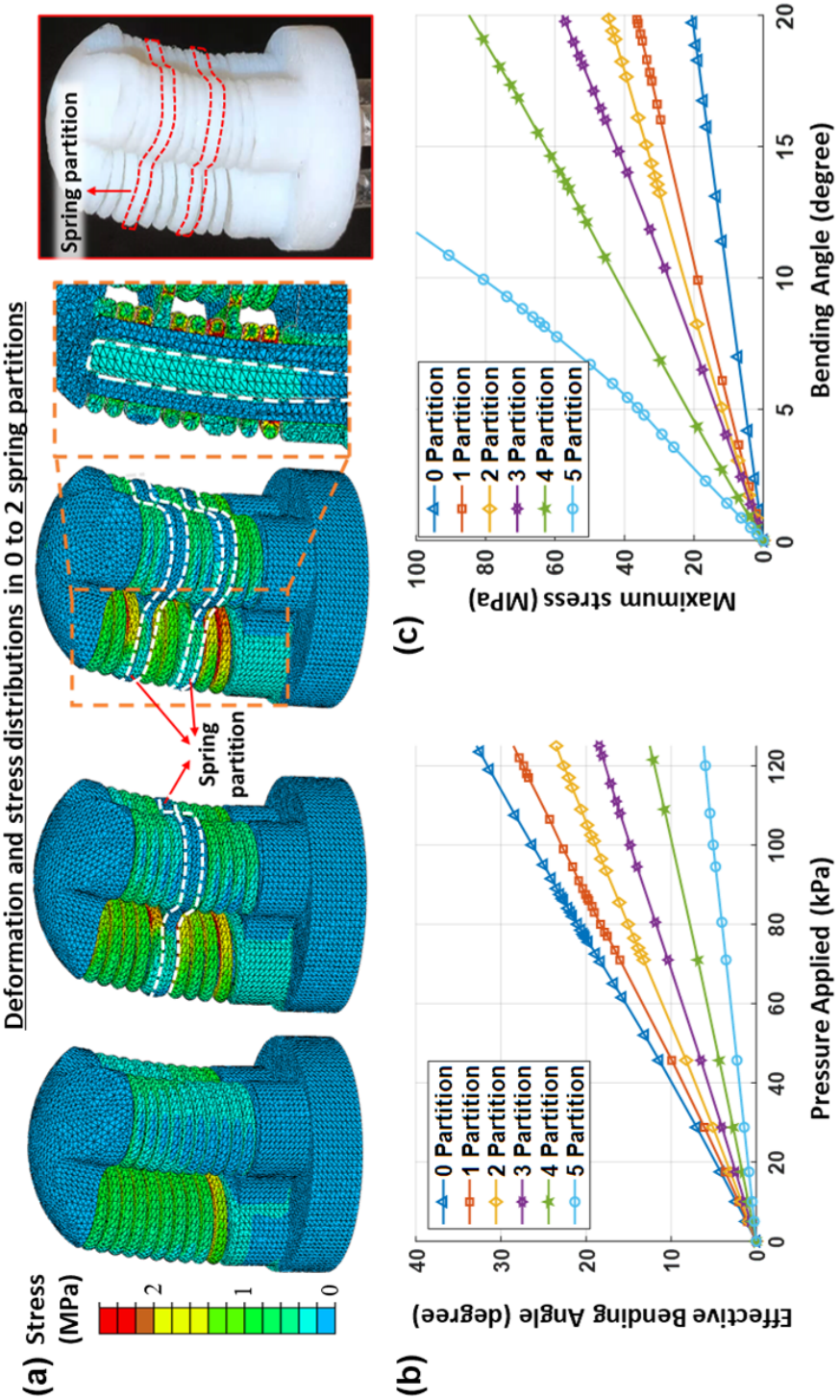


Fig. 5.4: Design optimization of the spring reinforcement constraint. **(a)** Finite-element-simulated deformation of the laser manipulator at a bending angle of 10°, and stress distributions under three different settings: zero, one and two spring partitions. A single chamber was pressurized to induce bending. **(b)** Effective bending angle versus applied pressure to the soft chamber, varied from zero to five spring partitions. **(c)** Maximum von Mises stresses in each model plotted against bending angle.

core diameter of 105 μm. Green laser (532 nm) with a power of 10 mw was coupled into the same fiber to provide visual guidance. A GRIN-lens collimator (≃Ø1.2×9.0 mm) was assembled at the fiber tip to emit a parallel-beam spot on tissue (**Fig. 5.5**). Compared with free-end fiber, the collimator could output a small laser spot (<1 mm) with high power density (>4 W/mm^2) even at a long distance (∼15 mm), hence allowing laser ablation without tissue-fiber contact. This minimizes the heat-affected zone on tissue and allows retention of the micro-lens after the procedure.

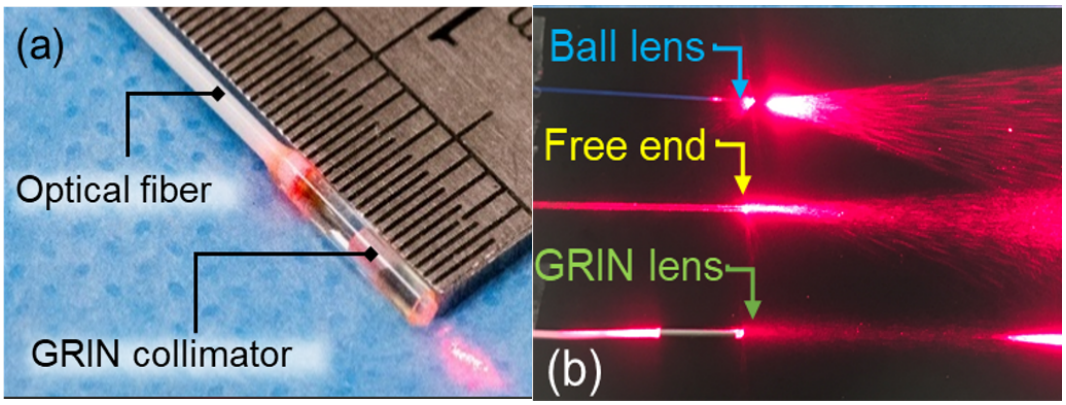


Fig. 5.5: (a) GRIN lens pigtailed with laser fiber. **(b)** Laser beam profile emitted from three types of fiber ends.

5.4.4 Mechanical Response Evaluation

Workspace: The maximum bending angle of the fabricated soft segments upon pressurization was measured using an EM tracking system (NDI Aurora®). The laser manipulator can steer the laser lens to a maximum angle of ±26° (**Fig. 5.6a**), which enables a scanning area of 15×15 mm^2 at a 15-mm projection distance. As for the active bending section, my design offers an angular workspace of ±60° (**Fig. 5.7a**), which can be further enlarged by elongating the segment length or reducing the number of spring partitions.

Frequency response: The dynamics response of soft manipulator with hydraulic actuation has been tested. The soft chamber was inflated/deflated through a hydraulic cylinder actuated by a servo motor. Ten-meter-long hydraulic pipeline was utilized to meet the transmission requirements between the MRI and control rooms. The positional output of the robot tip under periodic sinusoidal input from 0.1 Hz to 5 Hz was captured by an EM tracking marker. Fourier analysis was conducted to evaluate the harmonic response. As shown in **Fig. 5.6b**, the bandwidth was approximately 1 Hz (angular velocity: ∼20°/s), at which the magnitude response decreased -3 dB [196]. This indicates that the changing

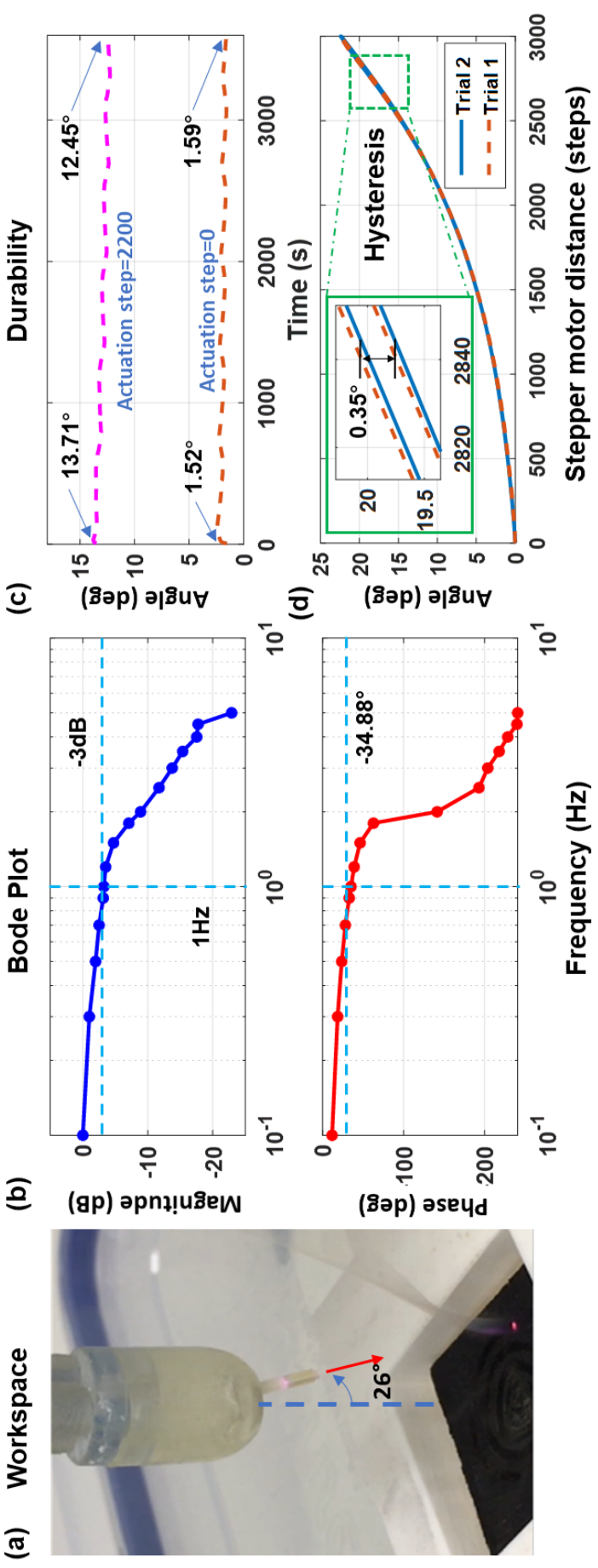


Fig. 5.6: Mechanical performance test. **(a)** Angular range of the soft laser manipulator. **(b)** Bode plot showing the dynamic response of laser lens steering. **(c)** Robot durability test performed by repeatedly actuating the laser manipulator 2,000 times over 1 hour. **(d)** Hysteresis test during fluidic inflation and deflation of the soft chambers.

frequency of positional command could be less than 1 Hz in order to avoid low-pass filtering by the soft robot composite. The bandwidth is intrinsically governed by the robot stiffness, which could be improved by fabricating the spring constraint with stiffer materials (e.g., PEEK). The phase lag was kept at less than 35° within the bandwidth. The transmission latency, or time delay from the computer signal to the distal output, was measured as 83ms on average with the input frequency less than 1 Hz. Note that in our setup, the laser parameters (e.g., power, pulse duration) could be modulated online based on the MR thermometry feedback. This could alleviate the need for high-speed laser scanning, which is primarily for avoiding local over-heating.

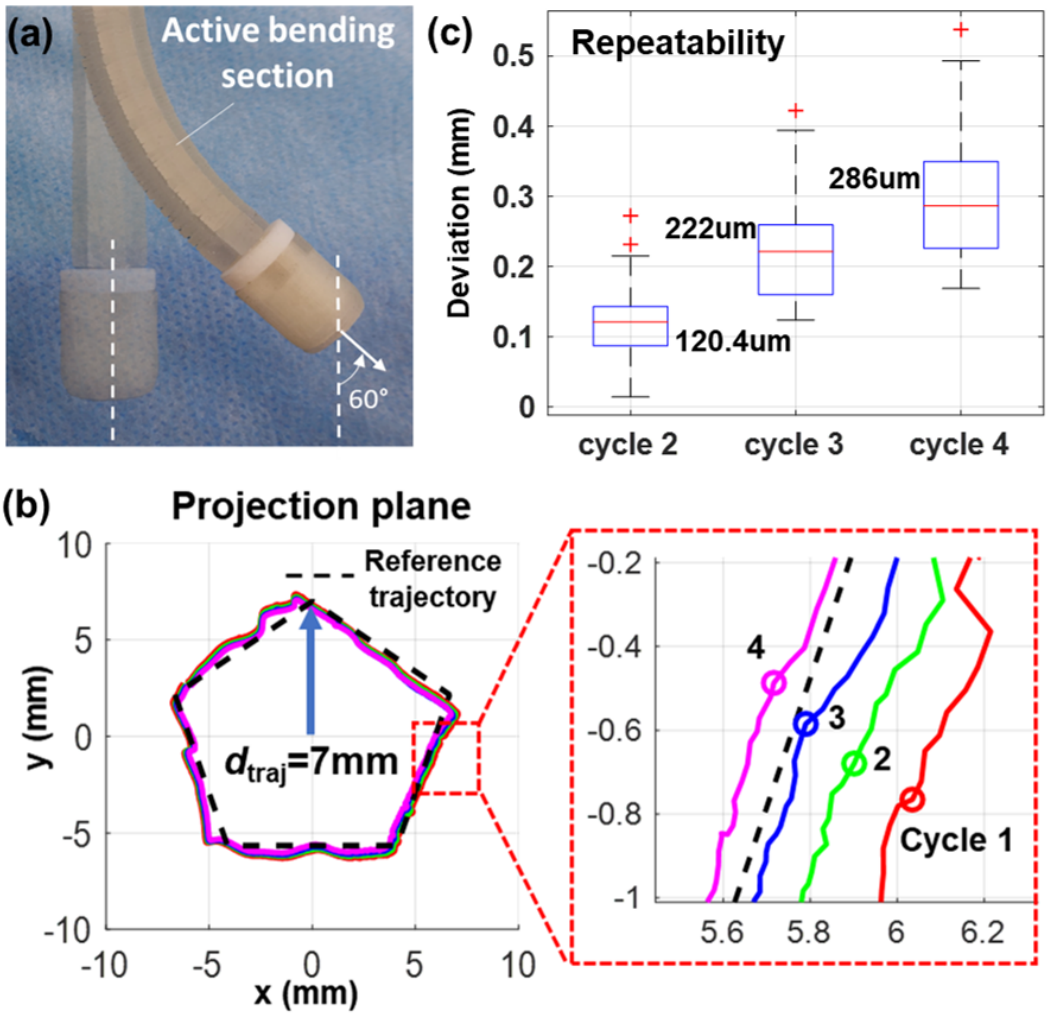


Fig. 5.7: Mechanical performance testing for active bending section. **(a)** Angular range. **(b)** Following a path on projection plane in numerous cycles. **(c)** Repeatability of the path following.

Durability: The reliability of soft chambers (e.g., material fatigue) should be maintained in a prescribed amount of usage, despite the intention to be a single-use device. As shown in **Fig. 5.6c**, the soft manipulator is inflated (actuation step 2,200) and deflated

(actuation step 0) repeatedly 2,000 times over a period of 1 hour. The robot remains functional while the absolute bending angle decreased only 1.5°, which can be compensated by a feedback controller or an online-updated learning model.

Hysteresis: The hysteresis of the soft manipulator has been tested by measuring its bending angle upon loading and unloading to the elastomer chambers. Two consecutive actuation cycles were performed. **Fig. 5.6d** shows the soft robot has a maximum hysteresis angle of 0.35°, which is negligible. This is accredited to the rigid spring reinforcement constraints that mechanically compensate for the hysteresis originating from the elastomeric material.

Repeatability: The repeatability of the active bending section has been evaluated by executing predefined actuation commands to follow a path in numerous cycles (**Fig. 5.7b**). The repeatability of robot footprint is defined as the deviation of the current trajectory from that tracked in the first cycle. Although the polygon reference path includes sudden changes of following directions, the soft robot still demonstrates a highly-repeatable path following with a mean deviation error of 254.5 μm over four cycles (**Fig. 5.7c**).

5.5 Docking and Insertion Mechanism

To perform laser surgery trans-orally, an anchorage platform is designed with instrument channels to house the soft robotic scope as well as other tools, e.g., a lighting guide. A hydraulic motor is integrated to provide advancement or retraction of the scope.

5.5.1 Design and Fabrication of Dental Guard

A docking mechanism is custom-made to provide stable robot mounting to patients in a non-invasive manner. A dental impression body was taken by using a fast-set dental alginate (Aroma Fine Plus) on the oral cavity (**Fig. 5.8a**) and then casted with plaster (**Fig. 5.8b**). The cast was optically scanned to form the 3D dental feature (**Fig. 5.8c**). The tailor-made dental guard (**Fig. 5.8d**) was constructed based on the 3D cast with an open bite of 35-40 mm, allowing 2 entrances for endoscopic instruments to work in a workspace $>$40×40×20 mm^3. The workspace can be further enlarged by pressing the tongue downward using a plastic depressor, enabling the accommodation of more instruments guided by the endoscope. Several mechanical interlocks are featured on the dental guard to anchor the entire robotic

mechanism. The robot can be securely mounted on the dental guard and allow docking through the oral cavity, thereby reducing the invasiveness to the soft palate and uvula. Thin and tiny RF wireless resonators [144] (<Ø3×10 mm, capsulized with Gadolinium-doped water) were fabricated and embedded inside the dental guard, providing high-contrast landmarks for registering the robot base in the MR image coordinates.

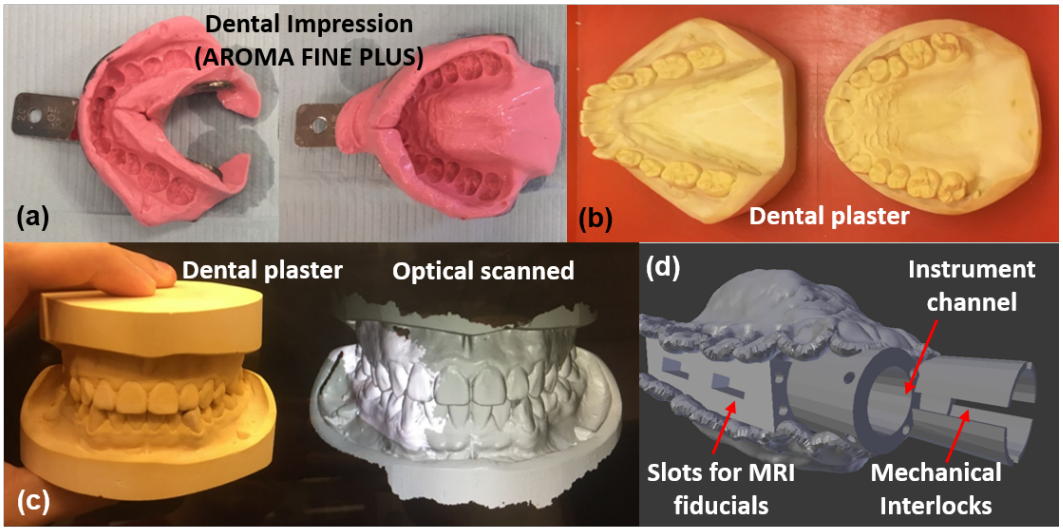


Fig. 5.8: **(a)** Dental impression with a fast set alginate. **(b)** Dental plaster molded using the impression body. **(c)** 3D dental feature obtained from optical scan. **(d)** Custom-made dental guard with two instrument channels.

5.5.2 Hydraulic Motor for Transoral Endoscope Insertion

A rolling-diaphragm sealed hydraulic motor (**Fig. 5.9**) [15, 135] has been integrated to advance the bending sections. Linear hydraulic actuation is transmitted to rotary movement using a pair of pinion-and-rack units. Rolling diaphragms offer excellent sealing as well as negligible sliding friction. A curved channel is designed to guide the passive bending section of the robotic scope. A pair of semi-rigid 10-meter-long pipes are used for the hydraulic transmission between the MRI control room and scanner room. The liquid pressure was pre-loaded to 1 bar to reduce the backlash and enhance the transmission efficiency. The hydraulic motor provides a maximum linear stroke of 20 mm, driving each pinion by 100.6° and corresponding to 50 mm robot insertion that is capable of reaching the oropharyngeal area. A concern related to the practical usage of hydraulic actuation under MRI is possible leakage or dripping when channeling the pipelines through the MRI waveguide. This problem can be addressed by the incorporation of an automatic fluid circulation system, with which water can be automatically filled or drained as needed during the setup.

