Robots intraluminaux flexibles et orientables lors de procédures complexes : assistance par un contrôle de haut niveau

Fernando Gonzalez-Herrera

UNIVERSITÉ DE STRASBOURG

DOCTORAL SCHOOL MSII
ICube Laboratory (UMR 7357)
and KU Leuven

THÈSE

presented by :

GONZALEZ HERRERA, Jose Fernando

Defended on: **Septembre 6th, 2024**

For obtaining the degree of **Doctor of Philosophy**

from the University of Strasbourg

Discipline: **Robotics**

Flexible steerable intraluminal robots during complex procedures: assistance through high-level control

Thesis directors:

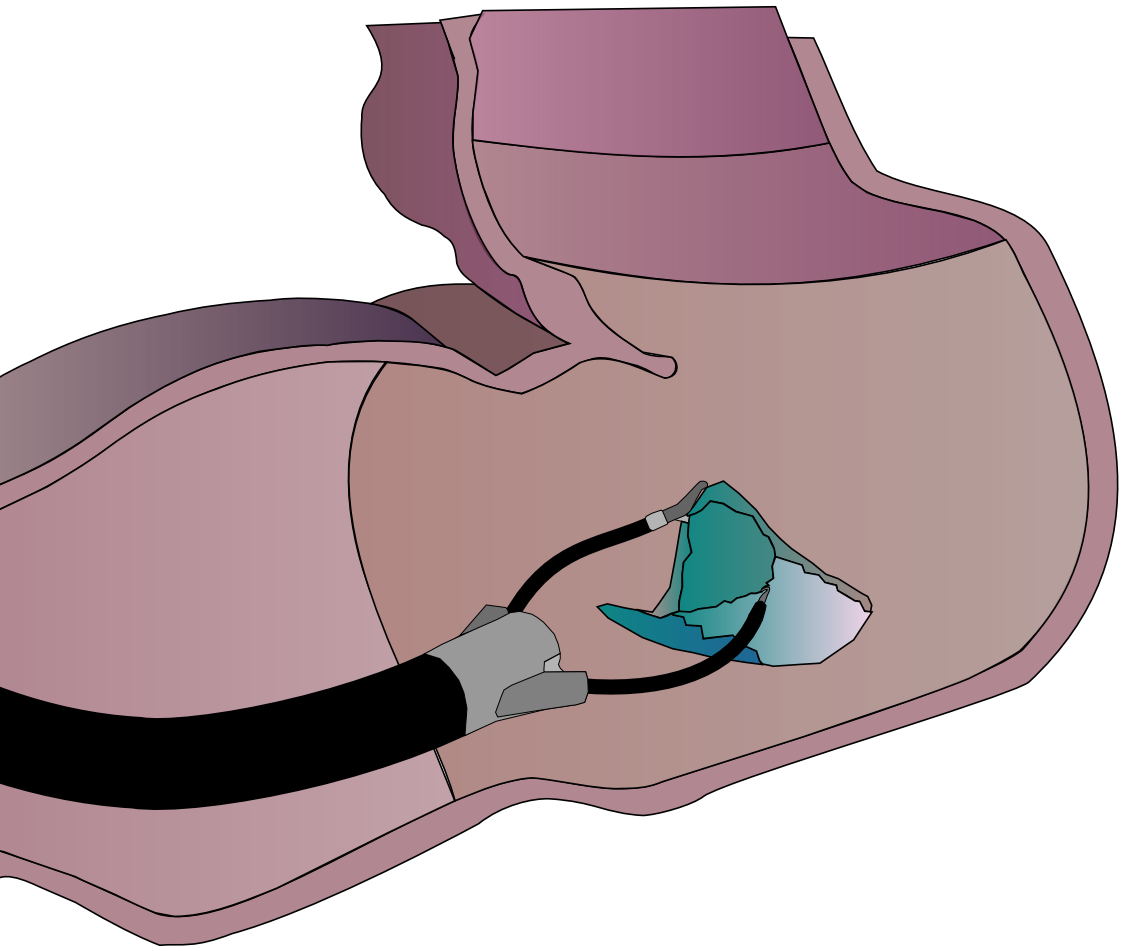
Prof. De Mathelin, Michel	Université de Strasbourg
Prof. De Schutter, Joris	KU Leuven

Reviewers:

Prof. Casals Gelpí, Alicia	Universitat Politècnica de Catalunya
Prof. Zemiti, Nabil	Université de Montpellier

Examiners:

Prof. Vander Poorten, Emmanuel	KU Leuven
Prof. Laroche, Edouard	Université de Strasbourg
Borghesan, Gianni	KU Leuven
Rosa, Benoit	Research Scientist, CNRS



Flexible steerable intraluminal robots during complex procedures: assistance through high-level control

Jose Fernando Gonzalez Herrera

Dissertation presented in partial fulfillment of
the requirements for the degree of:
Docteur at Université de Strasbourg; and of
Doctor of Engineering Science (PhD) at KU Leuven, Belgium

KU LEUVEN

Université de Strasbourg

UNIVERSITE DE STRASBOURG
and
KU LEUVEN

Flexible steerable intraluminal robots during complex procedures: assistance through high-level control

by

Jose Fernando Gonzalez Herrera

degree of Doctor of Engineering

in the

École doctorale Mathématiques, sciences de l'information et de l'ingénieur
Université de Strasbourg

October 2024

Preface

This manuscript is the result of several years of studies, as part of a joint doctorate between the Université de Strasbourg, as part of the RDH-ICube lab, and KU Leuven, as part of the Robot-Assisted Surgery (RAS) Group. Before presenting the research carried out over this *extended* time, I will use the next few lines to honour the help and support that was given to my journey.