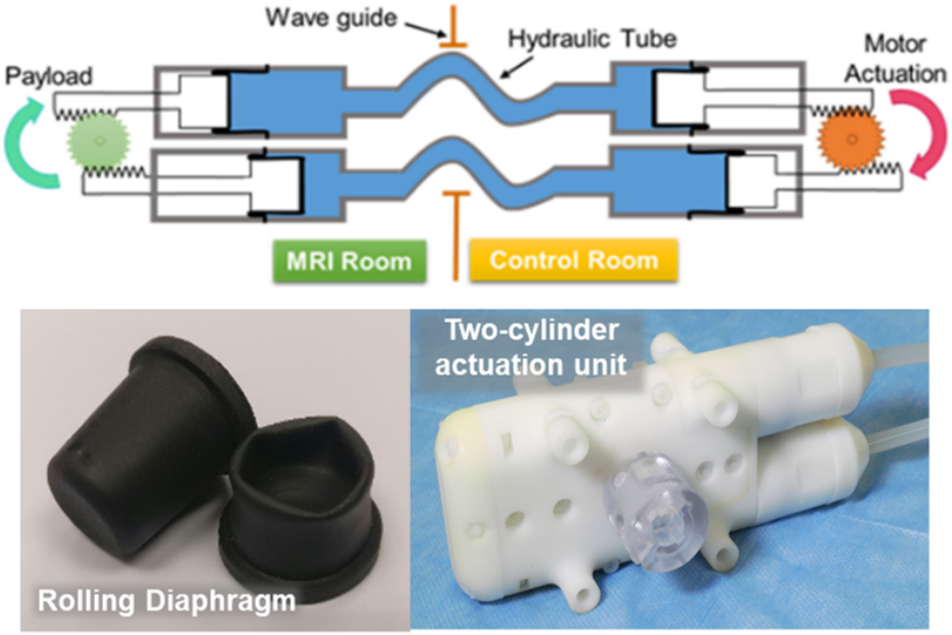


Fig. 5.9: Two-cylinder actuator driven by master-slave hydraulic transmission. Rolling diaphrams [53] are integrated to provide fluid sealing while ensuring minimal sliding friction.

5.5.3 Hydraulic Transmission Modelling and Optimization

To study the hydraulic motor characteristics and optimize its design, a dynamic model of the fluid transmission system is established. Based on the model, parametric analysis is conducted in simulation to identify the key factors that dominate the transmission response. This serves as a guideline for selecting the optimal design parameters governing the transmission performance in various aspects, such as stiffness and latency. Note that this model has the potential to be extended to characterize complicated hydraulic-driven systems with multiple DoFs. The modeling steps, design considerations, and simulation results are given as follows:

Dynamics Modeling of Hydraulic Transmission

This section details the dynamics model of the rolling diaphragm-based hydraulic transmission, which describes the relation between the input force F_{in} from the master side to the output force of F_{out} at the slave side. A spring-mass-damping system (**Fig. 5.10**) is adopted to approximate the interactions between the pistons and hydraulic fluid. Their interconnection is represented by the spring elements of stiffness $2K_t$, which account for the pipeline compliance and fluid compressibility. Assume that the system damping induced by the rolling diaphragm and fluid friction can be described by the coefficients c and B_f.

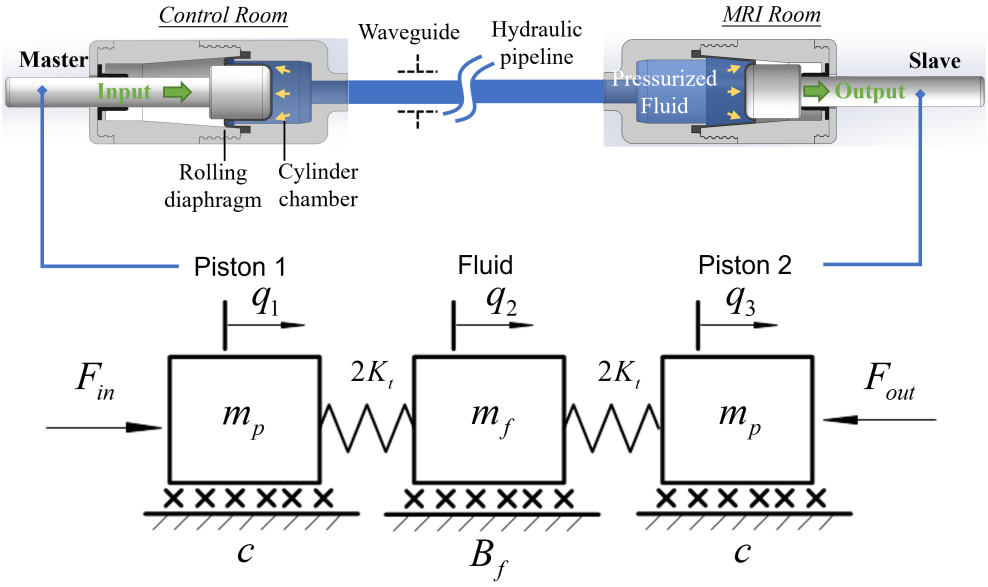


Fig. 5.10: Dynamics model of fluid transmission over a pipeline. Input force F_{in} is employed to push Piston 1, with force F_{out} applying at the output side. The cross markers "×" below three masses denote the fluid damping.

The pipeline compliance is first derived by considering that the input force F_{in} is applied at Piston 1 while the Piston 2 is blocked. Assume the pipe material is homogeneous and isotropic. The following equations in **Eq. 5.1** gives the hoop stress σ_θ and radial stress σ_r at the inner pipe wall in response to the fluid pressure P, based on Lamé Formula:

$$\sigma_\theta = P \cdot \frac{D_{pi}^2 + D_{po}^2}{D_{po}^2 - D_{pi}^2}, \ \sigma_r = -P \tag{5.1}$$

where D_{po} and D_{pi} denote the outer and inner diameters of the pipeline. Assuming that the pipe material is isotropic with a modulus of elasticity E and Poisson's ratio υ, we can obtain the hoop strain:

$$\varepsilon_\theta = \Delta D_{pi}/D_{pi} = (\sigma_\theta - \upsilon\sigma_r)/E \tag{5.2}$$

Considering the conversation of fluid volume in an enclosed system, the volume change induced by pushing the input piston should be equal to the volume change along the pipeline:

$$[(D_{pi} + \Delta D_{pi})^2 - D_{pi}^2]L_p = D_{in}^2 \cdot \Delta x_{in} \tag{5.3}$$

where D_{in} and Δx_{in} denote the diameter and displacement of the input piston, respectively. The length of pipeline is denoted by L_p. Thus, the increase of pipe inner diameter ΔD_{pi} can be calculated as:

$$\Delta D_{pi} \approx D_{in}^2 \cdot \Delta x_{in}/(2D_{pi}L_p) \tag{5.4}$$

With **Eq. 5.1** to **Eq. 5.4**, and the input force, $F_{in} = P \cdot (\pi {D_{in}}^2/4)$, which generates the fluid pressure, we can obtain the pipeline transmission stiffness K_p at the input piston:

$$K_p = \frac{F_{in}}{\Delta x_{in}} = \frac{\pi E D_{in}^4}{8 D_{pi}^2 L_p} \left[\frac{D_{pi}^2 + D_{po}^2 + \upsilon D_{po}^2 - \upsilon D_{pi}^2}{D_{po}^2 - D_{pi}^2} \right]^{-1} \tag{5.5}$$

Taking the fluid compressibility into account, we can calculate an additional stiffness:

$$K_f = E_v \cdot {A_{in}}^2 / V \tag{5.6}$$

where A_{in} represents the cross-sectional area of the piston with input force applied. The whole volume and bulk modulus of the enclosed fluid are denoted by V and E_v respectively. The two kinds of stiffness, K_p and K_f, can be regarded as two springs connected in serial, with the combined transmission stiffness K_t being calculated as:

$$K_t = K_p K_f / (K_p + K_f) \tag{5.7}$$

To model the damping effect of fluid transmission, the energy loss along the pipeline is analyzed. Assume that Piston 1 is moved by the input force F_{in} at a constant velocity, with zero loadings applied at the Piston 2 ($F_{out} = 0$). For simplification, assume the fluid as a steady and incompressible flow here. Therefore, the fluid head loss between input and output pistons, h_f, can be derived according to the Bernoulli equation:

$$\frac{P_{in}}{\rho} + \frac{v_{in}^2}{2} + g z_{in} = \frac{P_{out}}{\rho} + \frac{v_{out}^2}{2} + g z_{out} + h_f \tag{5.8}$$

where P_{in} and P_{out} are the fluid pressures, v_{in} and v_{out} are the fluid velocities, z_{in} and z_{out} denote the fluid elevations. The subscript "*in*" and "*out*" give the corresponding observing points, i.e., input and output piston, respectively. The density of fluid is represented by ρ. The gravitational constant is denoted by g.

As the lengths of the input and output cylinder are way shorter than the pipeline, the head loss in the cylinders is ignored. Local loss caused by the diameter change at the pipe connections is assumed to be negligible comparing to the pipeline loss, h_f, which can be approximately calculated as:

$$h_f = \lambda L_p {v_f}^2 / (2 D_{pi}) \tag{5.9}$$

where λ and v_f denote the flow coefficient and velocity [197], respectively. Considering a laminar pipe flow, we can obtain that $\lambda = 64/R_e$, with Reynolds number given as $R_e =$

$\rho v_f D_{pi}/\mu$, where μ denotes the fluid's dynamic viscosity.

Neglecting the change of pipe elevations, we can ignore two corresponding terms, gz_{in} and gz_{out}, in **Eq. 5.8**. The change of fluid velocities between the input and output piston is negligible. Considering the assumption of zero payloads ($F_{out} = 0$), the pressure P_{out} equals to zero. With the relation $F_{in} = P_{in} \cdot A_{in}$, we can calculate the equivalent damping coefficient as:

$$B_f = \frac{F_{in}}{v_{in}} = 8\pi\mu L_P \left(\frac{D_{in}}{D_{pi}}\right)^4 \tag{5.10}$$

The inertia of fluid mass is approximated by an equivalent mass as seen at the piston, which can be calculated in the aspect of energy:

$$m_f v_{in}^2 = m_{cyl} v_{in}^2 + m_{pi} {v_f}^2, \tag{5.11}$$

where m_{cyl} and m_{pi} are the masses of fluid in the cylinders and pipeline, respectively. The equivalent fluid mass, m_f, can be given as:

$$m_f = m_{cyl} + m_{pi}(D_{in}/D_{pi})^4 \tag{5.12}$$

Fig. 5.10 shows the key parameters and components governing the dynamics of the fluid transmission system. Applying Newton's Second Law for each mass, we can obtain a linear differential equation as:

$$\mathbf{K}\mathbf{q} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{M}\ddot{\mathbf{q}} = \mathbf{F} \tag{5.13}$$

where

$$\begin{cases} \mathbf{K} = \begin{bmatrix} 2K_t & -2K_t & 0 \\ -2K_t & 4K_t & -2K_t \\ 0 & -2K_t & 2K_t \end{bmatrix} & \mathbf{C} = diag\,(c, B_f, c) \\ \mathbf{M} = diag(m_p, m_f, m_p) & \mathbf{F} = \begin{bmatrix} F_{in} & 0 & -F_{out} \end{bmatrix}^T \end{cases} \tag{5.14}$$

In **Eq. 5.13**, a state vector $\mathbf{q} = [\ q_1 \quad q_2 \quad q_3\ ]^{\mathbf{T}}$ comprises the displacements of input piston (q_1), fluid (q_2), and output piston (q_3), as defined in **Fig. 5.10**. The damping coefficient of the rolling diaphragm is denoted by c. Arranging **Eq. 5.13** into the form of state-space function, we obtain:

$$\begin{bmatrix} \dot{\mathbf{q}} \\ \ddot{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{M}^{-1}\mathbf{F} \end{bmatrix} \tag{5.15}$$

Table 5.1: Nominal values of design parameters in simulation

Parameters	Value
Pipe inner diameter (mm)	4.0
Ratio of pipe inner diameter to thickness	4.0
Pipe material	PA 6
Pipe length (m)	Value
Piston diameter (mm)	13.0

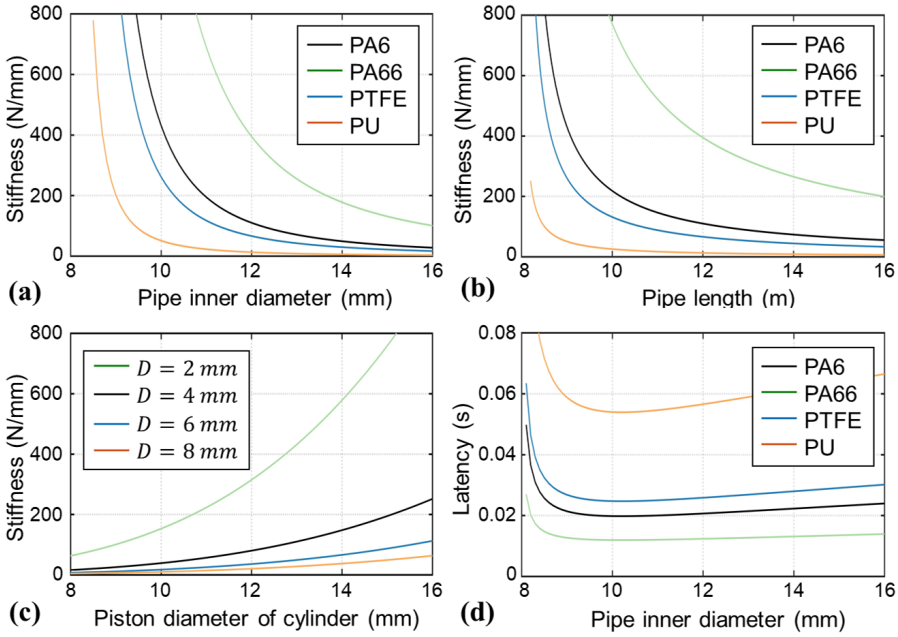


Fig. 5.11: Simulation showing the trend of transmission stiffness against: **(a)** pipe inner diameter; **(b)** pipe length; and **(c)** piston diameter. It can be observed that the stiffness profile is significantly affected by the pipe materials: PA 6, PA 66, PTFE and PU.

Design Considerations

Based on the dynamics model, numerical simulation can be conducted to investigate the effect of various design parameters on the transmission performances, such as stiffness and latency. This gives us a guideline in selecting the material, length, diameter, and thickness of fluid pipes, as well as the diameter of input/output cylinders. Pipes made of four common materials: polycaprolactam (PA) 6, PA 66, polytetrafluoroethylene (PTFE), and polyurethane (PU), are studied. Some nominal parameters in the simulation are given in **Table 5.1**.

Transmission stiffness: The transmission stiffness has a compelling effect on the system's repeatability and accuracy. **Fig. 5.11a-b** illustrates the trends of transmission stiffness in various pipe diameters and lengths. The stiffness decreases when the diameter or the length increases. It can also be observed that the choice of pipe material imposes a

Table 5.2: Physical parameters of the hydraulic transmission system

Parameter	Description	Parameter	Description
c_p	Velocity of pressure transient in fluid	υ	Poisson's ratio of pipe material
E_v	Bulk modulus of fluid	D_{pi}	Pipe inner diameter
E	Young's modulus of pipe material	e	Pipe wall thickness
ρ	Fluid density	ψ	Pipe support factor

significant influence on transmission stiffness. PA66 offers the highest transmission stiffness because it possesses the highest Young's modulus. However, the large bending stiffness of PA66 (0.17 N·m^2 with outer/inner diameter of Ø6/4 mm) hinders its application in conditions that requires flexible arrangement. **Fig. 5.11c** shows that the enlargement of piston diameter can also enhance the transmission stiffness, due to the increased piston diameter to pipe inner diameter ratio.

Transmission latency: The fluid pipe has a thick wall, given that the ratio between its inner diameter (D_{pi}) and wall thickness (T_p) is smaller than 10 ($D_{pi}/T_p < 10$). The power transmission in this system can be considered as the simultaneous occurrence of pressure and velocity changes. Such velocity of pressure transmission over enclosed fluid can be approximated as [198]:

$$c_p = \left(\sqrt{\frac{\rho \psi D_{pi}}{T_p \cdot E} + \frac{\rho}{E_v}} \right)^{-1} \tag{5.16}$$

where the parameters are summarized in **Table 5.2**. The pipe support factor ψ can be calculated as [198]:

$$\psi = 2e(1+\upsilon)/D_{pi} + D_{pi}/(D_{pi}+T_p) \tag{5.17}$$

Fig. 5.11d shows that the latency is around 21 ms using the baseline parameter values (**Table 5.1**). The ratio of pipe inner diameter to thickness varies while the pipe thickness is fixed at 1 mm. The enlargement of pipe inner diameter over 2 mm causes a slight increase of the latency. But changing the diameter D_{pi} to <1 mm will give a significant rise in latency. Meanwhile, the latency is also proportional to the pipe length. And the use of more rigid pipe material will reduce the transmission latency.

Overall, several design tradeoffs are developed to account for the required dynamics performance and dedicated operating conditions. Shorter pipelines are preferable as a general design rule to reduce the fluid inertia, give rise to a higher transmission stiffness, and decrease the transmission latency. For surgical applications that emphasize positional

accuracy, such as breast or prostate biopsy, pipelines with a smaller diameter and pistons with a larger diameter should be adopted to enhance the transmission stiffness. However, a small pipe diameter (< 2 mm) will dramatically increase transmission latency and damping, which will deteriorate the control and stability. Therefore, the pipeline diameter shall be expanded to accommodate for applications that require a rapid dynamic response, such as the tele-manipulation of a catheter in a master-slave manner.

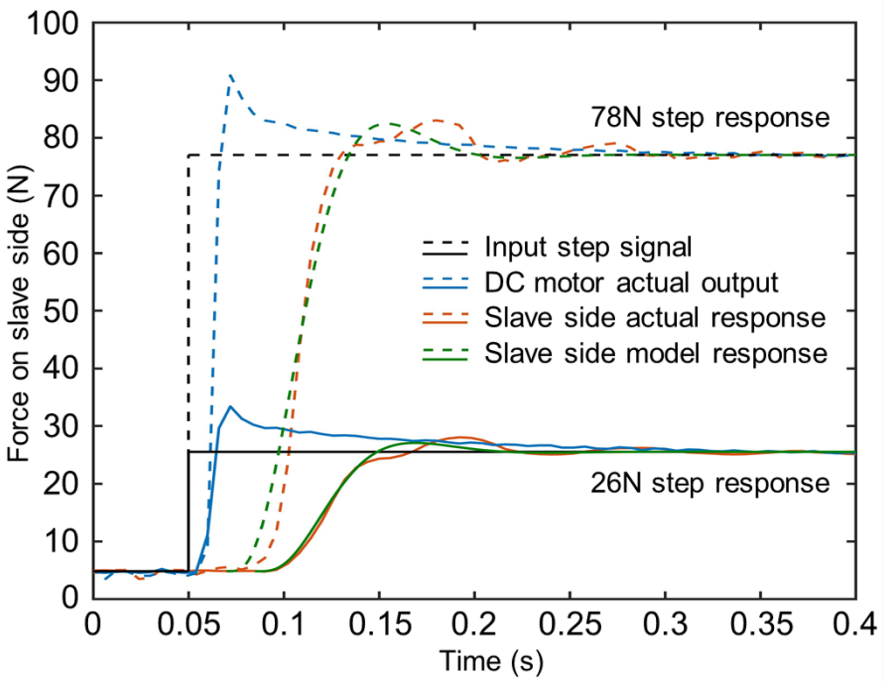


Fig. 5.12: Step response of a single cylinder transmission, which were measured in two different steps of magnitudes. Experimental and modeled responses are compared. The response time from the signal input is within 40 ms.

Step Response of Single-cylinder Actuation

Master and slave components of the transmission were connected by 10-meter semi-rigid nylon (PA 6) tubes, which had inner and outer diameters of ID:Ø4 mm and OD:Ø6 mm, respectively. Distilled water is selected as the transmission fluid, accrediting to its availability and ease of implementation. The water pressure pre-loading in all pipelines was controlled by a pressurized air supply system, a pressure regulating valve, and a fluid reservoir. Not only could this eliminate backlash, but it could also ensure the symmetric folding/unfolding of rolling diaphragms, thus reducing friction during fast motion transmission.

To evaluate the dynamics model, the force response of the fluid transmission has been experimentally tested and compared with simulation results. An input force was applied at the master side using a rack-and-pinon mechanism (**Fig. 5.9**), with the pinion driven by a

DC motor. The output piston was blocked, with its reaction force measured using a force sensor. Initial water pressure of 0.05 MPa was generated to eliminate the backlash between the rack and pinon, giving rise to an output force at 5 N. **Fig. 5.12** shows the measured force response to the step force input with two different magnitudes, i.e., 26 N and 78 N, which matches the corresponding simulation results calculated based on the dynamics model. The hydraulic transmission demonstrates a short response time of $<$40 ms, with the rise time from 10% to 90% appropriately 25 ms. The output force was settled within the 5% error band in 0.17 s. This indicates that forces can be transmitted rapidly using the proposed fluid transmission even with a long pipeline ($\sim$10 m), which facilitates responsive robotic teleoperation in MRI.

5.6 Real-time Positional Localization in MRI

Localization of the robot system in the MRI coordinate is the prerequisite of MRI-guided robotic intervention. The robot motion can be tracked or localized using MR markers (**Section 2.5.2**). In this section, an inductively coupled radio frequency (ICRF) marker for MR-based positional tracking is presented. The marker enables robust amplification of tracking signals at all magnetic resonance (MR) scanning orientations. The marker employs three curved resonant circuits fully covering a cylindrical surface that encloses the signal source. Each resonant circuit is a planar spiral inductor with parallel plate capacitors fabricated monolithically on a flexible printed circuit board (FPC) and bent to achieve the curved structure.

5.6.1 MR-based Wireless Marker

A wireless marker is a tuned LC tank circuit at resonant frequency matchable with the MRI scanner Larmor frequency. When a wireless marker is placed inside the scanner, only components ${B_1}^+$ and ${B_1}^-$ that are parallel to the marker's surface normal vector are effective for generating current through its inductor. The omnidirectional tracking arrangement is shown in **Fig. 5.13**, with three curved resonant circuits evenly wrapped on a 3 mm diameter hollow cylinder (wall thickness: 0.25 mm, length: 10 mm) surface. The cylinder is filled with 10 mM gadolinium-doped water which is a typical MRI contrast agent for enhancing signal amplification, acting as the internal signal source. Two sides of the cylinder were sealed with UV-curable adhesive.

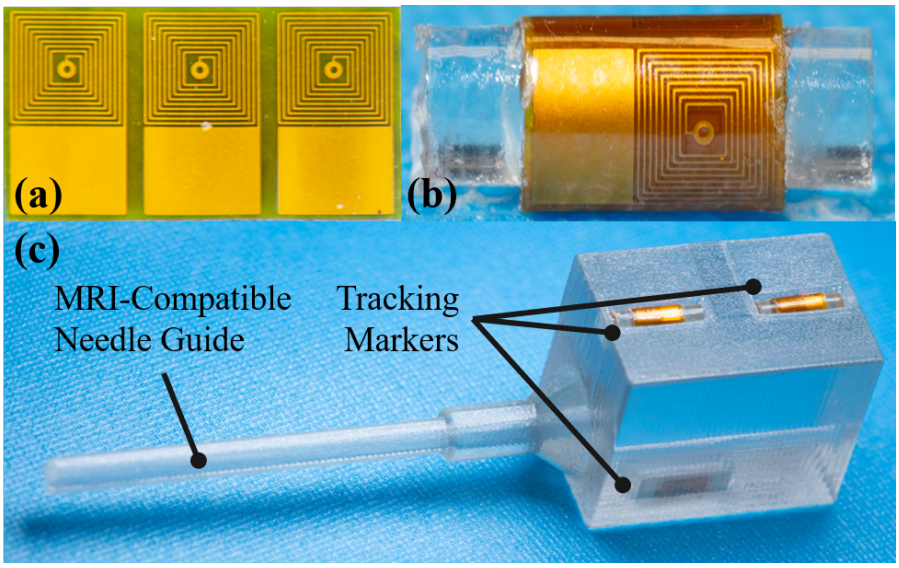


Fig. 5.13: **(a)** Original planar form of three monolithic resonant circuits. **(b)** Curved form of the resonant circuits forming the omnidirectional marker. **(c)** Three markers embedded inside an MRI-compatible needle guide in order to provide 6-DoFs positional tracking.

5.6.2 Marker Orientation Dependency

The marker orientation dependency has been evaluated with the marker embedded inside a 3D-printed MRI-compatible block. The block was mounted on a 3D-printed plastic protractor stand that can rotate in steps of 10° around X-, Y- and Z-axis from 0° to 90°. The protractor stand was affixed at the isocenter inside a clinical 1.5T MRI Scanner (Signa Explorer, GE Healthcare) with a standard 8-receiver imaging head coil. The MR images at different orientations were acquired using fast spoiled gradient-echo (FSPGR) sequence, with parameters: TE = 2.144 ms, TR = 7.185 ms, slice thickness = 2 mm, matrix size = 200 × 200, flip angle = 1°, FOV = 120 mm × 120 mm, pixel spacing = 0.6 mm. Image distortion caused by gradient nonlinearities was compensated with gradient warp correction. DICOM images were extracted and processed in MATLAB. The initial signal-to-background ratio is ∼57. Although the marker signal drops when the rotational angle around X-, and Z-directions increases, the signal remains bright and can be unambiguously identified from the background, as shown in **Fig. 5.14**. When the axis of the marker aligns with X- or Y-direction, the ratio drops to ∼30.

5.6.3 MR Imaging Test

MRI-based experiments have been conducted to validate the capability for visualizing subjects and instruments in a single MRI image. An MRI-compatible needle guide [15]

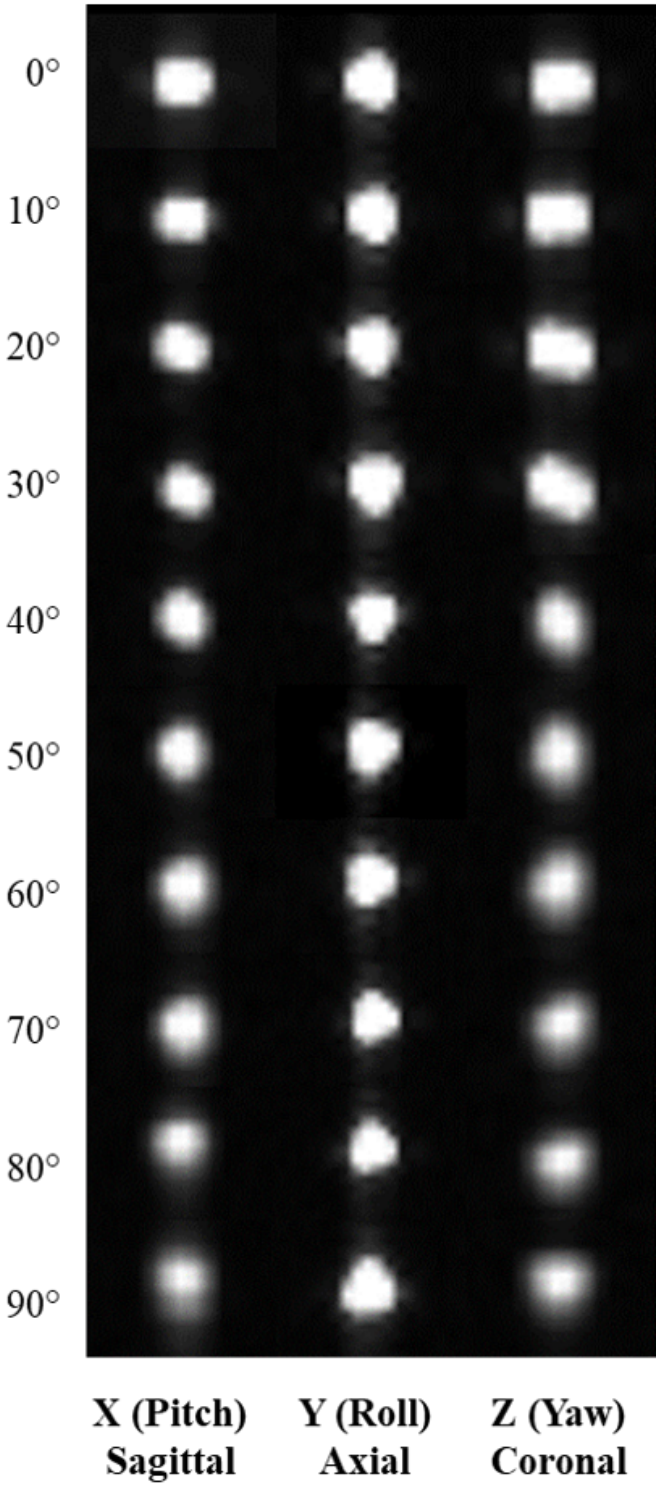


Fig. 5.14: Coronal, sagittal, and axial gradient echo images of the marker at low flip angle (1°) acquired at different orientations using a fast spoiled gradient-echo (FSPGR) sequence (with settings TE = 2.144 ms, TR = 7.185 ms, slice thickness = 2 mm, matrix size = 200 × 200, flip angle = 1°, FOV = 120 mm × 120 mm, pixel spacing = 0.6 mm). Marker signal can be unambiguously identified from the background at any orientation. The max. and min. signal-to-background ratio are, respectively, around 57 and 30.

(**Fig. 5.13c**) embedded with three markers was mounted on the forehead of the subject, who was scanned with a supine position. Fast gradient echo pulse sequence (TE = 2.692 ms, TR = 5.853 ms, slice thickness = 1 mm, matrix size = 256 × 256, pixel spacing = 1 mm) was applied to retrieve MR images on the aforesaid 1.5T MRI scanner. The coronal and sagittal MR images are shown in **Fig. 5.15a** and **Fig. 5.15b**, respectively. The performance plots demonstrate that the signal level from the marker is much higher than the noise from both the head and background (SNR $>$ 18), implying the marker can be easily visualized or tracked both visually and with a computer algorithm.

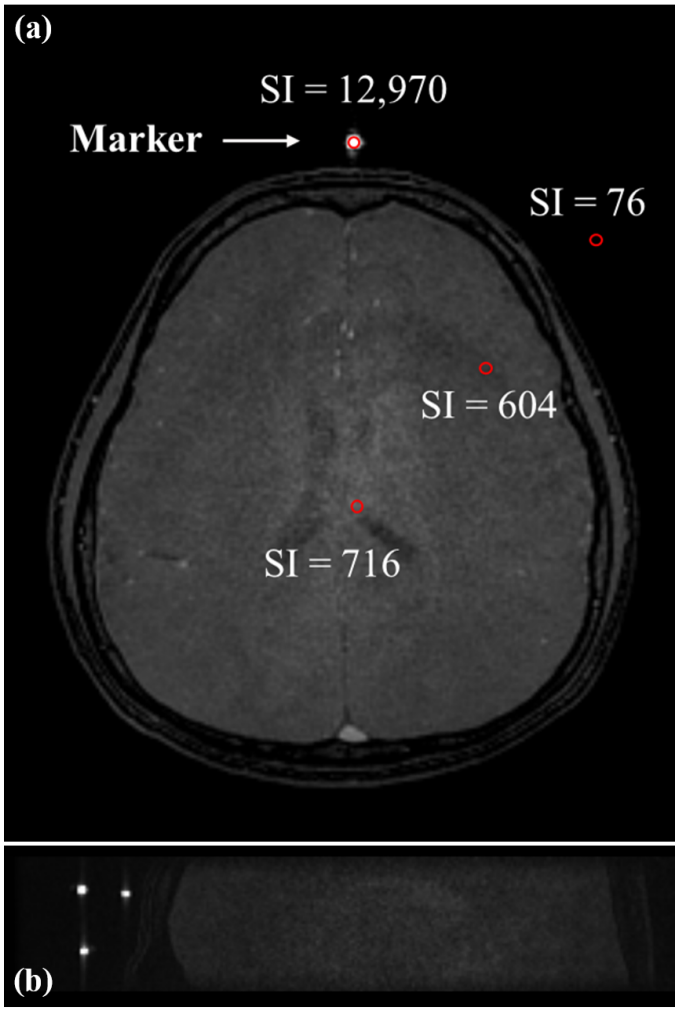


Fig. 5.15: (a) Low flip-angle (1°) MR scanning image with average signal intensities (SI) of ROI (2 × 2 pixels) indexed. The marker is affixed to a human head with a 3D-printed fixture. The average marker signal is about 18 times stronger than the signal from brain, and about 170 times stronger than background noise. **(b)** Image of the MR-compatible needle guide placed next to the head. Three bright spots can be clearly visualized and identified from the head.

5.6.4 Real-time Tracking Test

A real-time tracking test of three wireless markers has been conducted by continuously performing 2D imaging. The markers were attached along a catheter which was continuously steered. Gradient echo sequences, which have the parameters TR/TE = 3.9/1.168 ms, flip angle = 1°, slice thickness = 30 mm, are applied to enable fast 2D imaging (0.2 s). The slice could cover the three markers and enable 2D projection imaging (**Fig. 5.16a**). **Fig. 5.16b** gives the time series maximum intensity projection (MIP) along the y-axis in a period of 30 s, with a sampling frequency of 5 Hz. In the MR images, the signals of markers had high contrast against the background, which could be automatically detected to provide real-time positional feedback in MRI (**Fig. 5.16c**). Orthogonal 2D projection imaging could be further interleaved to achieve 3D localization of the three markers.

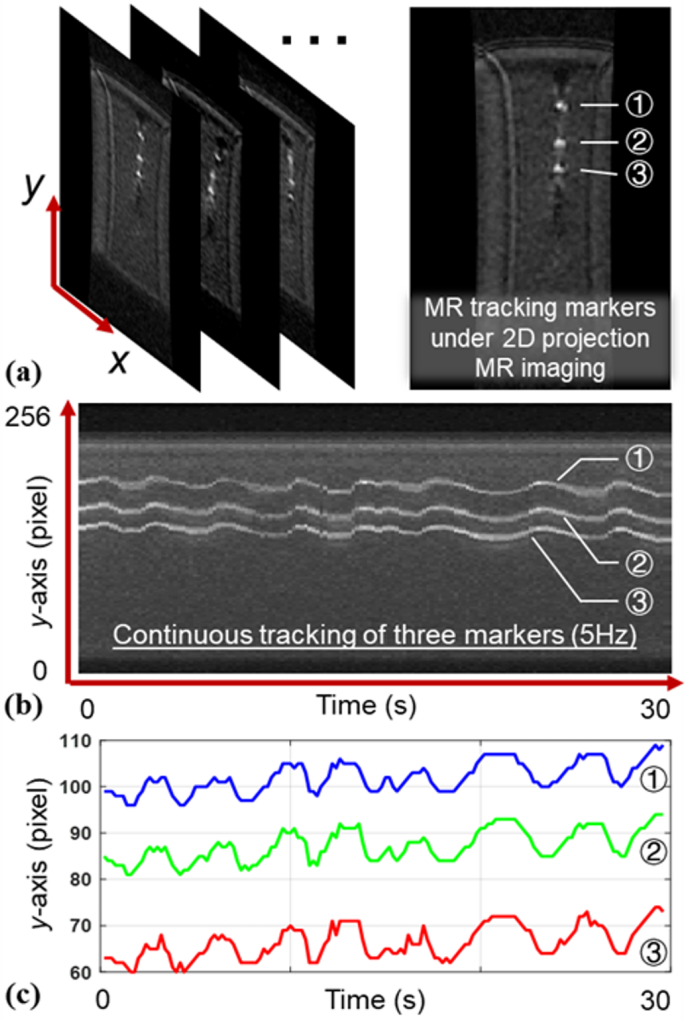


Fig. 5.16: **(a)** MR markers tracked using 2-D projection sequence. **(b)** Projection to the y-axis showing the continuous tracking of the three markers at update rate of 5 Hz. **(c)** Variation of their corresponding y-axis position over time.

5.7 Conclusion

This chapter presents the design of a soft robotic system for intra-op MRI-guided transoral laser microsurgery. The robot comprises two actively-steerable soft continuum segments which are driven by micro-volume liquid flow (<0.004 mL) to the internal elastomer chambers, which are individually integrated with spring reinforcement constraints. The spring reinforcements were designed and optimized through FEA to balance between robot bending stiffness and maximum local stresses, thus enhancing robot response with minimized hysteresis while maintaining adequate robot durability. The distal laser manipulator actuation is effectively decoupled from the active bending section, allowing stable and repeatable kinematics, and thus high-accuracy laser control. Integrated with an optical laser fiber, the robot enabled endoscopic laser delivery and operating in the oropharyngeal region with sub-millimeter accuracy. A patient-specific dental guard was designed to create an open-jaw cavity for robot anchorage and access of auxiliary instruments such as a fiberscope.

The active bending section offered a 60º omni-directional bending workspace for coarse navigation to the region of interest. The distal laser manipulator allowed fine laser beam pan-tilting (±26º) in a local target area, with a maximum scanning area of 15×15 mm^2 at a 15 mm projection distance. Combined with a learning-based controller, accurate laser spot manipulation can be achieved, with a mean tracking error <200 μm in path following tasks (**Chapter 4**).

A two-cylinder hydraulic motor was integrated to provide a DoF of advancement/retraction, with the design parameters optimized based on a dynamic model for hydraulic transmission. Wireless MR markers were fabricated and validated to enable real-time positional tracking in MRI.

Chapter 6

Preclinical Validation Studies

6.1 Introduction

With the unique capability of thermal imaging in 3D, MRI-guided laser ablative therapy is attracting increasing attention. It can provide real-time measurement of heat diffusion in tissue, enabling intra-op assessment of the ablation margins, which can play an important role in the treatment of oral and pharyngeal cancers with particular concern for function preservation. Aside from MR thermometry, anatomical MR imaging also provides major and proven benefits for delineating cancerous, healthy, and ablated tissue, which can be utilized for pre-op planning and intra-op tracking of resection completeness and depth. When paired with MR thermometry (and seamlessly registered in the same reference coordinate frame), there is tangible clinical motivation to utilize MRI in laser-based otolaryngologic procedures. The proposed MR-safe soft robotic system (**Chapter 5**) enables the introduction of intra-op MRI guidance to evaluate the laser/tissue interaction process. This could protect critical healthy structures, while eliminating the need for FSA, thus smoothening the surgical workflow and shortening operation time.

In this chapter, two eye-to-hand vision-based controllers (**Chapter 4**), i.e., epipolar geometry-based and learning-based, were validated. *Ex-vivo* laser ablation tests were performed, both in laboratory (**Section 6.2**) and MRI (**Section 6.4**), using the epipolar geometry-based controller (**Section 4.4**). *Ex-vivo* tissue ablation and a cadaveric head-and-neck trial were carried out in MRI (**Section 6.5**), where the learning-based controller (**Section 4.5**) was employed for laser spot steering control. MR thermometry was performed to monitor the tissue ablation margin and thermal diffusion intra-operatively. The

key contributions are listed as follows:

1. Experimental validation of the proposed laser spot visual servoing, as well as the ablation performance through lab-based and MRI-based *ex-vivo* tests;

2. Experimental validation of intra-op MRI-guided robotic laser ablation on a cadaveric head-and-neck model. Intra-op MR thermometry was used to monitor the laser ablation process.

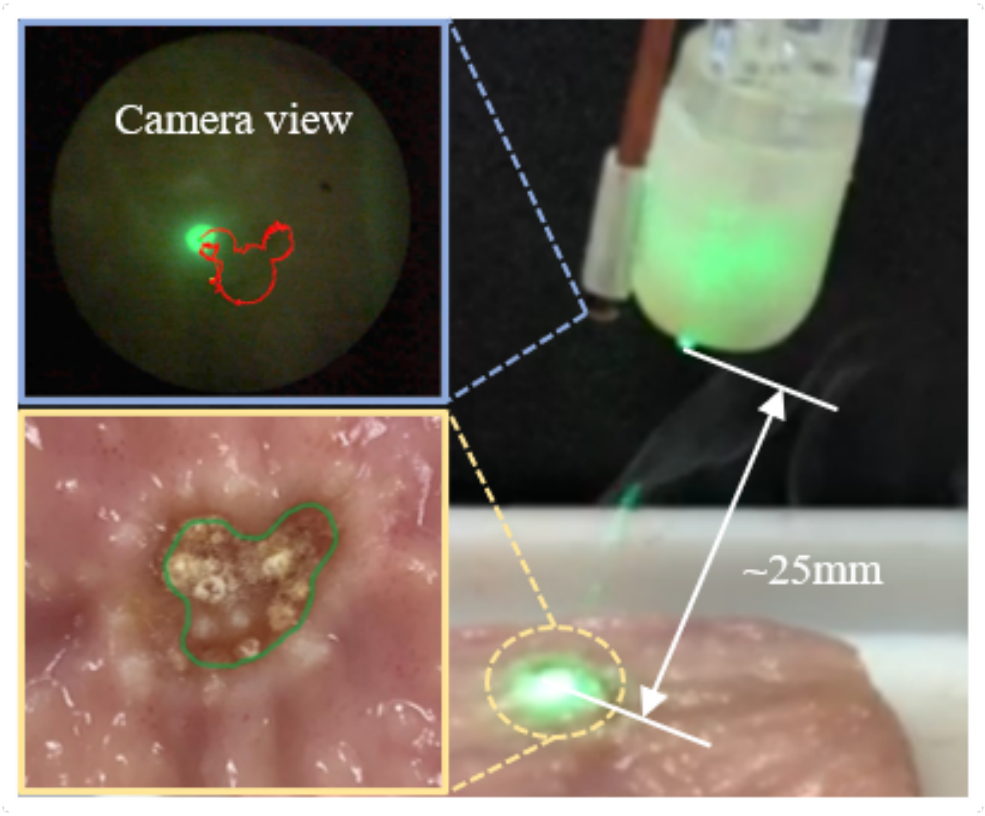


Fig. 6.1: Laser spot path following and ablation with *ex-vivo* swine stomach tissue. A camera view is presented showing the image of the laser spot and its footprint (left upper). A post-ablation image of the tissue is also shown (left bottom).