First of all, I would like to thank the members of the examination committee: Edouard Laroche (UNISTRA), Joris De Schutter (KUL), Michel de Mathelin (UNISTRA), Gianni Borghesan (KUL), Benoit Rosa (UNISTRA), Emmanuel Vander Poorten (KUL), Alicia Casals (Univ. Barcelone), Nabil Zemiti (Univ. Montpellier). It was an honour to have my work evaluated by such multidisciplinary and influential community members. Furthermore, I would like to thank also an additional member of the supervision committee and the midterm evaluation: Florent Nageotte, your insight and your feedback during the early stages helped me shape a proposal that is valuable for engineers and surgeons. Synergically, the feedback from all of you helped me craft a valuable solution. Thanks to Prof. de Mathelin and Prof. de Schutter, your feedback at different stages of this project was essential during the development of the core idea of the thesis. Special thanks to Benoit for being patient enough to test the multiple versions that were developed, your insight was key for developing a coherent, robust and simple enough solution that provides value in robotic assistance. Thanks to Gianni and Manu for providing fast and objective feedback during my PhD trajectory. I would also like to thank Philippe Zanne for helping me troubleshoot and solve technical issues with the STRAS system and for helping me at different stages of the proposal.

Secondly, I would like to thank all the clinical personnel at IHU as well as all the members of the ICube-RDH lab who tested multiple versions of the system and provided honest feedback on the proposal, without such objective feedback the solution would not be of value for both engineers and surgeons. Special thanks to the surgeons who were honest about the early developments of the approach, your valuable insight allowed me to improve the solution.

Thirdly, I would like to state that this work was supported by the ATLAS project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813782. This work was also partially supported by French State Funds managed by the Agence Nationale de la Recherche (ANR), Grant ANR-18-CE19-0012 (MACROS) and Grant ANR-10-IAHU-02 (IHU-Strasbourg).

Lastly and personally, I would like to thank all of my colleagues (master's and PhD students, and postdocs) at the ICube laboratory, and also mention those at the CAMMA lab, as well as those in the RAS group. Thanks to all the ATLAS ITN PhD fellows who

allowed me to share academic roads. Thanks to the members of the C1-Colonoscopy group, as part of the ATLAS ITN network, for your teamwork and collaboration in integrating the robotic colonoscopy system. Special thanks, in no particular order, to Vlad, Jorge, Beatriz, Guiqiu, and Sujit; you all made the journey joyful and fun even in the most challenging times. I would also like to thank my family for the support, without you, it would not have been possible, thanks. Por último gracias a mi buebi, gracias por ser y por estar.

Abstract

Flexible steerable intraluminal robots have been developed to navigate, inspect and perform complex surgical procedures inside the patient's digestive system, reducing the high incidence and mortality rate of colorectal cancer. Flexible steerable intraluminal robots are typically composed of two arms at the tip of the body, forming a coupled kinematic architecture with up to 10 degrees of freedom. This thesis focuses on the development of robotic assistance to handle the challenges (the coupled architecture and the high number of degrees of freedom) of flexible steerable intraluminal robots in complex surgical environments, reducing the risk of incorrectly performing complex surgical procedures, such as the dissection stage of endoscopic submucosal dissection and the colonoscopy and biopsy case. The surgical tasks are modelled and formulated as objectives and constraints in an optimization-based control framework based on the task-function approach. A validation study with 10 subjects for the dissection case shows a better performance of the surgical task at a time penalty, in comparison with the manual mode in a simulated environment. In the colonoscopy and biopsy case, results show that for a realistic colon model, the surgical task can be performed automatically by coordinated control in a benchtop experiment. Future work should perform in-depth validation of the proposed control approaches for preclinical and clinical transfer.

Resume

Des robots intraluminaux flexibles et orientables ont été mis au point pour naviguer, inspecter et effectuer des procédures chirurgicales complexes à l'intérieur du système digestif du patient, réduisant ainsi l'incidence élevée et le taux de mortalité du cancer colorectal. Ces robots sont généralement composés de deux bras à l'extrémité du corps, formant une architecture cinématique couplée avec jusqu'à 10 degrés de liberté. Cette thèse se concentre sur le développement d'une assistance robotique pour relever les défis (l'architecture couplée et le nombre élevé de degrés de liberté) des robots intraluminaux flexibles et orientables dans des environnements chirurgicaux complexes, en réduisant les risque associés à des procédures chirurgicales complexes, telles que l'étape de dissection de la dissection sous-muqueuse endoscopique et le cas de la coloscopie et de la biopsie. Une modélisation des tâches chirurgicales est proposée afin de les formuler sous forme d'objectifs et de contraintes dans une commande basée optimisation. Une étude de validation avec 10 sujets pour le cas de la dissection montre une meilleure performance de la tâche chirurgicale avec une légère pénalité temporelle, en comparaison avec le mode manuel dans un environnement simulé. Dans le cas de la coloscopie et de la biopsie, les résultats montrent que pour un modèle de côlon réaliste, la tâche chirurgicale peut être exécutée automatiquement grâce au contrôle coordonné proposé. Les travaux futurs devront réaliser une validation approfondie des approches de commande proposées en vue d'un transfert préclinique et clinique.