6.2 Lab-based *Ex-vivo* Laser Ablation Test

The laser steering system, including the visual servoing Controller (**Section 4.4.2**), has been validated by a laser ablation test on *ex-vivo* swine stomach tissue (**Fig. 6.1**). The tissue was taken out of the refrigerator and naturally defrosted at room temperature of 20 °C for two hours before the test. A 1,550-nm laser source with a high water-absorption rate is used to provide efficient vaporization in the *ex-vivo* ablation test. The laser source power is 3W

in continuous-wave mode. A pilot green laser (532 nm) is also coupled to provide visual feedback. To simulate the clinical setting in transoral laser surgery, the soft laser manipulator is mounted in a downward position, facing toward the tissue surface with a tilted angle of $\sim 30°$. Note that this tilting angle and the projection distance from the robot tip to the tissue surface are totally unknown to the controller. However, accurate path following of the laser spot was still achieved, with the tracking trajectory depicted in the camera view. Small noises in the trajectory are introduced by the light scattering on the tissue surface, resulting in an irregular shape of the laser spot which affects its detection. We can observe a "Mickey Mouse" pattern on the surface of ablated tissue. The trajectory is distorted with respect to that in the camera view, due to the shrinkage of tissue being vaporized and the irregular surface geometry.

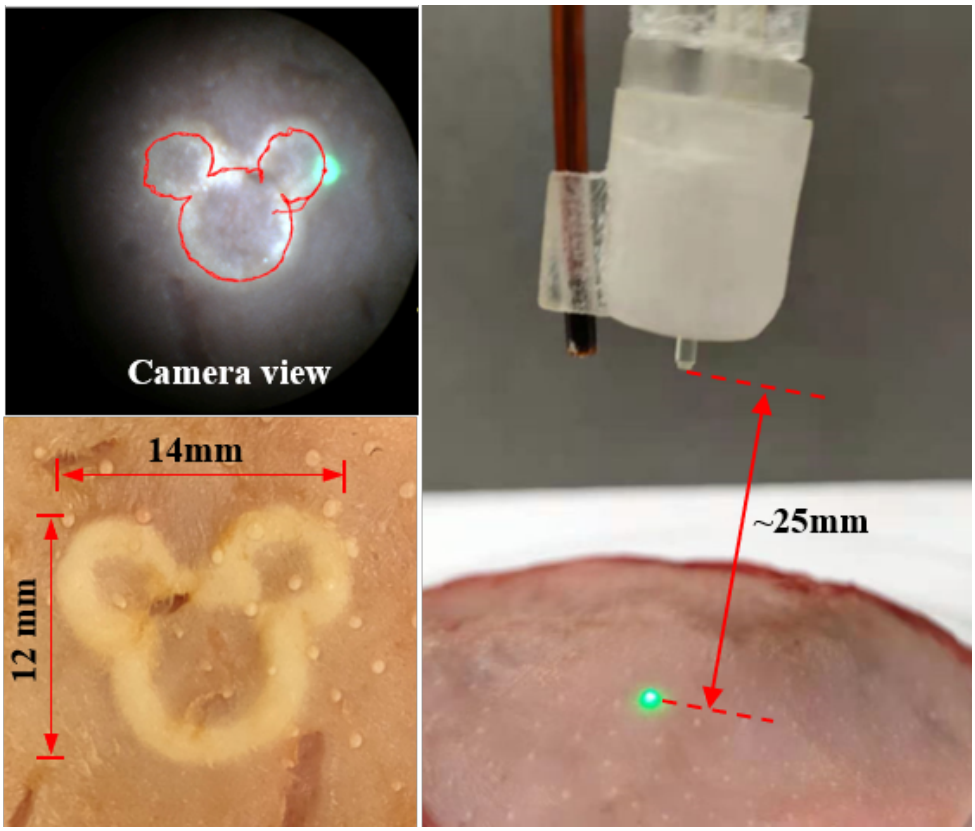


Fig. 6.2: Laser ablation test on *ex-vivo* pig tongue tissue. A camera view is presented showing the image of the laser spot and its footprint (left upper). A post-ablation image of tissue is shown with the ablated pattern (left bottom).

Another ablation test has been conducted with *ex-vivo* swine tongue tissue, which contains less fluid and has higher stiffness. The laser source power was increased to 4W so that an obvious ablation effect could be observed even with a faster laser steering speed. As shown in **Fig. 6.2**, the swine tongue tissue showed less shrinkage upon ablation than swine

stomach tissue (**Fig. 6.1**). A "Mickey Mouse" ablated path (14×12 mm) can be observed on the tissue surface, as well as the camera view. The RMS tracking error during the ablation is about 5.75 pixels (∼201.3 μm).

6.3 MRI-compatibility Test

SNR testing has become a common protocol for MRI compatibility evaluation in the MRI robotics field. During the testing, MR scanning was performed with both spin echo and gradient echo pulse sequences, as suggested in the ASTM F2119 standard [199]. SNR is calculated as the ratio of signal intensity in the ROI to the noise intensity in the periphery, according to the National Electrical Manufacturers Association standard [200]. As shown in **Fig. 6.3**, SNR is measured as the ratio of the mean intensity inside the red circle and the standard deviation of intensities inside the green circle. The percentage of SNR changes compared to the control image can be further calculated.

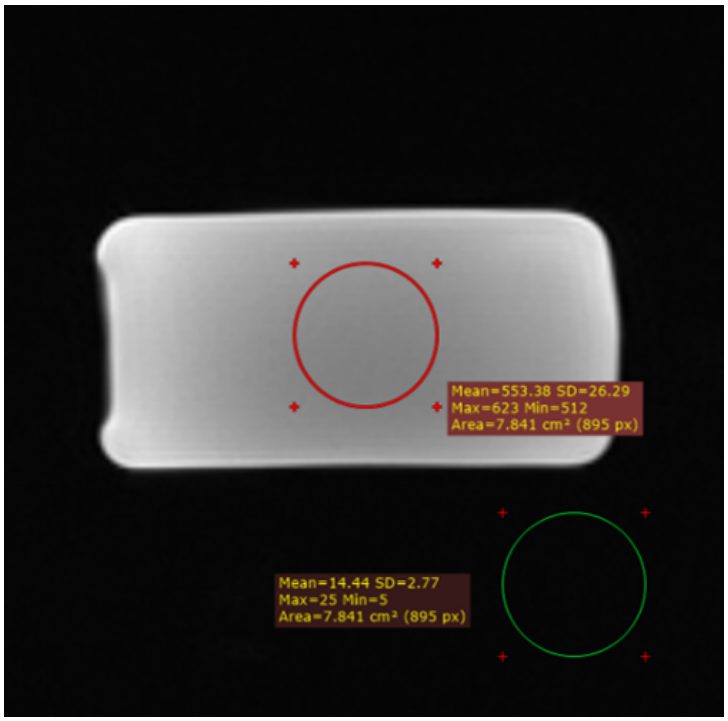


Fig. 6.3: MR T2 image of a water phantom.

6.3.1 Soft Laser Manipulator

The MRI compatibility of the soft laser manipulator has been investigated by the SNR test, comparing the reduction of SNR in MR images of a control phantom (J8931, J. M. Specialty

Parts, USA). T1 weighted gradient echo (GRE) images and T2-weighted fast spin echo (FSE) images were acquired with the robot operated inside a 1.5T MRI scanner. The MR images acquired under four operation conditions (i.e., only phantom as the baseline, robot static with power off, robot static with power on, and robot in operation) were compared [15], as shown in **Fig. 6.4**. There is no artifact observed in the phantom images. The loss of normalized SNR was smaller than 5% even when the robot was fully operated, which indicates that the laser steering system is MR compatible, and allows simultaneous robot manipulation and MR scanning.

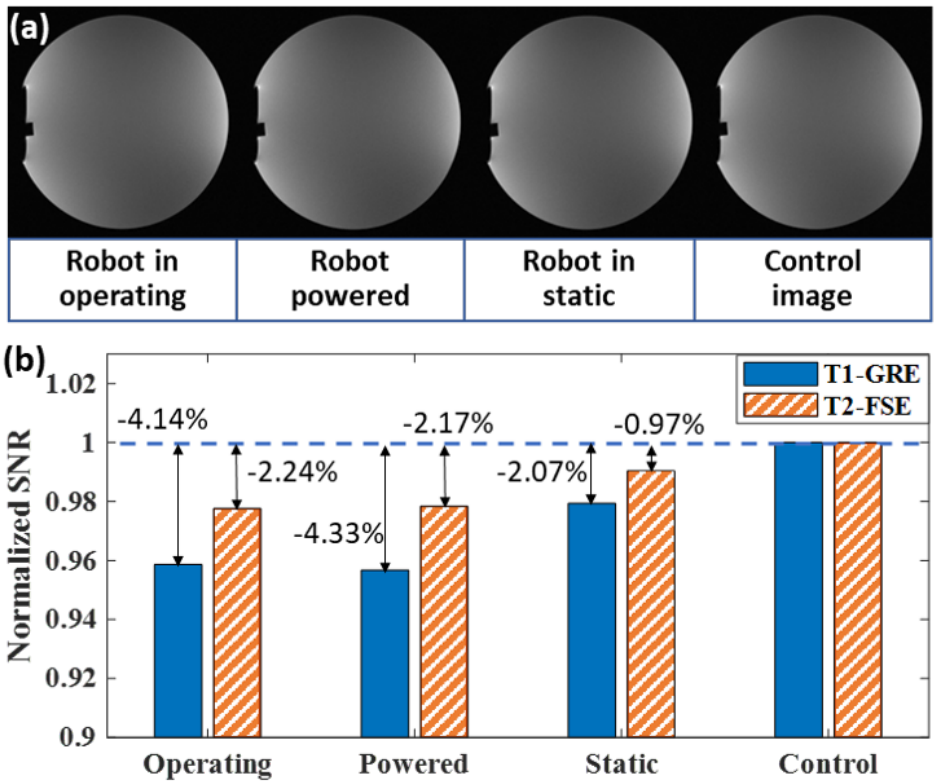


Fig. 6.4: SNR test of laser manipulator in T1-GRE and T2-FSE. **(a)** MR images (T1-GRE) of an MRI phantom placed aside the robot under four operating conditions. **(b)** Normalized deviation in artifact percentage calculated by SNR.

6.3.2 Transoral Robotic System

MRI compatibility of the entire robot system has been validated with a 1.5T MRI scanner (SIGNA™ HDxt, GE Healthcare) (**Fig. 6.5a**). Although the robot itself is made of non-metallic materials, wireless MR markers fabricated with micro circuits are integrated on the robot platform, which were also included in the MR-compatibility test. The robot was placed beside a rectangular water phantom and translated to the isocenter of the scanner. As suggested in the ASTM F2119 standard, T1-weighted GRE and T2-weighted FSE

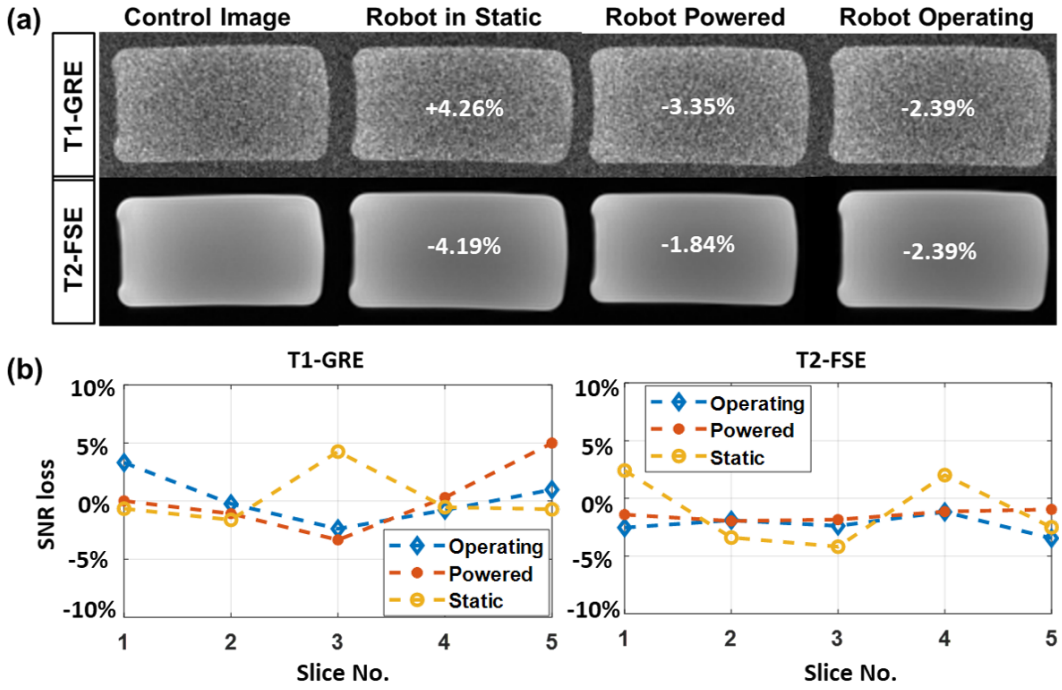


Fig. 6.5: SNR test of transoral robotic system in T1-GRE and T2-FSE. **(a)** MR images of a water phantom placed beside the robot, showing the normalized SNR loss during different states of robot operation. The robot is not present in the control case. **(b)** Normalized deviation in artifact percentage calculated by SNR.

sequences were both implemented for image acquisition and comparison. A control image of the phantom alone was acquired as the baseline. After positioning the robot next to the phantom, MR images were acquired with robot operated under three different conditions, which are: i) Static: robot powered off; ii) Powered: robot at rest, but the hydraulic and electric power are on; iii) Operating: robot is being actuated. The SNR in MR images was calculated following the guidelines [200] by NEMA. SNR losses compared to the control image were calculated over five image slices, which cover the majority of the water phantom. The mean SNR variation in T1-GRE and T2-FSE images with robot operated are 0.18% and -2.29%, respectively. An SNR loss up to 10% is suggested to be acceptable [201]. The SNR loss compared to the control image is within 5% even with the robot in full motion, as shown in **Fig. 6.5b**. Image intensity variations of 30% or above are considered as artifacts referring to the ASTM standard test method [199], and will appear in white pixels. No observable image artifacts are found within the phantom area, which indicates that the robot operation at the isocenter generates zero EM interference to tissue imaging.

6.4 MRI-based *Ex-vivo* Laser Ablation Test

Laser ablation experiments have been conducted in the MRI environment with *ex-vivo* swine tissues. The robot is controlled tele-operatively with hydraulic transmission over 10-meters long pipelines, which are channeled through the waveguide between the scan room and control room.

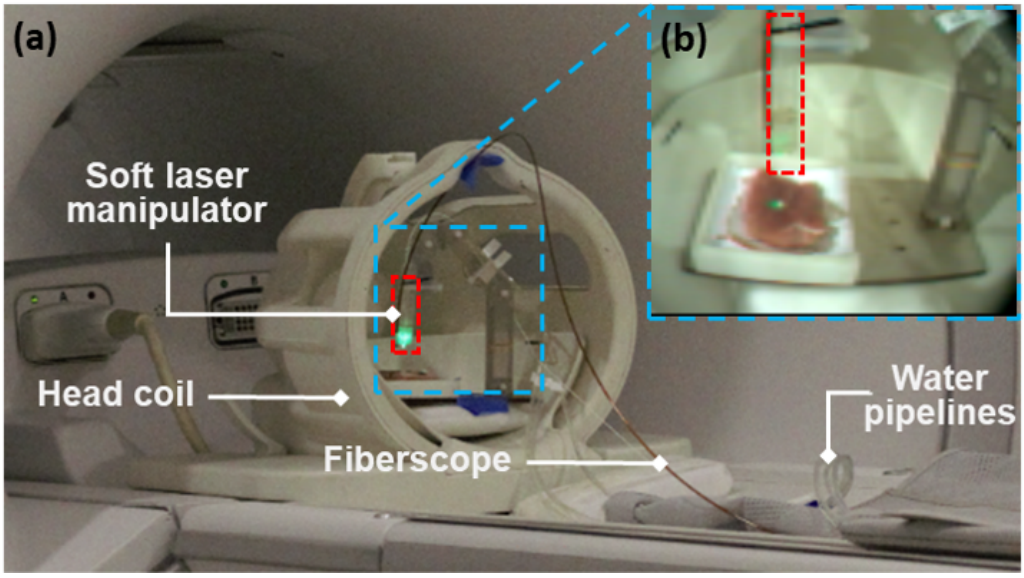


Fig. 6.6: (a) Experimental setup of the soft laser manipulator in a 1.5T MRI scanner. **(b)** Image captured by an MR-conditional camera.

6.4.1 Laser Spot Path Following

Experimental setup of the soft laser manipulator in a 1.5T MRI scanner is shown in **Fig. 6.6a**, with a head coil utilized for MR image acquisition. **Fig. 6.6b** is a local view showing the laser spot projected on swine tissue. A "Mickey Mouse" pattern was inputted as the laser spot trajectory. The *ex-vivo* tissue was ablated with a 4W 1,550nm laser, repeated for 3 loops. To incrementally monitor the laser ablation effect, 3D T1-GRE acquisition with parameters: TR/TE = 120/5.65 ms, Flip Angle = 40°, Slice Thickness = 6 mm, Pixel Spacing = 0.469×0.469 mm, was continuously performed with 18s temporal resolution. After the laser ablation with three repeated loops, we can observe a shadowed "Mickey Mouse" shape on the MR image (**Fig. 6.7**). The obvious difference between the ablated pattern and the input "Mickey Mouse" pattern results from the hyperthermia-related shrinkage of *ex-vivo* soft tissue once after the laser ablation. This difference does not reflect the laser targeting error, which can be proved by the path following experiment (**Chapter 4**). The tissue

shrinkage upon heating was mainly due to dehydration, which is significant for mucosa tissue containing abundant fluid. Currently, a routine pulse sequence was used to observe the tissue physiological change after ablation. Each acquisition of MR image takes about 18s; meanwhile, tissue shrinkage could happen because of the low temperature (20 °C) in the MRI room. In future work, the MRI sequences will be customized to achieve real-time and high-resolution imaging in the small ablation region.

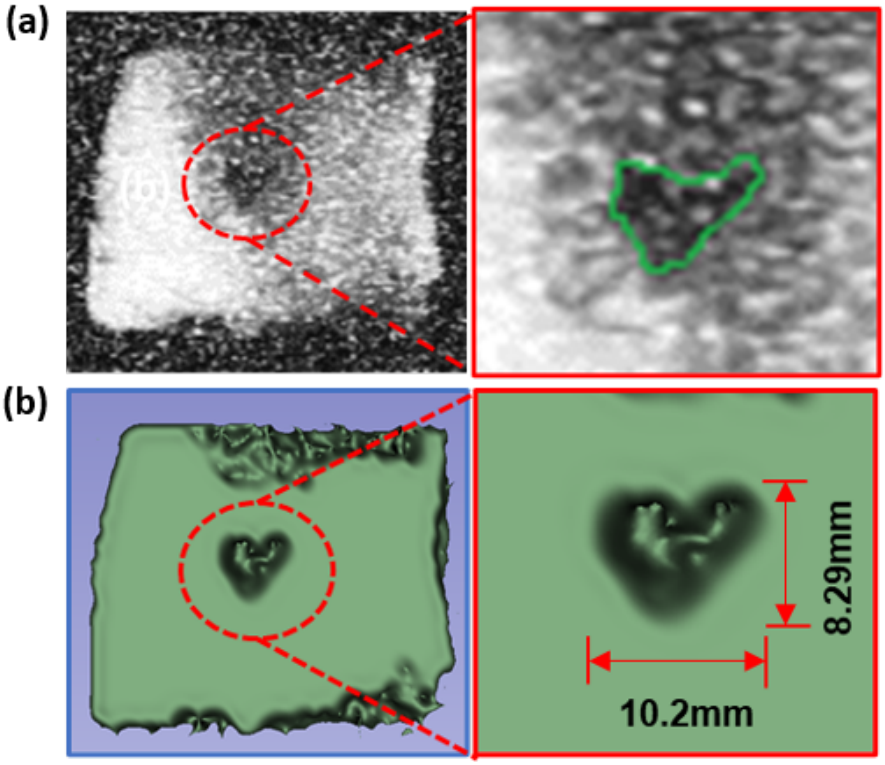


Fig. 6.7: **(a)** MR Image (T1-GRE) of *ex-vivo* swine tissue after ablation with 3 repeated loops showing a "Mickey Mouse" like pattern. **(b)** Ablated swine tissue model reconstructed from 6 slices of MR images.

6.4.2 Thermal Diffusion Monitoring

MRI enables temperature measurement through the proton resonance frequency (PRF) shift [202], which can be calculated from the phase images. Prior to heating, baseline images are acquired to determine the reference phase image, ϕ_{ref}. At each time step t, the temperature difference ΔT can be computed based on the difference between phase images ϕ_t and ϕ_{ref}:

$$\begin{aligned} \Delta T &= (\phi_t - \phi_{ref}) \cdot k \\ k &= \left(\frac{\gamma}{2\pi} \cdot \alpha \cdot B_0 \cdot \mathrm{TE}\right)^{-1} \end{aligned} \tag{6.1}$$

where the coefficient k is determined by the gyromagnetic ratio γ, PRF temperature coefficient α, magnetic field strength B_0, and echo time TE. Although these parameters could be obtained from prior arts, calibrating the coefficient k from experiments allows a more accurate temperature measurement. This can be conducted by matching the temperature sensing respectively using MR thermometry and a fiber-optic fluorescent temperature sensor (PRB-MR1 OSENSA, Canada), from a bottle of hot water that is cooling down naturally. As shown in **Fig. 6.8**, two sensing modalities show a similar trend by tuning the coefficient k.

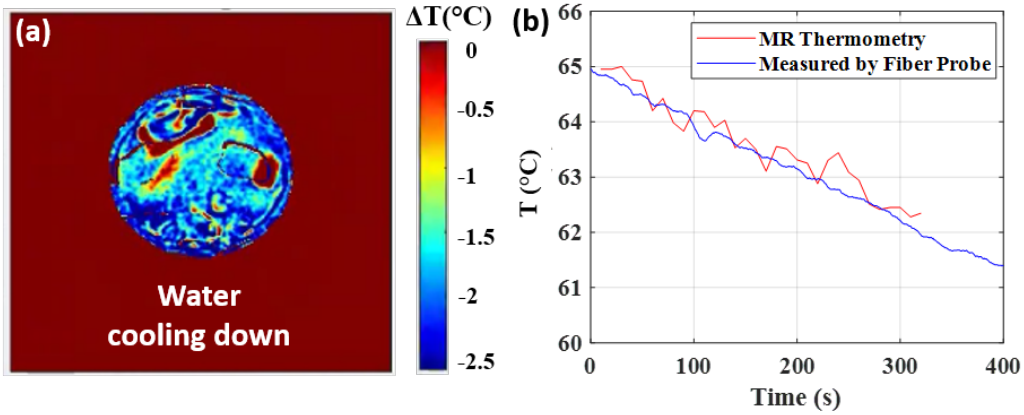


Fig. 6.8: **(a)** Map of temperature change measured from MR thermometry. **(b)** Calibration of MR thermometry with a fiber-optic temperature sensor.

***Ex-vivo* Laser Ablation Trial:**

To evaluate the feasibility of proposed soft robot tele-manipulation, laser beam steering, and its intra-op MRI settings, an *ex-vivo* laser ablation test has been conducted to monitor laser-induced changes on tissue. A head coil was utilized to house the robot and acquire MR images (**Fig. 6.9a**). The robot was remotely controlled to aim the laser spot along a near-circular trajectory. A continuous-wave laser source (1,550 nm, 3 W) with a high absorption rate in water was used. A green laser with a wavelength of 520 nm was also combined into the ablation laser beam to offer a visual indication of the ablation laser location using an MR-compatible fiberscope. A gradient echo sequence with parameters: TR/TE = 45/13.8 ms, flip angle = 12°, slice thickness = 8 mm, number of slices = 1, pixel size = 0.625×0.625 mm, field of view = 160×160 mm was implemented to acquire both a magnitude and phase image at a time resolution of 10 s. This could be further accelerated by introducing advanced pulse sequences [203]. The magnitude image directly gives the update of tissue morphology during the laser ablation. As shown in **Fig. 6.9b**, the laser ablation depth can be clearly seen and measured (2.5 mm) from the MR images. Based on the

phase images, the temperature increments relative to the baseline acquired before ablation could be calculated using the PRFS method [202]. The accumulated thermal diffusion on this cross-sectional slice overlaid on the anatomical tissue image is shown in **Fig. 6.9c**, with the laser spot region generating the highest temperature increase. To cause irreversible cell death, the tissue must be heated to >42 °C [204]. Thermal dosage to the tissue can be monitored accordingly, which could offer useful guidance to surgeons when controlling the laser projection and intensity. With advanced MR imaging sequences, multi slices of images can be taken in parallel [205] or orthogonally [206], thus providing 3D monitoring of the tissue thermal effect during laser ablation.

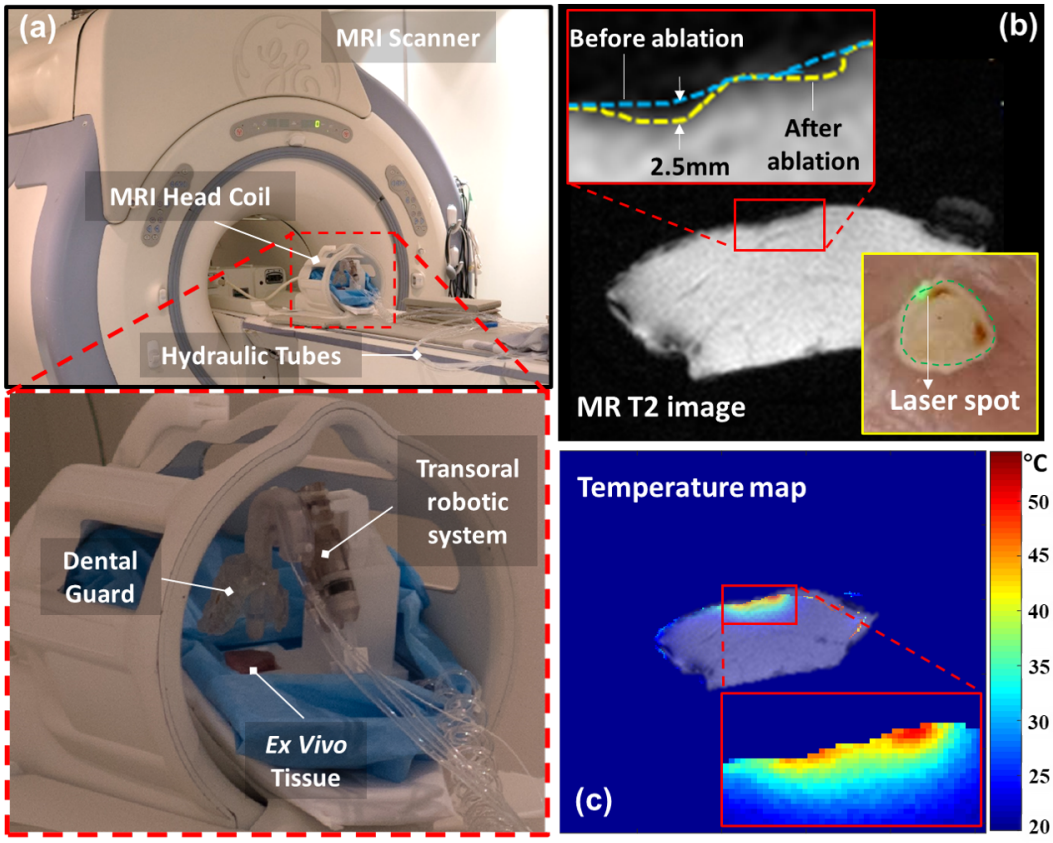


Fig. 6.9: MRI-based *ex-vivo* laser ablation test. **(a)** Robot setup in the MRI scanner, with an 8-channel transmit/receive head coil for image acquisition. **(b)** Observable ablation depth found in MR T2 images of swine tissue before and after ablation, with the laser spot controlled to follow a near-circular path. **(c)** MR thermometry applied to monitor thermal diffusion and accumulated temperature increments in tissue due to laser ablation.

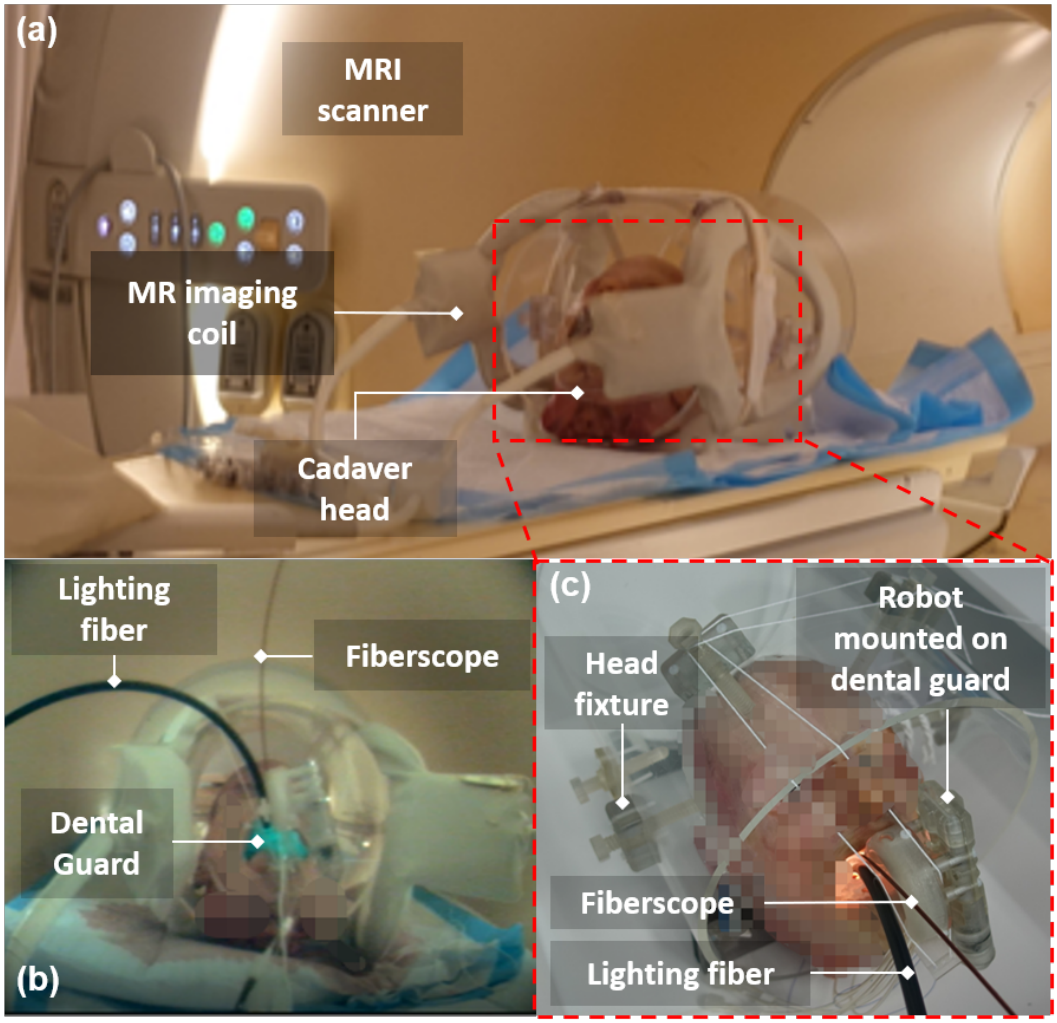


Fig. 6.10: MRI-based cadaver trial setup. **(a)** Experimental setup in the 3T MRI scanner (Philips Achieva). **(b)** Image acquired by an MRI-compatible camera, with robot moved to the isocenter. **(c)** Cadaveric head fixed in an acrylic tube mount. A fiberscope and lighting guide were inserted through the additional channel of the dental guard.

6.5 MRI-based Cadaveric Trial

To validate the proposed robotic platform in human anatomy, a cadaver trial of transoral laser ablation has been conducted in a 3T MRI scanner (Philips Achieva). The experimental setup is shown in **Fig. 6.10a-b**. A cadaver head was defrosted 24 hours before the test and fixed in an acrylic tube mount using plastic screws (**Fig. 6.10c**). The dental guard with an open bite of 35 mm was fitted in the mouth to create an open-jaw cavity. In addition, the tongue was retracted anteriorly with a 2/0 silk suture and an L-shape 3D-printed depressor (90×20×2 mm) was employed to push down the tongue and further increase the workspace. The depressor was tightened to the head fixture using fabric strings. The robot was then

mounted on the dental guard and fixed through mechanical interlocks. A 10-meter-long MR-compatible fiberscope, with an additional lighting fiber for illumination, provided the endoscopic view for simple navigation.

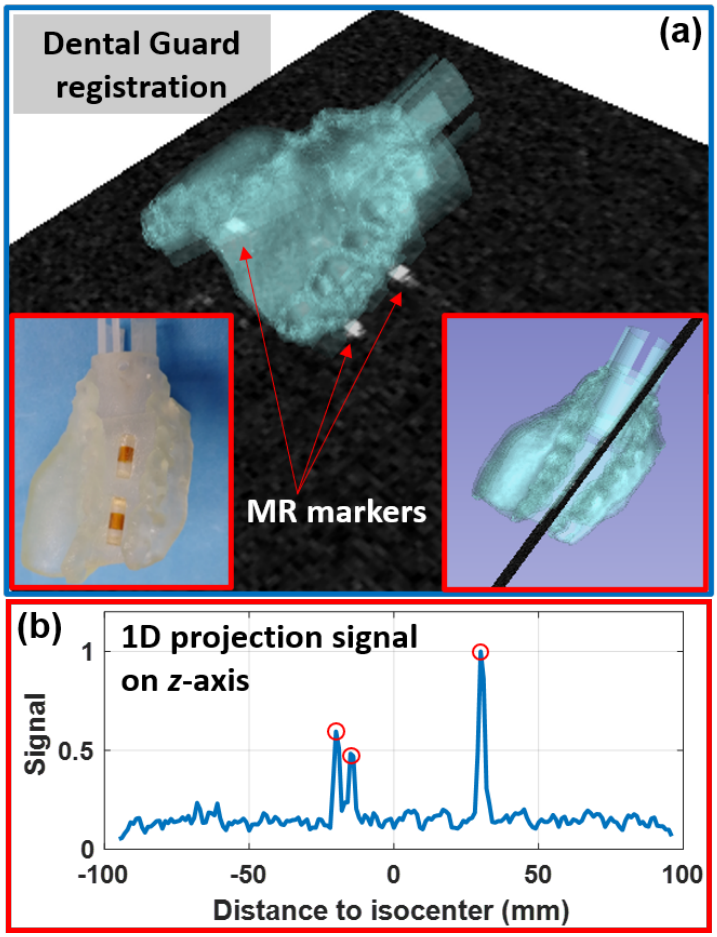


Fig. 6.11: MRI-based cadaver trial setup and registration. **(a)** Dental guard registered based on scanned MR tracking markers. **(b)** Corresponding z-axis 1-D projection scan for localizing the three MR tracking markers.

6.5.1 Image Registration with Dental Guard

Three MR fiducial markers **Section 5.6** fabricated from capsulized Gadolinium-doped water (concentration: 10 mM) were integrated with the dental guard to localize the robot base in MRI coordinates (**Fig. 6.11(a)**). Specifically-designed miniature wireless MR resonators [144] were integrated into the fiducial markers to locally amplify their signal, providing high SNR even under low flip angle excitation (1°). The resonators are flexible printed LC circuits tuned to the Larmor frequency (63.87 MHz). The fiducial markers show high contrast in anatomical MR images, thus enabling positional localization. Based on the pre-op 3D imaging, positions of the markers can be identified from the anatomical background. With

3D positions of three markers inside the dental guard, point-based rigid registration can be performed to locate the dental guard as well as the robot base in MR image coordinates. Alternatively, using 1D-projection gradient readouts, the marker positions could even be obtained in real-time (∼30 Hz) (**Fig. 6.10b**).

6.5.2 Laser Spot Path Following

The robot was positioned in the oropharynx of the cadaver, and a laser ablation trajectory was defined by the operator by "sketching" in the fiberscope camera view (**Fig. 6.12a**). Low-temperature laser visualized by MR thermometry could be applied to ensure that the ablation location is consistent with the pre-op plan made in the T2-weighted model. Control of the laser spot on the mucosa along the prescribed trajectory was achieved through machine learning-based visual servoing (**Section 4.5**). Accurate laser spot steering was achieved with a mean tracking error of <3 pixels (∼0.2 mm). The tracking performance of the first two ablation cycles is shown in **Fig. 6.12b**.

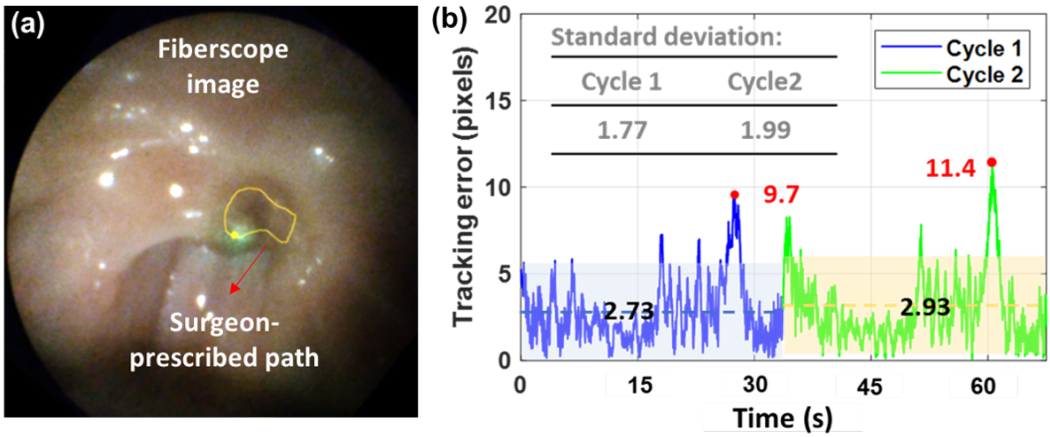


Fig. 6.12: MRI-based cadaver trial ablation result. **(a)** Laser spot controlled to follow a trajectory defined in the fiberscope view. **(b)** Tracking errors of laser spot path following over two initial cycles.

6.5.3 Intra-op MR Thermometry

For pre-operative planning, a T1-weighted imaging model of the head and neck region was acquired, using 3D gradient echo pulse sequence with parameters: TR/TE = 7.7/3.6 ms, slice thickness = 1.1 mm, flip angle = 8°, number of slices = 280, pixel size = 1.03×1.03 mm, field of view = 230×230 mm. The intra-op scanning plane for temperature measurement was determined based on the reconstructed 3D head volume (**Fig. 6.13a**).