Beknopte samenvatting

Flexibele bestuurbare intraluminale robots zijn ontwikkeld om te navigeren, inspecteren en complexe chirurgische procedures uit te voeren in het spijsverteringsstelsel van de patiënt, om de hoge incidentie en mortaliteit van darmkanker te verminderen. Flexibele bestuurbare intraluminale robots bestaan meestal uit twee armen aan het uiteinde van het lichaam, die een gekoppelde kinematische architectuur vormen met maximaal 10 vrijheidsgraden. Dit proefschrift richt zich op de ontwikkeling van robotassistentie om de uitdagingen (de gekoppelde architectuur en het hoge aantal vrijheidsgraden) van flexibele bestuurbare intraluminale robots in complexe chirurgische omgevingen aan te kunnen, waardoor het risico op het verkeerd uitvoeren van complexe chirurgische procedures, zoals de dissectiefase van endoscopische submucosale dissectie en de colonoscopie en biopsiecasus, wordt verminderd. De chirurgische taken worden gemodelleerd en geformuleerd als doelstellingen en beperkingen in een op optimalisatie gebaseerd besturingsraamwerk op basis van de taakfunctiebenadering. Een validatiestudie met 10 proefpersonen voor de dissectiecasus toont een betere prestatie van de chirurgische taak met een tijdstraf, in vergelijking met de handmatige modus in een gesimuleerde omgeving. In het geval van colonoscopie en biopsie laten de resultaten zien dat voor een realistisch colonmodel de chirurgische taak automatisch kan worden uitgevoerd door gecoördineerde besturing in een laboratoriumexperiment. In de toekomst moeten de voorgestelde besturingsbenaderingen grondig worden gevalideerd voor preklinische en klinische overdracht.

Contents

Preface i

Abstract iii

Resume iv

Beknopte samenvatting v

List of Figures ix

List of Tables xi

List of Acronyms xii

List of Symbols xiv

1 Introduction 1
- 1.1 From traditional surgery to no-scar minimally invasive surgery . . . 1
- 1.2 Flexible steerable intraluminal robots . . . 5
- 1.3 Research objectives . . . 7
- 1.4 Challenges in the control of flexible steerable intraluminal robotic platforms 8
- 1.5 Research question . . . 8
- 1.6 Thesis overview . . . 9

2 State of the art 11
- 2.1 Medical context . . . 11
  - 2.1.1 Identification of colorectal cancer . . . 12
  - 2.1.2 Removal of colorectal cancer . . . 14
- 2.2 Flexible steerable intraluminal platforms . . . 16
- 2.3 Current challenges of FSIR platforms . . . 19
- 2.4 Robotic assistance during complex surgical procedures inside the body . . 22
  - 2.4.1 Robotic assistance in minimally invasive surgery . . . 22
  - 2.4.2 Autonomous endoscope control of laparoscopes . . . 24
  - 2.4.3 Autonomous control of arms of laparoscopes . . . 26
- 2.5 Remaining challenges in FSIR and thesis contributions . . . 27

3 STRAS system **29**
3.1 Development of the STRAS robotic system . . . 29
3.2 STRAS system overview . . . 31
3.3 Modelling of FSIR platforms applied on the STRAS system . . . 34
3.3.1 Introduction . . . 34
3.3.2 Constant curvature model . . . 35
3.3.3 Forward kinematic modelling . . . 36
3.3.4 Differential kinematic modelling . . . 39
3.3.5 Workspace characterization . . . 41
3.4 STRAS system telemanipulation . . . 41
3.4.1 STRAS system ranges . . . 43
3.5 The STRAS system as validation platform . . . 44

4 Assistance through high-level control in flexible steerable intraluminal platforms **46**
4.1 Endoscopic submucosal dissection by the FSIR platforms . . . 47
4.2 Problem modelling . . . 49
4.3 Overview of the problem . . . 51
4.4 Related works . . . 53
4.4.1 Control frameworks . . . 54
4.4.2 Quadratic formulation . . . 57
4.5 Semi-autonomous arm-body control approach . . . 59
4.5.1 Overview . . . 59
4.5.2 Formulation of the control problems as objectives . . . 63
4.5.3 Optimization-Based High-Level Control . . . 64
4.5.4 Summary of behaviours . . . 67
4.6 Experiments . . . 69
4.6.1 Experimental setup . . . 70
4.6.2 Experimental task . . . 71
4.6.3 Protocol . . . 72
4.6.4 Quantitative evaluation metrics . . . 73
4.6.4.1 Duration metrics . . . 73
4.6.4.2 Kinematics metrics . . . 74
4.6.4.3 Surgical performance metrics . . . 75
4.7 Results . . . 76
4.8 Discussion and future work on semi-autonomous polyp removal . . . 79

5 Assisted colonoscopy in flexible steerable intraluminal robots **86**
5.1 Problem statement . . . 88
5.1.1 Current state of colonoscopy and biopsy . . . 88
5.1.2 Optical biopsies using OCT . . . 90
5.1.3 Problem overview . . . 92
5.2 Related works . . . 94
5.2.1 Visual servoing . . . 94
5.2.2 Assisted navigation of the digestive system . . . 96
5.2.3 Assisted in-situ scanning for tissue identification . . . 97
5.2.4 Conclusion and contributions . . . 98

5.3 Colonoscopy and biopsy performed by flexible steerable intraluminal robotic platforms . . . 98
5.3.1 Problem statement . . . 99
5.3.2 Problem modelling . . . 101
5.3.3 Overview of the approach . . . 102
5.4 Experiments . . . 105
5.4.1 Experimental setup . . . 106
5.5 Results . . . 108
5.5.1 Individual validation of each of the tasks . . . 108
5.5.2 Lumen alignment and arm in largest abnormal tissue . . . 108
5.6 Discussions and future work . . . 111