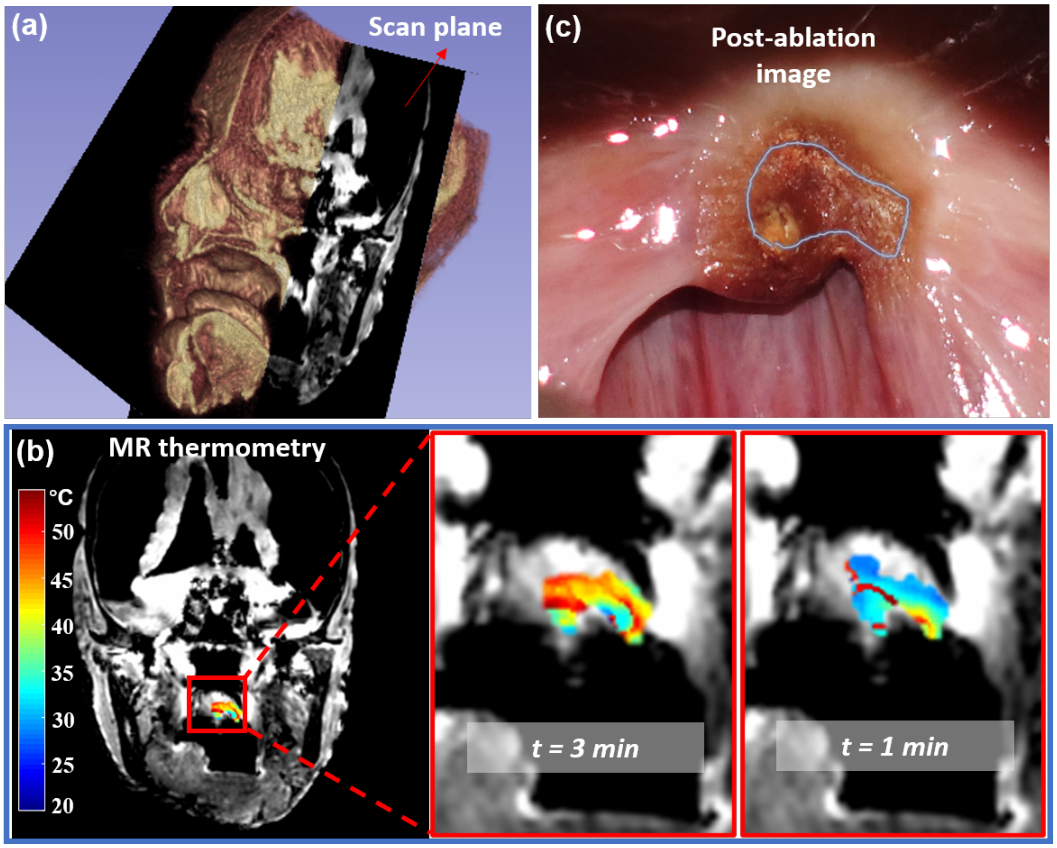


Fig. 6.13: MRI-based cadaver trial ablation result and MR thermometry. **(a)** Imaging slice prescribed for intra-op MR thermometry superimposed on the cadaver pre-operative 3D model. **(b)** Temperature increments map overlaid on the anatomical MR image after 1 and 3 minutes of ablation. **(c)** Post-ablation image of the tissue.

Once the lesion region was confirmed in the MR coordinates, the ablation laser was turned on and steered along the trajectory over multiple cycles (∼35 s each) for 3 minutes, with a power of 3 W. Meanwhile, anatomical and temperature interleaved imaging was performed using a routine 2D fast-spin-echo sequence with parameters: TR/TE= 15/5 ms, slice thickness = 2 mm, number of slices = 14, flip angle = 18°, pixel size = 0.479×0.479 mm, field of view = 230×230 mm, at a time resolution of 20 s. The update rate can be accelerated to even sub-seconds by tuning the sequence parameters, e.g., reducing the field of view and number of slices. As shown in **Fig. 6.13b** a map of temperature increments can be obtained to reflect the thermal effect on tissue upon laser energy absorption. To cause irreversible cell death, the tissue must be heated to >42 °C [204]. A post-ablation image was taken with a handheld camera (**Fig. 6.13c**), which showed an ablation pattern with a similar boundary to the predefined trajectory. There would also be undesired coagulation occurring around the center of region. In future work, laser pulse parameters (e.g., duration, power) could be modulated online based on the MR thermometry feedback to intensify

the tissue vaporization in a short duration, thus minimizing the undesired coagulation. Optical-fiber-guided thulium laser would be an alternative that can provide improved cutting and ablation properties comparable to that of the widely-used CO_2 laser.

6.6 Conclusion

Preclinical trials have been conducted to evaluate the robot performance both with *ex-vivo* swine tissue and a cadaver model. During *ex-vivo* tissue testing, the use of anatomical T2 imaging and MR thermometry on monitoring laser ablation are investigated, which showed the potential for measuring laser ablation depth and temperature diffusion region. A cadaver trial was performed to validate the robot setup/anchorage, intra-op MRI guidance, as well as visual servo-based laser beam manipulation in actual human anatomy. The robot achieved a precise and automated path following of the laser spot along a trajectory defined *in-situ* by the surgeon. During the ablation, MR thermometry was performed to measure the thermal diffusion process in the ablated lesion. The robot provided a stable platform during laser ablation due to the rigid anchorage of its base with the dental guard. In the current practice of treating oropharyngeal tumors, patients are placed under general anesthesia with muscle paralysis. Therefore, the proposed system is expected to maintain similar levels of motion stability as demonstrated in the cadaver testing. The outer cover capsulizing the distal laser manipulator also provides protection from contact with local anatomy.

With the flexible endoscopic approach, the proposed robotic system enabled direct exposure of the ablation laser to the target lesion, without the need for positioning patients with extreme neck extension to accommodate rigid laser micromanipulators. Moreover, by introducing MRI guidance, clinicians could have intuitive and precise feedback of the ablation process, which is critical for function preservation of surrounding structures. Temperature diffusion monitoring proved to be useful in laser surgery, as thermal dose and accumulated tissue damage could be calculated accordingly. MR parameters sensitive to temperature, such as PRF shift, were adjusted so that the MR susceptibility at high SNR can be used to measure small changes of temperature with <1°C accuracy. Combined with advanced MR sequences, even three-dimensional resection margins can be evaluated intra-operatively to determine the completeness of laser excision/resection. This could protect critical healthy structures, while eliminating the need for FSA, thus smoothening the surgical workflow and shortening operation time.

Chapter 7

Conclusion

7.1 Summary of Thesis Achievements

This thesis aims to address several technical challenges in MRI-guided non-contact laser surgery, which is demanding for a compact and dexterous manipulator, as well as its high-precision control. A fluid-driven soft manipulator is introduced to enable this MRI-guided minimally invasive intervention by leveraging its compactness and MR safety. More importantly, precise motion control of soft robots is studied to lay the foundation of its application in surgery. A soft robotic system enabling MRI-guided transoral laser surgery is designed based on the relevant clinical requirements. Preclinical trials were conducted in MRI so as to validate the overall performances of the integrated robot system. The materials presented in this thesis bring a new dimension for the development of MRI-guided robotics by providing improved accessibility, dexterity, and safety. The major achievements of the thesis are summarized as follows:

1. *Online-learning-based eye-in-hand visual servoing:* A nonparametric online learning control framework facilitating precise eye-in-hand visual servoing of a pneumatic-driven soft robot made from hyper-elastic materials. The controller can be initialized by learning the inverse transition mapping solely from collected sensing data, thus alleviating the need for analytically modeling the robot kinematics or calibrating the camera intrinsic/extrinsic. Localized GPR models enable fast online updates to accommodate new input data that reflect the latest robot status. The controller achieved precise path following for a fluid-driven soft robot with varying payload applied on its tip (up to 105% of the robot weight). Stable tracking was

maintained even when the load was quickly increased ($\sim$ 62.5% of robot weight over 1.5 s), while a slowly-varied payload can be effectively compensated. After removing the loading, path following accuracy further increased with an RMSE of 5.6 pixels and a maximum of 10.8 pixels. Combining strain sensing of FBGs and camera images enables enhanced motion estimation in feature-deficient scenes. With AR integrated, a virtual view can be created to extend the camera view beyond the normal FOV, thus offering enhanced guidance for operators.

2. *High-precision control for soft robotic laser steering*: An eye-to-hand visual servo controller enabling precise soft robotic laser steering. Epipolar geometry is *first* constructed to relate the camera frame with soft robot kinematics, thus modeling the inverse transition mapping. Based on this inverse mapping, a feedback controller is established, without having to incorporate prior knowledge of the tissue surface geometry. A state observer is also designed to provide a dynamic estimation of robot configurations, enabling accurate computation of the inverse transition mapping. The closed-loop controller can provide robustness to actuation disturbances compared to open-loop control. It brings new opportunities to precise laser projection using a soft manipulator, which could be equipped on an endoscope tip for endoscopic laser surgery. Moreover, the MRI-compatibility of soft laser manipulator enables further application in MRI-guided laser treatment, which could enhance the intra-op monitoring of laser ablation outcomes.

3. *Soft laser manipulator design and optimization based on FEA:* A miniaturized laser manipulator is designed to provide delicate laser beam steering (+/- 26°) in a confined region. The manipulator can be actuated through micro-volume (<0.004 mL) fluidic inflation of elastomer chambers, which are individually integrated with spring reinforcement constraints. FEA is conducted to simulate the robot characteristics upon pressurization, thus facilitating the design optimization. Based on FEA, an optimized spring reinforcement is designed, providing enhanced stiffness, improved responsiveness, and reduced hysteresis for the soft manipulator, thus allowing accurate laser beam control.

4. *Validation of wireless MR marker for MR-based positional localization*: A novel MRI-based omnidirectional positional marker was fabricated and evaluated. The markers were made of cylindrical glass tubes (Ø3 × 8 mm) with Gadolinium-doped water (concentration: 10 mM) sealed inside to provide the signal source. Miniaturized

RF-coil (1.5×6.7 mm) was attached on the tubes to locally amplify MR signal, thus offering high image contrast against the background even under low flip-angle ($<1°$) excitation. The wireless MR-markers enable convenient integration with instruments, allowing continuous real-time positional tracking in MRI coordinates. 1D-projection pulse-sequence can be applied so as to enable fast acquisition of positional signal (>30 Hz).

5. *Soft robotic system for MRI-guided transoral laser surgery*: The *first* soft robotic platform enabling intra-op MRI-guided transoral laser microsurery. The robot comprises two actively-steerable soft continuum segments, i.e., active bending section and distal laser manipulator. The active bending section offers a 60º omni-directional bending workspace for coarse navigation to the region of interest. The distal laser manipulator allows fine laser beam pan-tilting (±26º) in a local target area, with a maximum scanning area of 15×15 mm^2 at a 15 mm projection distance. Integrated with an fiber-based medical laser, the robot enables endoscopic laser delivery and operating in the oropharyngeal region with sub-millimeter accuracy ($<\sim$0.2 mm). A patient-specific dental guard is designed to create an open-jaw cavity for robot anchorage and access of auxiliary instruments such as a fiberscope. The entire robotic system passed MR compatibility testing with no EM artifacts, which allows full robot operation during imaging even at the MRI scanner isocenter.

6. *Preclinical validation of the transoral soft robotic system*: Preclinical trials were conducted to evaluate the robot performance both with *ex-vivo* swine tissue and a cadaver model. During *ex-vivo* tissue testing, the use of anatomical T2 imaging and MR thermometry on monitoring laser ablation was investigated, which showed the potential for measuring laser ablation depth and temperature diffusion region. A cadaver trial was performed to validate the robot setup/anchorage, intra-op MRI guidance, as well as visual servo-based laser beam manipulation in actual human anatomy. The robot achieved a precise and automated path following of the laser spot along a trajectory defined *in-situ* by the surgeon. During the ablation, MR thermometry was performed to measure the thermal diffusion process in the ablated lesion. The robot provided a stable platform during laser ablation due to the rigid anchorage of its base with the dental guard.

7.2 Future Work

In the future work, the vision-based controllers (**Chapter 3** and **Chapter 4**) will be extended by combining an extra sensing modality, which is optical FBG. Optical fiber allows convenient integration with soft continuum robots. FBGs on the optical fiber would offer rich configuration-related strain data, which can be trained to achieve pose estimation of the flexible manipulator. More excitingly, the FBG sensor is intrinsically MR safe, which provides a promising direction to track flexible instruments in MRI. Through the synergetic use of FBG data and camera image feedback, we could develop a soft robot motion controller with improved robustness to disturbances in the camera view, such as varying lighting and occlusion.

For the transoral soft robotic system (**Chapter 5** and **Chapter 6**), a plan is to implement real-time MR imaging sequences customized for the laser ablation monitoring, optimizing for higher temporal resolution while maintaining sufficient imaging (spatial) resolution. Custom real-time sequences interleaving multi-slice anatomical and thermometry images bring the potential for greatly enhanced ablation monitoring and could enable automated laser ablation control based on MRI feedback. The *in-situ* MR images will also be overlaid on the fiberscope camera view, providing an augmented reality environment to further assist surgeons in evaluating the ablation process. A promising extension of this work is to couple optical coherence tomography (OCT) with the ablation setting by adopting the lens-pigtailed optical fiber as an OCT sampling arm, which can provide tissue depth images in micron-scale resolution. Alongside the proposed laser steering mechanism, scanning paths/patterns could be performed to form 3D forward-viewing OCT images underlying the lesion. Visual servo of the laser spot projected on tissue would enable the operator to prescribe any preferred OCT scanning profile *in-situ* in the endoscopic view, thus facilitating tissue/tumor characterization at the cellular scale, apart from using MRI.

References

[1] P. Tamije, V. Palanisamy, and T. Purusothaman, "Performance analysis of clustering algorithms in brain tumor detection of mr images," *European journal of scientific research, ISSN*, pp. 321–330.

[2] K. Y. Park, J. J. Lee, D. Dierker, L. M. Marple, C. D. Hacker, J. L. Roland, D. S. Marcus, M. Milchenko, M. M. Miller-Thomas, T. L. Benzinger, *et al.*, "Mapping language function with task-based vs. resting-state functional mri," *PloS one*, vol. 15, no. 7, p. e0236423, 2020.

[3] C. H. Cunningham, J. Y. Lau, A. P. Chen, B. J. Geraghty, W. J. Perks, I. Roifman, G. A. Wright, and K. A. Connelly, "Hyperpolarized 13c metabolic mri of the human heart: initial experience," *Circulation research*, vol. 119, no. 11, pp. 1177–1182, 2016.

[4] "Hitachi inc." https://www.hitachi-medical-systems.eu/products/mri/oasis-12t.html. Accessed: 2021-06-06.

[5] "Hyperfine inc." https://hyperfine.io/. Accessed:2021-06-06.

[6] "General electric inc." https://www.gehealthcare.com. Accessed:2021-06-06.

[7] V. Seifert and C. Senft, "Utilization of low-field intraoperative mri in glioma surgery–an overview," *Intraoperative MRI-Guided Neurosurgery. Stuttgart: Thieme*, pp. 99–107, 2011.

[8] "Hitachi inc." https://www.hitachi-medical.com.sg/solutions/intraoperative-mri. Accessed: 2021-06-06.

[9] "Imris inc." https://www.imris.com/hybrid-operating-suite/. Accessed: 2021-06-06.

[10] "Brigham and women's hospital, harvard medical school." https://ncigt.org/amigo. Accessed: 2021-06-06.

[11] "Clearepoint neuro inc." https://www.clearpointneuro.com. Accessed: 2021-06-08.

[12] "Monteris medical inc." https://www.monteris.com. Accessed: 2021-06-08.

[13] C. J. Nycz, R. Gondokaryono, P. Carvalho, N. Patel, M. Wartenberg, J. G. Pilitsis, and G. S. Fischer, "Mechanical validation of an mri compatible stereotactic neurosurgery robot in preparation for pre-clinical trials," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1677–1684, IEEE, 2017.

[14] C. Jun, S. Lim, J.-P. Wolinsky, T. Garzon-Muvdi, D. Petrisor, K. Cleary, and D. Stoianovici, "Mr safe robot assisted needle access of the brain: preclinical study," *Journal of Medical Robotics Research*, vol. 3, no. 01, p. 1850003, 2018.

[15] Z. Guo, Z. Dong, K.-H. Lee, C. L. Cheung, H.-C. Fu, J. D. Ho, H. He, W.-S. Poon, D. T.-M. Chan, and K.-W. Kwok, "Compact design of a hydraulic driving robot for intraoperative mri-guided bilateral stereotactic neurosurgery," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2515–2522, 2018.

[16] A. Melzer, B. Gutmann, T. Remmele, R. Wolf, A. Lukoscheck, M. Bock, H. Bardenheuer, and H. Fischer, "Innomotion for percutaneous image-guided interventions," *IEEE Engineering in Medicine and Biology Magazine*, vol. 27, no. 3, pp. 66–73, 2008.

[17] E. Franco, M. Ristic, M. Rea, and W. M. Gedroyc, "Robot-assistant for mri-guided liver ablation: A pilot study," *Medical physics*, vol. 43, no. 10, pp. 5347–5356, 2016.

[18] G. Li, N. A. Patel, J. Hagemeister, J. Yan, D. Wu, K. Sharma, K. Cleary, and I. Iordachita, "Body-mounted robotic assistant for mri-guided low back pain injection," *International journal of computer assisted radiology and surgery*, vol. 15, no. 2, pp. 321–331, 2020.

[19] Z. He, Z. Dong, G. Fang, J. D.-L. Ho, C.-L. Cheung, H.-C. Chang, C. C.-N. Chong, J. Y.-K. Chan, D. T. M. Chan, and K.-W. Kwok, "Design of a percutaneous mri-guided needle robot with soft fluid-driven actuator," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2100–2107, 2020.

[20] M. Waqar, M. Nawaz Abro, Q. Soomro, M. Shahban, and S. Khatoon, "Retrospective incidence analysis of head and neck cancer patients in rural areas of sindh, pakistan," *Jundishapur Journal of Chronic Disease Care*, vol. 8, no. 4, 2019.

[21] R. Misirovs, I. Gartner, and J. Manickavasagam, "Double pyramid technique of transoral laser partial laryngectomy for radiorecurrent laryngeal cancer," *BMJ Case Reports CP*, vol. 11, no. 1, 2018.

[22] S. S. Lo and J. J. Lu, "Natural history, presenting symptoms, and diagnosis of nasopharyngeal carcinoma," in *Nasopharyngeal Cancer*, pp. 41–51, Springer, 2010.

[23] M. Bannister and K. W. Ah-See, "Is oropharyngeal cancer being misdiagnosed as acute tonsillitis?," *British Journal of General Practice*, vol. 64, no. 628, pp. e742–e744, 2014.

[24] K. Okami, "Clinical features and treatment strategy for hpv-related oropharyngeal cancer," *International journal of clinical oncology*, vol. 21, no. 5, pp. 827–835, 2016.

[25] N. Andreff, S. Dembélé, B. Tamadazte, and Z. E. Hussnain, "Epipolar geometry for vision-guided laser surgery.," in *10th International Conference on Informatics in Control, Automation and Robotics, ICINCO'2013.*, pp. 1–6, 2013.

[26] "A.r.c laser inc." https://www.arclaser.de/en/products/co2-laser/c-las-articulated-arm/. Accessed: 2021-06-11.

[27] "Luminens inc." https://lumenis.com/. Accessed: 2021-06-11.

[28] "Kls martin inc." https://www.klsmartin.com. Accessed: 2021-06-11.

[29] "Deka inc." https://www.dekalaser.com/en-GB/. Accessed: 2021-06-11.

[30] L. S. Mattos, G. Dagnino, G. Becattini, M. Dellepiane, and D. G. Caldwell, "A virtual scalpel system for computer-assisted laser microsurgery," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1359–1365, IEEE, 2011.

[31] N. Deshpande, G. Peretti, F. Mora, L. Guastini, J. Lee, G. Barresi, D. G. Caldwell, and L. S. Mattos, "Design and study of a next-generation computer-assisted system for transoral laser microsurgery," *OTO open*, vol. 2, no. 2, p. 2473974X18773327, 2018.

[32] M. M. Chen, R. K. Orosco, G. C. Lim, and F. C. Holsinger, "Improved transoral dissection of the tongue base with a next-generation robotic surgical system," *The Laryngoscope*, vol. 128, no. 1, pp. 78–83, 2018.

[33] "Intuitive surgical inc." https://www.intuitive.com. Accessed: 2021-06-11.

[34] "Medtronic inc." https://medrobotics.com/gateway/flex-system-int/. Accessed: 2021-06-11.

[35] R. Renevier, B. Tamadazte, K. Rabenorosoa, L. Tavernier, and N. Andreff, "Endoscopic laser surgery: Design, modeling, and control," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 99–106, 2016.

[36] A. Acemoglu, N. Deshpande, and L. S. Mattos, "Towards a magnetically-actuated laser scanner for endoscopic microsurgeries," *Journal of Medical Robotics Research*, vol. 3, no. 02, p. 1840004, 2018.

[37] S. Patel, M. Rajadhyaksha, S. Kirov, Y. Li, and R. Toledo-Crow, "Endoscopic laser scalpel for head and neck cancer surgery," in *Photonic Therapeutics and Diagnostics VIII*, vol. 8207, p. 82071S, International Society for Optics and Photonics, 2012.

[38] M. Zhao, T. J. O. Vrielink, A. A. Kogkas, M. S. Runciman, D. S. Elson, and G. P. Mylonas, "Laryngotors: A novel cable-driven parallel robotic system for transoral laser phonosurgery," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1516–1523, 2020.