6 Conclusion **115**

7 List of publications of this thesis **121**

A NASA Task Load Index **122**

B Extra concepts **124**
B.1 Smoothness by spectral arc length . . . 124
B.2 Effect size . . . 125

C Résumé en français **126**
C.1 Introduction . . . 126
C.2 Robots intraluminaux flexibles et orientables . . . 127
C.3 Problème . . . 128
C.4 Proposition . . . 130
C.5 Résultats . . . 132
C.6 Conclusion . . . 133

Bibliography **136**

List of Figures

1.1 Overview of a colonoscopy procedure performed with manual endoscopes . 2
1.2 Overview of a minimally invasive procedure performed with manual endoscopes . . . 5
1.3 Minimally invasive surgery inside the digestive system using manual endoscopes. . . . 5
1.4 FLEX robotic platform. . . . 7

2.1 Colonoscopy overview. . . . 13
2.2 Overview of endoscopic mucosal resection. . . . 15
2.3 Overview of endoscopic mucosal resection. . . . 16
2.4 Flexible endoscopes and flexible steerable intraluminal robots. . . . 17
2.5 Examples of flexible steerable intraluminal robots. . . . 18
2.6 Commonly proposed levels of autonomy for minimally invasive surgery. . . . 23
2.7 Overview of the common approaches for autonomous laparoscopy. . . . 25

3.1 Anubiscope system developed by Karl Storz. . . . 30
3.2 Motorized modules for actuation of the arms. . . . 31
3.3 Overview of the STRAS system. . . . 32
3.4 STRAS arms from the endoscopic camera. . . . 33
3.5 Handle of the leader side of the STRAS system. . . . 33
3.6 The three spaces and mappings of the constant curvature model. . . . 35
3.7 Overview of the constant curvature model in the STRAS robot. . . . 36
3.8 Frames of interest of the STRAS system. . . . 37
3.9 Tip positions of the arms and body of the STRAS system. . . . 38
3.10 Theoretical workspace of the arms from the top view. . . . 42
3.11 Workspace of the dominant arm. . . . 42
3.12 Leader side of the STRAS system. . . . 44

4.1 Overview of the envisioned scenario. . . . 48
4.2 Overview of the problem. . . . 49
4.3 Overview of the control problem. . . . 52
4.4 Example similar arm-body robotic structures. . . . 54
4.5 Overview of the main components of the simulator setup. . . . 61
4.6 Overview of the defined safe zones for the dominant arm. . . . 65
4.7 Overview of the control framework. . . . 68
4.8 Experimental setup used. . . . 71
4.9 Overview of manual and semi-autonomous mode. . . . 79
4.10 NASA TLX results from the manual and semi-autonomous arm-body control mode. . . . 80

4.11 Example of distances of the arms to the objectives in manual and semi-autonomous arm-body control mode 82

5.1 Diggestive system, as is explored during colonoscopy. 89
5.2 Degrees of freedom in manual flexible colonoscopes. 89
5.3 OCT schematic and OCT images overview. 91
5.4 Rotational scanning by the OCT with healthy and abnormal tissue. . . . 93
5.5 Visual servoing configurations. a) eye-in-hand; b) eye-to-hand 95
5.6 Overview of the colonoscopy and biopsy. 99
5.7 Overview of the proposed solution for assisted colonoscopy and biopsy. . . 103
5.8 Overview of the visual detection module outputs. 104
5.9 Overview of the colon model. 107
5.10 Individual validation of the $\mathscr{B}_1$ and $\mathscr{B}_2$ objectives. 109
5.11 Experiment 1: Tracking errors during the experiment 110
5.12 Experiment 2: Tracking errors during the experiment 111

C.1 Architecture commune aux plateformes FSIR. 128
C.2 Exemples de structures robotiques bras-corps. 129
C.3 Vue d'ensemble du robot STRAS. 130
C.4 Vue d'ensemble des bras à partir de la caméra corporelle. 131
C.5 Aperçu du cadre de contrôle. 132
C.6 Exemple de distances entre les bras et les objectifs en mode manuel et semi-autonome de contrôle bras-corps 134
C.7 Expérience 1 : Erreurs de suivi pendant l'expérience 135
C.8 Expérience 2 : Erreurs de suivi pendant l'expérience 135

List of Tables

3.1 Ranges of the STRAS system leader, follower and configuration space. . . 45

4.1 Table of the behaviours of the assisted telemanipulation control mode. . . 69
4.2 Parameters used in the experimental setup. 71
4.3 Summary of the results of the experiments with the semi-autonomous arm-body and the manual telemanipulation mode. 77

5.1 Parameters used in the experimental setup for assisted colonoscopy. . . . 108

List of Acronyms

2D Two dimensional.

3D Three dimensional.

CCM Constant curvature model.

CRC Colo rectal cancer.

DoF Degree(s) of Freedom.

EASE Endoluminal Assistant for Surgical Endoscopy.

EMR Endoscopic submucosal ressection.

ERP Endoluminal robotic platforms.

ESD Endoscopic submucosal dissection.

FKM Forward kinematic model.

FOV Field of view.

FSIR Flexible steerable intraluminal robots.

GI Gastrointestinal.

HTM Homogeneous transformation matrix.

IBVS Image-based visual servoing.

IKM Inverse kinematic model.

IQR Interquartile range.

LoA Level of Autonomy.

MIS Minimally invasive surgery.

NOTES Natural Orifice Transluminal Endoscopic Surgery.

OCT Optical coherence tomography.

OR Operating room.

PBVS Position-based visual servoing.

QP Quadratic programming.

SPARC Spectral arc length measure.

STRAS Single port and Transluminal Robotic Assistant for Surgeons.

TLX Task load index.

VS Visual servoing.

WQP Weighted Quadratic Programming.

List of Symbols

$\boldsymbol{c}$	Constant vector
$\boldsymbol{K}$	Diagonal matrix of proportional gains
$\boldsymbol{e}$	Error vector
$\boldsymbol{\gamma}$	Gamma vector
$\boldsymbol{H}_t$	Hessian matrix at t time for the QP controller
${}_b\boldsymbol{T}^a$	Homogeneous transformation matrix expressing frame a with respect to frame b
${}_b\boldsymbol{J}_a$	Task Jacobian of frame a expressed in frame b
$\mathscr{B}_i$	Optimization objective i for the colonoscopy and biopsy contribution
$\mathscr{H}_i$	Optimization objective i for the dissecction contribution
${}_b\boldsymbol{p}^a$	Position of frame a with respect to frame b
$\{a\}$	Reference frame a
k_i	Scalar proportional gain i
γ_i	Scalar gamma value for objective i
$\dot{\boldsymbol{q}}_a$	Speeds for the a in the leader side
${}_b\dot{\boldsymbol{p}}^a$	Speeds of frame a with respect to frame b
$\boldsymbol{q_a}$	Vector of configuration variables for the a *leader* side
$\dot{\boldsymbol{q}}_{\boldsymbol{a}}$	Vector of joint speeds for a
$\dot{\boldsymbol{q}}_a$	Vector of joint speeds for a virtual a
$\boldsymbol{k}$	Vector of proportional gains

Chapter 1

Introduction

1.1 From traditional surgery to no-scar minimally invasive surgery

Traditional surgery, often referred to as open-body or conventional surgery, relies on opening the body of patients to perform surgical procedures. Traditional surgery requires performing large incisions on the patient to allow surgeons to reach and have a complete view of the surgical site to carry out surgical gestures inside the patient. The large opening allows knives, scalpels, and graspers to be positioned freely whereas the patient's tissue and skin is being held in place by surgical tools —retractors and hooks. Large incisions pose a heightened risk of infection (Shabanzadeh and Sørensen, 2012) and necessitate extended recovery periods lasting at least a couple of days (Fisher, 2004). Traditional surgery typically involves sedating the patient and administering pain medication to manage discomfort during the prolonged recovery phase, which is worsen by the body's strain from the incisions. An example of traditional surgery is open colectomy. In this procedure, the abdomen of the patient is cut open and the infected part of the patient's digestive system is removed. After the open colectomy, the user necessitates a couple of days to a full week of hospital rest to fully recover. After leaving the hospital, patients will need to rest at home, which can limit their daily activities due to reduced mobility. Overall, traditional surgery such as open colectomy is a long and painful procedure for patients who have to face unnecessary secondary effects due to the anaesthesia and the large and noticeable scars (Schäfer et al., 1998).

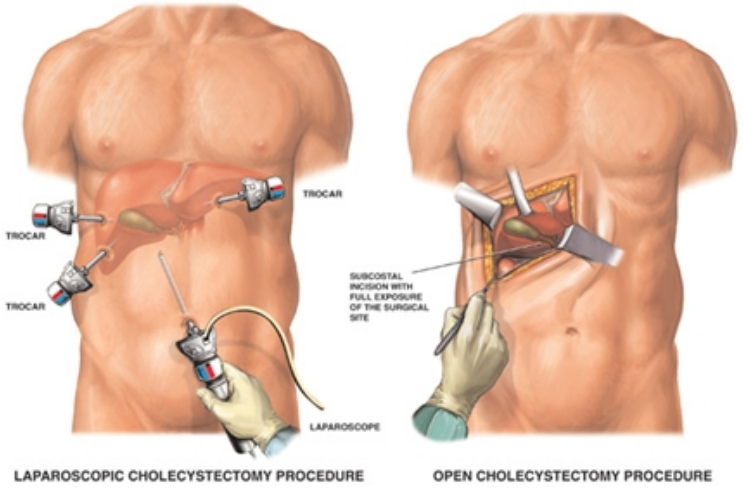


FIGURE 1.1: Overview of a colonoscopy procedure performed with manual endoscopes. Image adapted from WCH (2020).

Minimally invasive surgery (MIS) poses an improvement over traditional surgery as no large incisions are made (Fuchs, 2002), as depicted in fig. 1.1. Minimally invasive surgery employs small incisions in the body of patients that allow surgeons to pass small surgical instruments, accessories, and miniature cameras inside the patient (Schäfer et al., 1998). The surgical site can be viewed through cameras, while surgical instruments and accessories are used to operate inside the patient's body through the incisions. The surgeon can interact with the tissue from the outside by cutting, marking, pulling, or pushing the tissue from the patient's inside while moving through the small incisions. Each incision allows one or more surgical tools to be inserted. A camera is inserted through a dedicated incision to allow visualization of the surgical site and tools. An example of minimally invasive surgery is laparoscopic polypectomy. This procedure requires at least 3 small incisions, one for the camera and another two for the surgical tools. Compared to open colectomy, laparoscopic polypectomy is more precise as it removes only the infected part of the patient's digestive system, instead of the large block (Aimaq et al., 2011). Most of the laparoscopy polypectomy surgeries require local or general anaesthesia. In some cases, the procedure might even require monitored sedation for a few days after the surgery, allowing patients to endure the pain (Stiff et al., 1994).