[39] A. Marzullo, G. Serio, L. Madami, F. Pezzuto, F. Fortarezza, N. Quaranta, M. L. Fiorella, T. Lettini, L. Resta, M. Marrelli, *et al.*, "Intraoperative frozen section as a reliable ancillary technique in salivary gland surgery: A cross sectional study," *F1000Research*, vol. 7, no. 231, p. 231, 2018.

[40] J.-O. Park, S.-L. Jung, Y.-H. Joo, C.-K. Jung, K.-J. Cho, and M.-S. Kim, "Diagnostic accuracy of magnetic resonance imaging (mri) in the assessment of tumor invasion depth in oral/oropharyngeal cancer," *Oral oncology*, vol. 47, no. 5, pp. 381–386, 2011.

[41] M. Huisman, R. M. Staruch, M. Ladouceur-Wodzak, M. A. Van Den Bosch, D. K. Burns, A. Chhabra, and R. Chopra, "Non-invasive targeted peripheral nerve ablation using 3d mr neurography and mri-guided high-intensity focused ultrasound (mr-hifu): pilot study in a swine model," *PLoS One*, vol. 10, no. 12, p. e0144742, 2015.

[42] D. B. Comber, E. B. Pitt, H. B. Gilbert, M. W. Powelson, E. Matijevich, J. S. Neimat, R. J. Webster III, and E. J. Barth, "Optimization of curvilinear needle trajectories for transforamenal hippocampotomy," *Operative Neurosurgery*, vol. 13, no. 1, pp. 15–22, 2017.

[43] T. Kato, I. Okumura, H. Kose, K. Takagi, and N. Hata, "Tendon-driven continuum robot for neuroendoscopy: validation of extended kinematic mapping for hysteresis operation," *International journal of computer assisted radiology and surgery*, vol. 11, no. 4, pp. 589–602, 2016.

[44] "Endotics inc." http://www.endotics.com/. Accessed: 2021-06-14.

[45] M. Cianchetti, T. Ranzani, G. Gerboni, I. De Falco, C. Laschi, and A. Menciassi, "Stiff-flop surgical manipulator: Mechanical design and experimental characterization of the single module," in *2013 IEEE/RSJ international conference on intelligent robots and systems*, pp. 3576–3581, IEEE, 2013.

[46] R. K. Katzschmann, C. Della Santina, Y. Toshimitsu, A. Bicchi, and D. Rus, "Dynamic motion control of multi-segment soft robots using piecewise constant curvature matched with an augmented rigid body model," in *2019 2nd IEEE International Conference on Soft Robotics (RoboSoft)*, pp. 454–461, IEEE, 2019.

[47] R. J. Webster III and B. A. Jones, "Design and kinematic modeling of constant curvature continuum robots: A review," *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1661–1683, 2010.

[48] D. Trivedi and C. D. Rahn, "Shape sensing for soft robotic manipulators," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 49040, pp. 335–343, 2009.

[49] Z. Zhang, J. Dequidt, A. Kruszewski, F. Largilliere, and C. Duriez, "Kinematic modeling and observer based control of soft robot using real-time finite element method," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5509–5514, IEEE, 2016.

[50] K.-H. Lee, M. C. Leong, M. C. Chow, H.-C. Fu, W. Luk, K.-Y. Sze, C.-K. Yeung, and K.-W. Kwok, "Fem-based soft robotic control framework for intracavitary navigation," in *2017 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pp. 11–16, IEEE, 2017.

[51] B. Banko, V. Djukic, J. Milovanovic, G. Lilic, M. Kratovac-Dunjic, R. Milenkovic, R. Maksimovic, *et al.*, "Preoperative staging of laryngeal carcinoma: comparison of ct and mri with histopathology," European Congress of Radiology 2013, 2013.

[52] D.-M. Denk and R. Schöfl, "Endoscopy of the pharynx and esophagus," in *Radiology of the Pharynx and the Esophagus*, pp. 167–181, Springer, 2004.

[53] J. P. Whitney, M. F. Glisson, E. L. Brockmeyer, and J. K. Hodgins, "A low-friction passive fluid transmission and fluid-tendon soft actuator," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2801–2808, IEEE, 2014.

[54] W. E. L. Grimson, R. Kikinis, F. A. Jolesz, and P. M. Black, "Image-guided surgery," *Scientific American*, vol. 280, no. 6, pp. 62–69, 1999.

[55] P. D. Afonso and V. Mascarenhas, "Imaging techniques for the diagnosis of soft tissue tumors," *Reports in Medical Imaging*, vol. 8, pp. 63–70, 2015.

[56] B.-K. Han, M. D. Schnall, S. G. Orel, and M. Rosen, "Outcome of mri-guided breast biopsy," *American Journal of Roentgenology*, vol. 191, no. 6, pp. 1798–1804, 2008.

[57] C. G. Overduin, J. J. Fütterer, and J. O. Barentsz, "Mri-guided biopsy for prostate cancer detection: a systematic review of current clinical results," *Current urology reports*, vol. 14, no. 3, pp. 209–213, 2013.

[58] A. Tzifa, G. A. Krombach, N. Krämer, S. Krüger, A. Schütte, M. von Walter, T. Schaeffter, S. Qureshi, T. Krasemann, E. Rosenthal, *et al.*, "Magnetic resonance–guided cardiac interventions using magnetic resonance–compatible devices: a preclinical study and first-in-man congenital interventions," *Circulation: Cardiovascular Interventions*, vol. 3, no. 6, pp. 585–592, 2010.

[59] K.-H. Lee, K. C. D. Fu, Z. Guo, Z. Dong, M. C. Leong, C.-L. Cheung, A. P.-W. Lee, W. Luk, and K.-W. Kwok, "Mr safe robotic manipulator for mri-guided intracardiac catheterization," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 2, pp. 586–595, 2018.

[60] A. Arepally, "Targeted drug delivery under mri guidance," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 27, no. 2, pp. 292–298, 2008.

[61] M. S. Breen, M. Breen, K. Butts, L. Chen, G. M. Saidel, and D. L. Wilson, "Mri-guided thermal ablation therapy: Model and parameter estimates to predict cell death from mr thermometry images," *Annals of biomedical engineering*, vol. 35, no. 8, pp. 1391–1403, 2007.

[62] J. D. Hazle, C. J. Diederich, M. Kangasniemi, R. E. Price, L. E. Olsson, and R. J. Stafford, "Mri-guided thermal therapy of transplanted tumors in the canine prostate using a directional transurethral ultrasound applicator," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 15, no. 4, pp. 409–417, 2002.

[63] G. I. Ogbole, A. O. Adeyomoye, A. Badu-Peprah, Y. Mensah, and D. A. Nzeh, "Survey of magnetic resonance imaging availability in west africa," *Pan African Medical Journal*, vol. 30, no. 1, 2018.

[64] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince, *MRI from Picture to Proton*. Cambridge university press, 2017.

[65] O. Gröhn, S. Deleye, D. Bertoglio, and S. Dedeurwaerdere, "In vivo imaging in rodents," in *Models of Seizures and Epilepsy*, pp. 197–215, Elsevier, 2017.

[66] M. Uecker, S. Zhang, D. Voit, A. Karaus, K.-D. Merboldt, and J. Frahm, "Real-time mri at a resolution of 20 ms," *NMR in Biomedicine*, vol. 23, no. 8, pp. 986–994, 2010.

[67] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 2018.

[68] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, *et al.*, "Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1310–1321, 2017.

[69] I. Roifman, J. Gutierrez, E. Wang, L. Biswas, J. Sparkes, K. A. Connelly, and G. A. Wright, "Evaluating a novel free-breathing accelerated cardiac mri cine sequence in patients with cardiomyopathy," *Magnetic resonance imaging*, vol. 61, pp. 260–266, 2019.

[70] S. Okamoto, Y. Matsui, T. Hiraki, T. Iguchi, T. Komaki, T. Yamauchi, M. Uka, K. Tomita, J. Sakurai, H. Gobara, *et al.*, "Needle artifact characteristics and insertion accuracy using a 1.2 t open mri scanner: A phantom study," *Diagnostic and Interventional Imaging*, 2021.

[71] K. Stephens, "Fda clears deep learning application for hyperfine's swoop portable mri system," *AXIS Imaging News*, 2021.

[72] J. R. Corea, A. M. Flynn, B. Lechêne, G. Scott, G. D. Reed, P. J. Shin, M. Lustig, and A. C. Arias, "Screen-printed flexible mri receive coils," *Nature communications*, vol. 7, no. 1, pp. 1–7, 2016.

[73] B. Collick, B. Behzadnezhad, S. A. Hurley, N. Mathew, N. Behdad, S. Lindsay, F. Robb, R. Stormont, and A. McMillan, "Rapid development of application-specific flexible mri receive coils," *Physics in Medicine & Biology*, vol. 65, no. 19, p. 19NT01, 2020.

[74] M. Moche, D. Zajonz, T. Kahn, and H. Busse, "Mri-guided procedures in various regions of the body using a robotic assistance system in a closed-bore scanner: preliminary clinical experience and limitations," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 31, no. 4, pp. 964–974, 2010.

[75] ASTM, "Standard practice for marking medical devices and other items for safety in the magnetic resonance environment," *ASTM F2503-13*, 2013.

[76] U. Salem, V. A. Kumar, J. E. Madewell, D. F. Schomer, D. C. de Almeida Bastos, P. O. Zinn, J. S. Weinberg, G. Rao, S. S. Prabhu, and R. R. Colen, "Neurosurgical applications of mri guided laser interstitial thermal therapy (litt)," *Cancer Imaging*, vol. 19, no. 1, pp. 1–13, 2019.

[77] G. Allegretti, P. Saccomandi, F. Giurazza, M. Caponero, G. Frauenfelder, F. Di Matteo, B. B. Zobel, S. Silvestri, and E. Schena, "Magnetic resonance-based thermometry during laser ablation on ex-vivo swine pancreas and liver," *Medical engineering & physics*, vol. 37, no. 7, pp. 631–641, 2015.

[78] G. Müller-Lisse, A. Heuck, M. Stehling, M. Frimberger, M. Thoma, P. Schneede, R. Muschter, A. Hofstetter, and M. Reiser, "Mri monitoring before, during and after interstitial laser-induced hyperthermia of benign prostatic hyperplasia. initial clinical experiences," *Der Radiologe*, vol. 36, no. 9, pp. 722–731, 1996.

[79] F. Orsi, L. Zhang, P. Arnone, G. Orgera, G. Bonomo, P. D. Vigna, L. Monfardini, K. Zhou, W. Chen, Z. Wang, *et al.*, "High-intensity focused ultrasound ablation: effective and safe therapy for solid tumors in difficult locations," *American journal of roentgenology*, vol. 195, no. 3, pp. W245–W252, 2010.

[80] D. M. Quinlan, M. J. Naslund, and C. B. Brendler, "Application of argon beam coagulation in urological surgery," *The Journal of urology*, vol. 147, no. 2, pp. 410–412, 1992.

[81] A. N. Mirza, B. D. Fornage, N. Sneige, H. M. Kuerer, L. A. Newman, F. C. Ames, and S. E. Singletary, "Radiofrequency ablation of solid tumors.," *Cancer journal (Sudbury, Mass.)*, vol. 7, no. 2, pp. 95–102, 2001.

[82] "Cls inc." https://clinicallaser.com/tranberg/. Accessed: 2021-06-08.

[83] "Medtronic inc." https://global.medtronic.com. Accessed: 2021-06-08.

[84] J. S. Perlmutter and J. W. Mink, "Deep brain stimulation," *Annu. Rev. Neurosci.*, vol. 29, pp. 229–257, 2006.

[85] P. A. Starr, L. C. Markun, P. S. Larson, M. M. Volz, A. J. Martin, and J. L. Ostrem, "Interventional mri–guided deep brain stimulation in pediatric dystonia: first experience with the clearpoint system," *Journal of Neurosurgery: Pediatrics*, vol. 14, no. 4, pp. 400–408, 2014.

[86] A. M. Mohammadi and J. L. Schroeder, "Laser interstitial thermal therapy in treatment of brain tumors–the neuroblate system," *Expert review of medical devices*, vol. 11, no. 2, pp. 109–119, 2014.

[87] M. Ahmed, C. L. Brace, F. T. Lee Jr, and S. N. Goldberg, "Principles of and advances in percutaneous ablation," *Radiology*, vol. 258, no. 2, pp. 351–369, 2011.

[88] M. M. Fleming, A. I. Holbrook, and M. S. Newell, "Update on image-guided percutaneous ablation of breast cancer," *American Journal of Roentgenology*, vol. 208, no. 2, pp. 267–274, 2017.

[89] R. N. Uppot, S. G. Silverman, R. J. Zagoria, K. Tuncali, D. D. Childs, and D. A. Gervais, "Imaging-guided percutaneous ablation of renal cell carcinoma: a primer of how we do it," *American Journal of Roentgenology*, vol. 192, no. 6, pp. 1558–1570, 2009.

[90] G. Foltz, "Image-guided percutaneous ablation of hepatic malignancies," in *Seminars in interventional radiology*, vol. 31, pp. 180–186, Thieme Medical Publishers, 2014.

[91] D. Stoianovici, C. Jun, S. Lim, P. Li, D. Petrisor, S. Fricke, K. Sharma, and K. Cleary, "Multi-imager compatible, mr safe, remote center of motion needle-guide robot," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 1, pp. 165–177, 2017.

[92] D. J. Coluzzi, "Fundamentals of dental lasers: science and instruments," *Dental Clinics*, vol. 48, no. 4, pp. 751–770, 2004.

[93] L. Q. Chow, "Head and neck cancer," *New England Journal of Medicine*, vol. 382, no. 1, pp. 60–72, 2020.

[94] J. Attra and N. Kokot, "Transoral robotic surgery and lasers," in *Biomedical Optics in Otorhinolaryngology*, pp. 183–201, Springer, 2016.

[95] G. Sankaranarayanan, R. R. Resapu, D. B. Jones, S. Schwaitzberg, and S. De, "Common uses and cited complications of energy in surgery," *Surgical endoscopy*, vol. 27, no. 9, pp. 3056–3072, 2013.

[96] D. Ginat, S. Sammet, and G. Christoforidis, "Mr thermography–guided head and neck lesion laser ablation," *American Journal of Neuroradiology*, vol. 39, no. 9, pp. 1593–1596, 2018.

[97] M. C. Jäckel, A. Martin, and W. Steiner, "Twenty-five years experience with laser surgery for head and neck tumors," *European archives of oto-rhino-laryngology*, vol. 264, no. 6, pp. 577–585, 2007.

[98] S. M. Zeitels and J. A. Burns, "Oncologic efficacy of angiolytic ktp laser treatment of early glottic cancer," *Annals of Otology, Rhinology & Laryngology*, vol. 123, no. 12, pp. 840–846, 2014.

[99] N. P. Parker, M. S. Weidenbecher, A. D. Friedman, B. A. Walker, and D. G. Lott, "Ktp laser treatment of early glottic cancer: A multi-institutional retrospective study," *Annals of Otology, Rhinology & Laryngology*, vol. 130, no. 1, pp. 47–55, 2021.

[100] M. Karaman, T. Gün, B. Temelkuran, E. Aynacı, C. Kaya, and A. M. Tekin, "Comparison of fiber delivered co 2 laser and electrocautery in transoral robot assisted tongue base surgery," *European Archives of Oto-Rhino-Laryngology*, vol. 274, no. 5, pp. 2273–2279, 2017.

[101] F. C. Holsinger, A. D. Sweeney, K. Jantharapattana, A. Salem, R. S. Weber, W. Y. Chung, C. M. Lewis, and D. G. Grant, "The emergence of endoscopic head and neck surgery," *Current oncology reports*, vol. 12, no. 3, pp. 216–222, 2010.

[102] M. K. Bhayani, F. C. Holsinger, and S. Y. Lai, "A shifting paradigm for patients with head and neck cancer: transoral robotic surgery (tors)," *Oncology*, vol. 24, no. 11, p. 1010, 2010.

[103] D. G. Grant, M. L. Hinni, J. R. Salassa, W. C. Perry, R. E. Hayden, and J. D. Casler, "Oropharyngeal cancer: a case for single modality treatment with transoral laser microsurgery," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 135, no. 12, pp. 1225–1230, 2009.

[104] W. Steiner, O. Fierek, P. Ambrosch, C. P. Hommerich, and M. Kron, "Transoral laser microsurgery for squamous cell carcinoma of the base of the tongue," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 129, no. 1, pp. 36–43, 2003.

[105] M. Rubinstein and W. B. Armstrong, "Transoral laser microsurgery for laryngeal cancer: a primer and review of laser dosimetry," *Lasers in medical science*, vol. 26, no. 1, pp. 113–124, 2011.

[106] G. J. Petruzzelli, "Transoral laser microsurgery: applications in head and neck oncology," *Expert review of medical devices*, vol. 6, no. 6, pp. 599–602, 2009.

[107] D. Kundrat, R. Graesslin, A. Schoob, D. Friedrich, M. Scheithauer, T. Hoffmann, T. Ortmaier, L. A. Kahrs, and P. Schuler, "Preclinical performance evaluation of a robotic endoscope for non-contact laser surgery," *Annals of Biomedical Engineering*, vol. 49, no. 2, pp. 585–600, 2021.

[108] M. Remacle, "Laser-assisted microphonosurgery," in *Surgery of larynx and trachea*, pp. 51–56, Springer, 2009.

[109] M. Remacle, G. Lawson, M.-C. Nollevaux, and M. Delos, "Current state of scanning micromanipulator applications with the carbon dioxide laser," *Annals of Otology, Rhinology & Laryngology*, vol. 117, no. 4, pp. 239–244, 2008.

[110] J. A. Burns, J. B. Kobler, J. T. Heaton, G. Lopez-Guerra, R. R. Anderson, and S. M. Zeitels, "Thermal damage during thulium laser dissection of laryngeal soft tissue is reduced with air cooling: ex vivo calf model study," *Annals of Otology, Rhinology & Laryngology*, vol. 116, no. 11, pp. 853–857, 2007.

[111] S. M. Zeitels, J. A. Burns, L. M. Akst, R. E. Hillman, M. S. Broadhurst, and R. R. Anderson, "Office-based and microlaryngeal applications of a fiber-based thulium laser," *Annals of Otology, Rhinology & Laryngology*, vol. 115, no. 12, pp. 891–896, 2006.

[112] S. Ayari-Khalfallah, C. Fuchsmann, and P. Froehlich, "Thulium laser in airway diseases in children," *Current opinion in otolaryngology & head and neck surgery*, vol. 16, no. 1, pp. 55–59, 2008.

[113] E. Passacantilli, M. Antonelli, A. D'Amico, C. P. Delfinis, G. Anichini, J. Lenzi, and A. Santoro, "Neurosurgical applications of the 2-μm thulium laser: histological evaluation of meningiomas in comparison to bipolar forceps and an ultrasonic aspirator," *Photomedicine and laser surgery*, vol. 30, no. 5, pp. 286–292, 2012.

[114] K. M. Van Abel, E. J. Moore, M. L. Carlson, J. A. Davidson, J. J. Garcia, S. M. Olsen, and K. D. Olsen, "Transoral robotic surgery using the thulium: Yag laser: a prospective study," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 138, no. 2, pp. 158–166, 2012.

[115] J. D. Richmon, "Transoral palate-sparing nasopharyngectomy with the f lex® s ystem: Preclinical study," *The Laryngoscope*, vol. 125, no. 2, pp. 318–322, 2015.

[116] M. Benazzo, P. Canzi, and A. Occhini, "Transoral robotic surgery with laser for head and neck cancers: a feasibility study," *Orl*, vol. 74, no. 3, pp. 124–128, 2012.

[117] D. Kundrat, A. Schoob, T. Piskon, R. Grässlin, P. J. Schuler, T. K. Hoffmann, L. A. Kahrs, and T. Ortmaier, "Toward assistive technologies for focus adjustment in teleoperated robotic non-contact laser surgery," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 3, pp. 145–157, 2019.

[118] S. M. Olson, M. Hussaini, and J. S. Lewis, "Frozen section analysis of margins for head and neck tumor resections: reduction of sampling errors with a third histologic level," *Modern Pathology*, vol. 24, no. 5, pp. 665–670, 2011.

[119] D. Pardo, L. Fichera, D. G. Caldwell, and L. S. Mattos, "Thermal supervision during robotic laser microsurgery," in *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, pp. 363–368, IEEE, 2014.

[120] A. Acemoglu, L. Fichera, I. E. Kepiro, D. G. Caldwell, and L. S. Mattos, "Laser incision depth control in robot-assisted soft tissue microsurgery," *Journal of Medical Robotics Research*, vol. 2, no. 03, p. 1740006, 2017.