Recovery time is reduced even further when the natural orifices of patients are used to perform minimally invasive surgery (Schwaitzberg et al., 2018). Using natural orifices –instead of performing incisions in the patient's body– reduces the discomfort of patients

(Zhu et al., 2021) and in some cases allows no-sedation screening procedures. The mouth, nose, urethra, or anus can be used to pass surgical instruments and cameras inside the body of the patient to perform minimally invasive surgery. Using only natural orifices of patients leaves no scars, which is why this type of surgery is commonly referred to as '*no-scar surgery*'. No-scar surgery requires flexible and slender devices inserted through patients' natural orifices. The flexible devices navigate through the tissues and organs of patients and reach the desired surgical site. To navigate the inside of the patient, flexible and slender devices are actuated at their tip which allows them to navigate through the natural paths inside the patient. Once the desired site is reached, the flexible and slender devices can perform the surgery.

Current flexible endoscopes allow surgeons to carry out minimally invasive surgery and procedures inside the body of the patient by reducing the need for incisions (Yoshizumi et al., 2009) when the conditions of the procedure and the patient allows it. A flexible endoscope is a tubular device that allows operators to see inside a patient's body (fig. 1.2). Most flexible endoscopes are narrow and flexible, with a camera at its tip. Flexible endoscopes contain channels that permit the passage of instruments, accessories, and fluids during diagnosis or endoscopic surgery. Flexible endoscopes allow surgeons to navigate inside the patient of the body while screening the digestive system and allowing for tool-tissue interaction. Endoscopic Mucosal Resection (EMR) through flexible endoscopy is a minimally invasive procedure in which surgeons introduce a flexible endoscope that navigates from the anus to the position of the infected tissue in the digestive system of the patient. In EMR, the surgeon uses the tip of the flexible endoscope and a surgical tool at its tip to remove malignant tissue. Surgeons commonly use a snare [1] to wrap the infected tissue, cut it and extract it from the patient's body. In EMR through flexible endoscopy, the recovery period at the hospital can be just 1 or 2 days with fewer movement constraints for the patient, e.g. decreased physical pain when walking or standing, in comparison with laparoscopic polypectomy. The reduced recovery time is due to the lack of incisions, instead using the natural orifices of patients, eliminating the requirement of anaesthesia and post-operative medication. No-scars surgery allow the surgeon to pass all the necessary tools in the flexible endoscope. Patients prefer incision-less

[1] a snare is a rope-like surgical instrument that allows removing tissue by grabbing, holding and cutting the tissue, it can be electronically activated –referred as hot-snare– or manual –referred as cold-snare (Trivedi et al., 2022).

surgery (Varadarajulu et al., 2008; Swanstrom et al., 2009; Kim et al., 2022) to laparoscopic MIS, due to the reduced recovery times and pain throughout the surgery and the recovery time.

No-scar surgery poses many benefits to patients yet is not straightforward for surgeons due to the technical difficulty of the operation (Thele et al., 2008) related to the control of flexible endoscopes (Levy et al., 2008). Flexible endoscopes are slender compliant platforms that can be positioned and oriented in multiple ways. Surgeons are required to invest a high amount of training hours to master their control, starting by assisting with the procedures until enough experience is built. One of the challenges in controlling such devices is the need for 2 hand coordination while adapting to a non-intuitive control mapping. Surgeons control the insertion and rotation with one hand, and the bending in the two orthogonal directions with the other hand. Surgeons command the control handle to control the bending. Knobs at the control handle regulate how much the endoscope is bent and in which direction. Rotation is possible by rotating the flexible endoscope body at the insertion point –commonly referred to as the insertion tube– or by twisting the control handle by the hand. Furthermore, rotation can be extended by twisting in the opposite direction of the control handle and the endoscope insertion point. Due to the technical complexity of their control, the main surgeon focuses on the control of the flexible endoscope, while assistants commonly control surgical tools, such as snares or forceps[2]. To perform surgical procedures through flexible endoscopes, surgeons and assistants have to cooperate and communicate effectively to control the flexible endoscope and the surgical tools, respectively.

Flexible endoscopes allow surgeons to carry out procedures such as endoscopic mucosal resection requiring surgeons and assistants to cooperate in the operating room to correctly perform such procedures. During EMR (depicted in fig. 1.3), surgeons have to control the flexible endoscope while an assistant controls the insertion and actuation of forceps or a snare. Correctly performing EMR involves a surgeon controlling the compliant bodies of flexible endoscopes and an assistant controlling the surgical tool while establishing correct and opportune communication between the two. A high amount of training hours are required for the surgeon and the assistant (Levy et al., 2008; Patel and Thakkar, 2011) to correctly perform EMR through flexible endoscopy (Moghul et al.,

[2]a forceps is a surgical instrument used to grab tissue and needles, it is built of two actuators with shark-like metal teeth

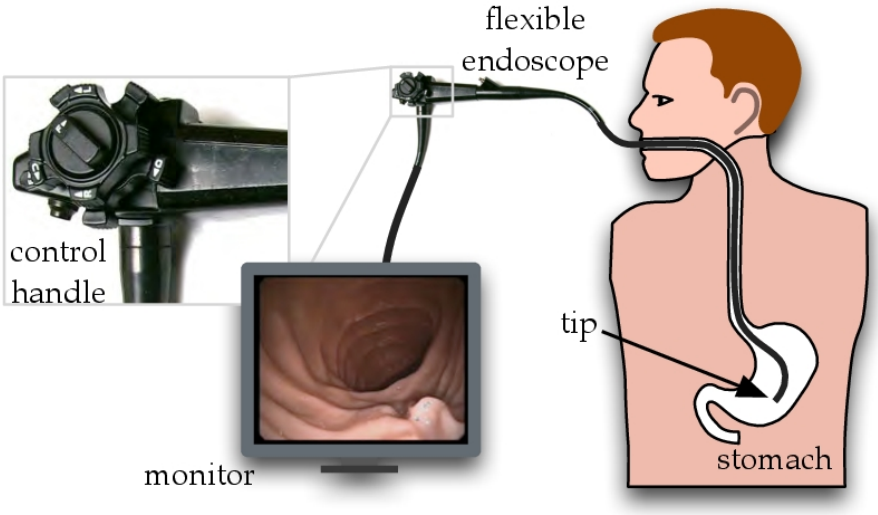


FIGURE 1.2: Overview of a minimally invasive procedure performed with manual endoscopes. Image adapted from Reilink (2013).

2013). EMR through flexible endoscopy could have a prolonged duration which requires a high mental demand to both surgeons and assistants (Zheng et al., 2012). The high mental demand impacts their performance over time and poses a safety risk for patients (Mehta et al., 2014). Control of flexible endoscopes by a single user could reduce the intricacy of surgical procedures such as EMR. Single-user control reduces the number of persons in the operating room, eliminating communication issues. Subsequently, a less crowded operating room reduces errors of surgeons and assistants due to problems in communication and collaboration in mentally demanding procedures.

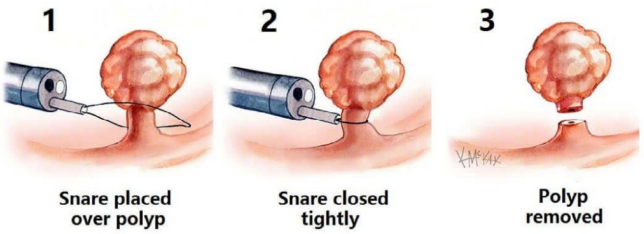


FIGURE 1.3: Minimally invasive surgery inside the digestive system using manual endoscopes. Image adapted from Chiu et al. (2007).

1.2 Flexible steerable intraluminal robots

Flexible steerable intraluminal robotic (FSIR) platforms appeared in the 2010s (Ewing et al., 2004; Pugin et al., 2011; Li and Chiu, 2018; Seeliger and Swanström, 2020) as an evolution of flexible endoscopes with robotic capabilities developed for single-user

control. A main flexible *body* has multiple channels where surgical instruments, actuators and sensors can be placed, referred to as *arms*. Robotic actuation of the body and arms allows surgeons to adopt a telemanipulation configuration, where surgeons are not required to be physically in the same room as the patient, or even in the same country (Marescaux et al., 2001). Telemanipulation is performed by a leader-follower configuration, where surgeons take command of a leader console that controls the follower. The body, arms, and accessories constitute the follower side, while the leader console can be constituted by joysticks, handles, buttons, and/or pedals. Effectively, the surgeons control the leader console that translates their commands by the teleoperation controller to the body and arms inside the patients by the follower side. FSIR platforms should adopt intuitive robotic telemanipulation relationships that map Degrees of Freedom (DoF) of the leader console to the follower side. The configuration with the camera at the tip of the body allows users to focus on controlling the tools as they perform the most important surgical tasks. The arms perform the most important surgical tasks as they have sensors and actuator tools at the end effector. These platforms expand the capabilities of manual endoscopes and can reduce the number of people in the operation room –*with manual endoscopes*– to a single user –*with FSIR*. Platforms such as the k-Flex (Hwang and Kwon, 2020a), FLEX (Mandapathil et al., 2015), and STRAS (De Donno et al., 2013b) allow surgeons to visualize and explore the digestive system with the possibility of acting with it, e.g., as to perform sensing on the tissue or even removing sections of the tissue if it is required. The usage of FSIR platforms poses a mentally-reduced way, in comparison to manual flexible endoscopes or other manual surgical instruments, to perform surgical procedures (Moore et al., 2015) and increase the capabilities of surgeons, however, FSIR platforms still have areas of opportunities to further simplify their control.

FSIR platforms have many DoFs, thus the control of such a highly articulated system may be difficult, as the user can lose track of a specific part of the system while focusing on another, e.g. the simultaneous motion of the two arms and the body in the demanding surgical procedure (Lee et al., 2019; Levy et al., 2008). The body has a camera at its tip, and the arms are attached to it. The relationship between the body and the arms is referred to as *coupled architecture*. Due to the coupled architecture, surgeons have to consider the movement of the body when controlling the arms and vice-versa. Flexible steerable intraluminal robotic platforms use a leader-follower telemanipulation

FIGURE 1.4: FLEX robotic platform. Image adapted from Mandapathil et al. (2015).

configuration that allows the introduction of autonomy to reduce the current challenges. During surgical procedures, the challenges and limitations of FSIR platforms can pose a risk to patient safety. The user has to deal with multiple DoFs –the arms and the body– while it is required to fulfil the interconnected effect they have on each other, the coupled architecture. Dealing with multiple DoFs while considering the coupled architecture is mentally demanding for surgeons. To reduce the amount of time required for surgeons to properly perform surgical tasks with FSIR platforms, telemanipulation could be supported by autonomous robot tasks. This thesis proposes a control framework to assist users with surgical tasks while telemanipulating FSIR. The proposed control framework would allow surgeons to delegate some of the intricate tasks related to an autonomous controller, enabling experts to focus on the more critical parts of the surgical tasks.