[121] H. Alsaffar, D. Goldstein, E. King, J. de Almeida, D. Brown, R. Gilbert, P. J. Gullane, O. Espin-Garcia, W. Xu, and J. Irish, "Correlation between clinical and mri

assessment of depth of invasion in oral tongue squamous cell carcinoma," *Journal of Otolaryngology-Head & Neck Surgery*, vol. 45, no. 1, pp. 1–5, 2016.

[122] B. D. de Senneville, C. Mougenot, B. Quesson, I. Dragonu, N. Grenier, and C. T. Moonen, "Mr thermometry for monitoring tumor ablation," *European radiology*, vol. 17, no. 9, pp. 2401–2410, 2007.

[123] M. Marx, P. Ghanouni, and K. Butts Pauly, "Specialized volumetric thermometry for improved guidance of mr g fus in brain," *Magnetic resonance in medicine*, vol. 78, no. 2, pp. 508–517, 2017.

[124] M. Lepetit-Coiffé, H. Laumonier, O. Seror, B. Quesson, M.-B. Sesay, C. T. Moonen, N. Grenier, and H. Trillaud, "Real-time monitoring of radiofrequency ablation of liver tumors using thermal-dose calculation by mr temperature imaging: initial results in nine patients, including follow-up," *European radiology*, vol. 20, no. 1, pp. 193–201, 2010.

[125] D. Silva, M. Sharma, R. Juthani, A. Meola, and G. H. Barnett, "Magnetic resonance thermometry and laser interstitial thermal therapy for brain tumors," *Neurosurgery Clinics*, vol. 28, no. 4, pp. 525–533, 2017.

[126] R. J. McNichols, A. Gowda, M. Kangasniemi, J. A. Bankson, R. E. Price, and J. D. Hazle, "Mr thermometry-based feedback control of laser interstitial thermal therapy at 980 nm," *Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery*, vol. 34, no. 1, pp. 48–55, 2004.

[127] R. J. McNichols, A. Gowda, and S. M. Wright, "Closed-loop feedback control of laser therapy using magnetic resonance imaging," in *Thermal Treatment of Tissue: Energy Delivery and Assessment*, vol. 4247, pp. 158–165, International Society for Optics and Photonics, 2001.

[128] J. Yuan, C.-S. Mei, L. P. Panych, N. J. McDannold, and B. Madore, "Towards fast and accurate temperature mapping with proton resonance frequency-based mr thermometry," *Quantitative imaging in medicine and surgery*, vol. 2, no. 1, p. 21, 2012.

[129] A. Krieger, S.-E. Song, N. B. Cho, I. I. Iordachita, P. Guion, G. Fichtinger, and L. L. Whitcomb, "Development and evaluation of an actuated mri-compatible robotic system for mri-guided prostate intervention," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 1, pp. 273–284, 2011.

[130] Y. Wang, G. A. Cole, H. Su, J. G. Pilitsis, and G. S. Fischer, "Mri compatibility evaluation of a piezoelectric actuator system for a neural interventional robot," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6072–6075, IEEE, 2009.

[131] D. Stoianovici, A. Patriciu, D. Petrisor, D. Mazilu, and L. Kavoussi, "A new type of motor: pneumatic step motor," *IEEE/ASME Transactions On Mechatronics*, vol. 12, no. 1, pp. 98–106, 2007.

[132] Y. Chen, K.-W. Kwok, and Z. T. H. Tse, "An mr-conditional high-torque pneumatic stepper motor for mri-guided and robot-assisted intervention," *Annals of biomedical engineering*, vol. 42, no. 9, pp. 1823–1833, 2014.

[133] Z. Guo, T. Lun, Y. Chen, H. Su, D. Chan, and K. Kwok, "Novel design of an mr-safe pneumatic stepper motor for mri-guided robotic interventions," in *Proceedings of The Hamlyn Symposium on Medical Robotics*, pp. 25–28, Imperial College London and the Royal Geographical Society London. The . . . , 2016.

[134] G. Ganesh, R. Gassert, E. Burdet, and H. Bleuler, "Dynamics and control of an mri compatible master-slave system with hydrostatic transmission," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 2, pp. 1288–1294, IEEE, 2004.

[135] Z. Dong, Z. Guo, K.-H. Lee, G. Fang, W. L. Tang, H.-C. Chang, D. T. M. Chan, and K.-W. Kwok, "High-performance continuous hydraulic motor for mr safe robotic teleoperation," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1964–1971, 2019.

[136] R. A. Omary, O. Unal, D. S. Koscielski, R. Frayne, F. R. Korosec, C. A. Mistretta, C. M. Strother, and T. M. Grist, "Real-time mr imaging-guided passive catheter tracking with use of gadolinium-filled catheters," *Journal of Vascular and Interventional Radiology*, vol. 11, no. 8, pp. 1079–1085, 2000.

[137] O. Unal, J. Li, W. Cheng, H. Yu, and C. M. Strother, "Mr-visible coatings for endovascular device visualization," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 23, no. 5, pp. 763–769, 2006.

[138] D. L. Rubin, A. V. Ratner, and S. W. Young, "Magnetic susceptibility effects and their application in the development of new ferromagnetic catheters for magnetic resonance imaging.," *Investigative radiology*, vol. 25, no. 12, pp. 1325–1332, 1990.

[139] O. Unal, F. R. Korosec, R. Frayne, C. M. Strother, and C. A. Mistretta, "A rapid 2d time-resolved variable-rate k-space sampling mr technique for passive catheter tracking during endovascular procedures," *Magnetic resonance in medicine*, vol. 40, no. 3, pp. 356–362, 1998.

[140] S. Patil, O. Bieri, P. Jhooti, and K. Scheffler, "Automatic slice positioning (asp) for passive real-time tracking of interventional devices using projection-reconstruction imaging with echo-dephasing (pride)," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 62, no. 4, pp. 935–942, 2009.

[141] W. R. Nitz, A. Oppelt, W. Renz, C. Manke, M. Lenhart, and J. Link, "On the heating of linear conductive structures as guide wires and catheters in interventional mri," *Journal of Magnetic Resonance Imaging*, vol. 13, no. 1, pp. 105–114, 2001.

[142] S. Wildermuth, C. L. Dumoulin, T. Pfammatter, S. E. Maier, E. Hofmann, and J. F. Debatin, "Mr-guided percutaneous angioplasty: assessment of tracking safety, catheter handling and functionality," *Cardiovascular and interventional radiology*, vol. 21, no. 5, pp. 404–410, 1998.

[143] A. Alipour, S. Gokyar, O. Algin, E. Atalar, and H. V. Demir, "An inductively coupled ultra-thin, flexible, and passive rf resonator for mri marking and guiding purposes: Clinical feasibility," *Magnetic resonance in medicine*, vol. 80, no. 1, pp. 361–370, 2018.

[144] C.-L. Cheung, J. D.-L. Ho, V. Vardhanabhuti, H.-C. Chang, and K.-W. Kwok, "Design and fabrication of wireless multilayer tracking marker for intraoperative mri-guided interventions," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 2, pp. 1016–1025, 2020.

[145] H. Su, G. Li, D. C. Rucker, R. J. Webster III, and G. S. Fischer, "A concentric tube continuum robot with piezoelectric actuation for mri-guided closed-loop targeting," *Annals of biomedical engineering*, vol. 44, no. 10, pp. 2863–2873, 2016.

[146] A. Orekhov, C. Abah, and N. Simaan, "Snake-like robots for minimally invasive, single-port, and intraluminal surgeries," *The Enciclopedia of Medical Robotics. World Scientific*, pp. 203–243, 2018.

[147] B. Mosadegh, P. Polygerinos, C. Keplinger, S. Wennstedt, R. F. Shepherd, U. Gupta, J. Shim, K. Bertoldi, C. J. Walsh, and G. M. Whitesides, "Pneumatic networks for soft robotics that actuate rapidly," *Advanced functional materials*, vol. 24, no. 15, pp. 2163–2170, 2014.

[148] S. Kim, C. Laschi, and B. Trimmer, "Soft robotics: a bioinspired evolution in robotics," *Trends in biotechnology*, vol. 31, no. 5, pp. 287–294, 2013.

[149] C. Laschi and M. Cianchetti, "Soft robotics: new perspectives for robot bodyware and control," *Frontiers in bioengineering and biotechnology*, vol. 2, p. 3, 2014.

[150] C. Blanes, M. Mellado, and P. Beltran, "Novel additive manufacturing pneumatic actuators and mechanisms for food handling grippers," in *Actuators*, vol. 3, pp. 205–225, Multidisciplinary Digital Publishing Institute, 2014.

[151] E. Tumino, R. Sacco, M. Bertini, M. Bertoni, G. Parisi, and A. Capria, "Endotics system vs colonoscopy for the detection of polyps," *World Journal of Gastroenterology: WJG*, vol. 16, no. 43, p. 5452, 2010.

[152] F. Cosentino, E. Tumino, G. R. Passoni, E. Morandi, and A. Capria, "Functional evaluation of the endotics system, a new disposable self-propelled robotic colonoscope: in vitro tests and clinical trial," *The International journal of artificial organs*, vol. 32, no. 8, pp. 517–527, 2009.

[153] B. Vucelic, D. Rex, R. Pulanic, J. Pfefer, I. Hrstic, B. Levin, Z. Halpern, and N. Arber, "The aer-o-scope: proof of concept of a pneumatic, skill-independent, self-propelling, self-navigating colonoscope," *Gastroenterology*, vol. 130, no. 3, pp. 672–677, 2006.

[154] M. Cianchetti, T. Ranzani, G. Gerboni, T. Nanayakkara, K. Althoefer, P. Dasgupta, and A. Menciassi, "Soft robotics technologies to address shortcomings in today's minimally invasive surgery: the stiff-flop approach," *Soft robotics*, vol. 1, no. 2, pp. 122–131, 2014.

[155] A. Arezzo, Y. Mintz, M. E. Allaix, S. Arolfo, M. Bonino, G. Gerboni, M. Brancadoro, M. Cianchetti, A. Menciassi, H. Wurdemann, *et al.*, "Total mesorectal excision using

a soft and flexible robotic arm: a feasibility study in cadaver models," *Surgical endoscopy*, vol. 31, no. 1, pp. 264–273, 2017.

[156] T. George Thuruthel, Y. Ansari, E. Falotico, and C. Laschi, "Control strategies for soft robotic manipulators: A survey," *Soft robotics*, vol. 5, no. 2, pp. 149–163, 2018.

[157] D. B. Camarillo, C. F. Milne, C. R. Carlson, M. R. Zinn, and J. K. Salisbury, "Mechanics modeling of tendon-driven continuum manipulators," *IEEE transactions on robotics*, vol. 24, no. 6, pp. 1262–1273, 2008.

[158] R. S. Penning, J. Jung, N. J. Ferrier, and M. R. Zinn, "An evaluation of closed-loop control options for continuum manipulators," in *2012 IEEE International Conference on Robotics and Automation*, pp. 5392–5397, IEEE, 2012.

[159] M. Luo, Y. Pan, E. H. Skorina, W. Tao, F. Chen, S. Ozel, and C. D. Onal, "Slithering towards autonomy: a self-contained soft robotic snake platform with integrated curvature sensing," *Bioinspiration & biomimetics*, vol. 10, no. 5, p. 055001, 2015.

[160] E. L. White, J. C. Case, and R. Kramer-Bottiglio, "A soft parallel kinematic mechanism," *Soft robotics*, vol. 5, no. 1, pp. 36–53, 2018.

[161] D. C. Rucker and R. J. Webster, "Mechanics of continuum robots with external loading and general tendon routing," in *Experimental Robotics*, pp. 645–654, Springer, 2014.

[162] J. Till, C. E. Bryson, S. Chung, A. Orekhov, and D. C. Rucker, "Efficient computation of multiple coupled cosserat rod models for real-time simulation and control of parallel continuum manipulators," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5067–5074, IEEE, 2015.

[163] F. Largilliere, V. Verona, E. Coevoet, M. Sanz-Lopez, J. Dequidt, and C. Duriez, "Real-time control of soft-robots using asynchronous finite element modeling," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2550–2555, IEEE, 2015.

[164] M. Giorelli, F. Renda, M. Calisti, A. Arienti, G. Ferri, and C. Laschi, "Neural network and jacobian method for solving the inverse statics of a cable-driven soft arm with nonconstant curvature," *IEEE Transactions on Robotics*, vol. 31, no. 4, pp. 823–834, 2015.

[165] T. George Thuruthel, E. Falotico, M. Manti, A. Pratesi, M. Cianchetti, and C. Laschi, "Learning closed loop kinematic controllers for continuum manipulators in unstructured environments," *Soft robotics*, vol. 4, no. 3, pp. 285–296, 2017.

[166] M. C. Yip and D. B. Camarillo, "Model-less feedback control of continuum manipulators in constrained environments," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 880–889, 2014.

[167] K.-H. Lee, D. K. Fu, M. C. Leong, M. Chow, H.-C. Fu, K. Althoefer, K. Y. Sze, C.-K. Yeung, and K.-W. Kwok, "Nonparametric online learning control for soft continuum robot: An enabling technique for effective endoscopic navigation," *Soft robotics*, vol. 4, no. 4, pp. 324–337, 2017.

[168] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Computed torque control with nonparametric regression models," in *2008 American Control Conference*, pp. 212–217, IEEE, 2008.

[169] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE transactions on robotics and automation*, vol. 12, no. 5, pp. 651–670, 1996.

[170] K. Hashimoto, *Visual servoing*, vol. 7. World scientific, 1993.

[171] M. Azizian, M. Khoshnam, N. Najmaei, and R. V. Patel, "Visual servoing in medical robotics: a survey. part i: endoscopic and direct vision imaging–techniques and applications," *The international journal of medical robotics and computer assisted surgery*, vol. 10, no. 3, pp. 263–274, 2014.

[172] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE Robotics & Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.

[173] Y.-H. Liu, H. Wang, C. Wang, and K. K. Lam, "Uncalibrated visual servoing of robots using a depth-independent interaction matrix," *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 804–817, 2006.

[174] H. Wang, W. Chen, X. Yu, T. Deng, X. Wang, and R. Pfeifer, "Visual servo control of cable-driven soft robotic manipulator," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 57–62, IEEE, 2013.

[175] Y. Lu, C. Zhang, S. Song, and M. Q.-H. Meng, "Precise motion control of concentric-tube robot based on visual servoing," in *2017 IEEE International Conference on Information and Automation (ICIA)*, pp. 299–304, IEEE, 2017.

[176] J. D. Greer, T. K. Morimoto, A. M. Okamura, and E. W. Hawkes, "Series pneumatic artificial muscles (spams) and application to a soft continuum robot," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5503–5510, IEEE, 2017.

[177] H. Wang, B. Yang, Y. Liu, W. Chen, X. Liang, and R. Pfeifer, "Visual servoing of soft robot manipulator in constrained environments with an adaptive controller," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 41–50, 2016.

[178] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.

[179] C. Rasmussen and C. Williams, "Gaussian processes for machine learning.,(mit press: Cambridge, ma)," 2006.

[180] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Model learning with local gaussian process regression," *Advanced Robotics*, vol. 23, no. 15, pp. 2015–2034, 2009.

[181] G. Fang, X. Wang, K. Wang, K.-H. Lee, J. D. Ho, H.-C. Fu, D. K. C. Fu, and K.-W. Kwok, "Vision-based online learning kinematic control for soft robots using local gaussian process regression," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1194–1201, 2019.

[182] R. M. Viglialoro, N. Esposito, S. Condino, F. Cutolo, S. Guadagni, M. Gesi, M. Ferrari, and V. Ferrari, "Augmented reality to improve surgical simulation: Lessons learned towards the design of a hybrid laparoscopic simulator for cholecystectomy," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 7, pp. 2091–2104, 2018.

[183] J. A. Paydarfar, X. Wu, and R. J. Halter, "Initial experience with image-guided surgical navigation in transoral surgery," *Head & neck*, vol. 41, no. 1, pp. E1–E10, 2019.

[184] H. Abidi, G. Gerboni, M. Brancadoro, J. Fras, A. Diodato, M. Cianchetti, H. Wurdemann, K. Althoefer, and A. Menciassi, "Highly dexterous 2-module soft robot for intra-organ navigation in minimally invasive surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 14, no. 1, p. e1875, 2018.

[185] N. Andreff and B. Tamadazte, "Laser steering using virtual trifocal visual servoing," *The International Journal of Robotics Research*, vol. 35, no. 6, pp. 672–694, 2016.

[186] G. Dagnino, L. S. Mattos, and D. G. Caldwell, "A vision-based system for fast and accurate laser scanning in robot-assisted phonomicrosurgery," *International journal of computer assisted radiology and surgery*, vol. 10, no. 2, pp. 217–229, 2015.

[187] A. M. Andrew, "Multiple view geometry in computer vision," *Kybernetes*, 2001.

[188] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[189] N. Punhani, G. R. Dongarwar, H. Mahajan, M. J. Daniel, K. Chalapathi, A. S. Nayyar, *et al.*, "Tumor size and its relation to cervical lymph node metastasis and its significance as a prognostic indicator for oral squamous cell carcinomas," *Clinical Cancer Investigation Journal*, vol. 6, no. 3, p. 153, 2017.

[190] J.-A. Seon, B. Tamadazte, and N. Andreff, "Decoupling path following and velocity profile in vision-guided laser steering," *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 280–289, 2015.

[191] D. B. Comber, J. E. Slightam, V. R. Gervasi, J. S. Neimat, and E. J. Barth, "Design, additive manufacture, and control of a pneumatic mr-compatible needle driver," *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 138–149, 2016.

[192] W. I. Wei and W.-K. Ho, "Transoral robotic resection of recurrent nasopharyngeal carcinoma," *The Laryngoscope*, vol. 120, no. 10, pp. 2011–2014, 2010.

[193] L. Fichera, *Cognitive supervision for robot-assisted minimally invasive laser surgery*. Springer Nature, 2016.

[194] R. V. Martinez, C. R. Fish, X. Chen, and G. M. Whitesides, "Elastomeric origami: programmable paper-elastomer composites as pneumatic actuators," *Advanced functional materials*, vol. 22, no. 7, pp. 1376–1384, 2012.

[195] Z. Wang, P. Polygerinos, J. T. Overvelde, K. C. Galloway, K. Bertoldi, and C. J. Walsh, "Interaction forces of soft fiber reinforced bending actuators," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 2, pp. 717–727, 2016.

[196] Z. Bubnicki, *Modern control theory*, vol. 2005925392. Springer, 2005.

[197] Y. Nakayama, *Introduction to fluid mechanics*. Butterworth-Heinemann, 2018.

[198] D. A. Thorley, *Fluid transients in pipeline systems*. No. BOOK, ASME Press, 2004.

[199] ASTM, "Standard test method for evaluation of mr image artifacts from passive implants," 2001.

[200] N. E. M. Association *et al.*, "Determination of signal-to-noise ratio (snr) in diagnostic magnetic resonance imaging," *NEMA Standards Publication MS 1-2001*, 2001.

[201] K. Chinzei and K. Miller, "Mri guided surgical robot," in *Australian conference on robotics and automation*, pp. 50–55, 2001.

[202] S. Toupin, P. Bour, M. Lepetit-Coiffé, V. Ozenne, B. D. de Senneville, R. Schneider, A. Vaussy, A. Chaumeil, H. Cochet, F. Sacher, *et al.*, "Feasibility of real-time mr thermal dose mapping for predicting radiofrequency ablation outcome in the myocardium in vivo," *Journal of Cardiovascular Magnetic Resonance*, vol. 19, no. 1, pp. 1–12, 2017.

[203] Z. Zhang, T. Michaelis, and J. Frahm, "Towards mri temperature mapping in real time—the proton resonance frequency method with undersampled radial mri and nonlinear inverse reconstruction," *Quantitative imaging in medicine and surgery*, vol. 7, no. 2, p. 251, 2017.

[204] A. L. Gough-Palmer and W. M. W. Gedroyc, "Laser ablation of hepatocellular carcinoma-a review," *World journal of gastroenterology: WJG*, vol. 14, no. 47, p. 7170, 2008.

[205] M. O. Köhler, C. Mougenot, B. Quesson, J. Enholm, B. Le Bail, C. Laurent, C. T. Moonen, and G. J. Ehnholm, "Volumetric hifu ablation under 3d guidance of rapid mri thermometry," *Medical physics*, vol. 36, no. 8, pp. 3521–3535, 2009.

[206] N. J. Mickevicius and E. S. Paulson, "Simultaneous orthogonal plane imaging," *Magnetic resonance in medicine*, vol. 78, no. 5, pp. 1700–1710, 2017.