1.3 Research objectives

This thesis is framed in the field of robotic-assisted minimally invasive surgery and addresses the surgical procedures in the digestive system. The general objective ($\mathcal{O}$) of this thesis is to *develop a control framework to assist surgeons in controlling flexible, steerable intraluminal robots during complex surgical procedures*. The control framework addresses the drawbacks and challenges of flexible steerable intraluminal robotic platforms and aims to ease control for the user. During the robotic assistance delivered by the proposed control framework, the user remains the main actor in the surgery and can override the automatic control to regain direct control. The proposed approach takes into account the coupled architecture of the FSIR platforms and establishes appropriate

behaviours for the arms and the body based on the surgical task being performed. The control approach is required to fulfil the objectives:

- $\mathcal{O}_1$: reducing the inherent complexity in telemanipulated control of FSIR platforms while executing demanding tasks, typical of minimally invasive surgery,
- $\mathcal{O}_2$: setting appropriate behaviours based on the surgical tasks carried out by each arm's tip set at the end of the surgical arms of FSIR platforms,
- $\mathcal{O}_3$: allowing the surgeon to regain control of the system, overruling the automatically-devised actions, at any time, timely, and with minimal effort, while
- $\mathcal{O}_4$: maintaining patient safety by following the guidelines of no-scar surgery procedures in the digestive system.

1.4 Challenges in the control of flexible steerable intraluminal robotic platforms

Although flexible steerable intraluminal robotic platforms pose benefits in comparison to the control of flexible endoscopes or laparoscopic robotic platforms, challenges to facilitate their control remain. Users telemanipulate flexible steerable intraluminal robotic platforms, which are constructed with flexible bodies containing arms. The identified challenges of FSIR (section 1.2) are: *i)* a high number of degree of freedom, and *ii)* a coupled architecture . In order to address the research objectives, dealing with the identified challenges of FSIR, while following safety guidelines, is required.

1.5 Research question

The general and specific objectives help to define the research question which comes from related works on laparoscopes, catheters and needles, and individual control of the subsystems of FSIR. Since FSIR platforms allow for single-user control of multi-DoF platforms that can perform complex surgical gestures based on their flexible and coupled architecture, an interesting research question is:

How can a control framework assist users to manage the multiple degrees of freedom, coupled architecture, and complex tool-tissue interaction to enable accurate execution of demanding surgical procedures?

This thesis proposes a control framework that assists single-user control of FSIR platforms while dealing with the identified challenges. The control framework is validated in a simulator with a physical leader interface and a physical environment with a physical follower and leader side.

1.6 Thesis overview

To attain the objectives, this thesis is structured in five chapters. Chapter 2 starts by introducing medical context (section 2.1) and overviews the selected surgical scenarios which are considered for the validation of the proposed framework. Then, an overview of FSIR platforms is presented in section 2.2, their challenges are summarised and general assistance during robotic surgery is reviewed. Chapter 2 describes the remaining challenges and outlines the contribution of this thesis. Chapter 3 reviews the STRAS system, which is used to validate the control framework developed here. Chapter 3 describes the architecture, specifications, and modelling of the STRAS system, which is needed to formulate the proposals for the autonomous techniques. Chapter 4 formalizes the first contribution of this thesis, which is a control framework for assisting during the dissection in endoscopic submucosal dissection. The chapter overviews the surgical scenario and allows to perform an overview of the state of the art. Problem modelling and the selected control framework allow the development of a semi-autonomous arm-body control mode which is evaluated in a simulator. This chapter ends with the results, ongoing work, and future directions of the proposed framework. The second surgical scenario considered is the colonoscopy and biopsy case, which are introduced early in chapter 5. The second contribution of this thesis expands the control framework proposed in chapter 4 to further validate it using the physical follower of a FSIR platform The results, conclusions and future work are described at the end of chapter 4. Lastly, this thesis ends with an overall conclusion and future work in chapter 6, which is followed by the list of publications in chapter 7.

Context within the ATLAS project This thesis work has been developed within the AuTonomous intraLuminAl Surgery (ATLAS) project, an Innovative Training Network - European Joint Doctorate project, funded under the Marie Skłodowska-Curie grant agreement N. 813782. This thesis project was developed under the first use-case of the ATLAS project which focuses on minimally invasive surgery on the digestive system. This thesis developed a high-level control scheme that is tested in a standalone surgical case. Furthermore, the high-level control scheme is used in collaboration with the projects of other PhD students within the ATLAS project to target another surgical case. The contributions within the ATLAS ITN project are expanded in chapter 5.

Chapter 2

State of the art

This chapter expands on the clinical scenario which is minimally invasive surgery through natural orifices addressing digestive system diseases. Section 2.1 states the importance of the clinical scenario in identifying and removing abnormal tissue from the digestive system. Section 2.2 describes flexible steerable intraluminal platforms. Then, Section 2.3 introduces the main challenges of FSIR platforms. Following, section 2.4 overviews approaches to general assistance and autonomy in minimally invasive surgery. This chapter ends with section 2.5 stating the remaining challenges that should be addressed by assistance or autonomy in FSIR.

2.1 Medical context

Colorectal cancer (CRC) ranks fourth in terms of both incidence and mortality rates, with a projected 52,550 deaths worldwide in 2023 (Siegel et al., 2023), and increases are anticipated in the coming years. In 2020, colorectal cancer had an incidence rate of 19.5 and a mortality rate of 9.0 per 100,000 persons globally (The International Agency for Research on Cancer (IARC), 2020) and it is predicted Morgan et al. (2023) to increase to 3.2 million new cases and 1.6 million deaths by 2040. Colorectal cancer is mainly caused by a small group of abnormal cells that grow into polyps. These polyps adhere and spread into the colonic walls and cause CRC (De Leon and Di Gregorio, 2001). Polyps in the colon can evolve into metastasis and even spread into the liver or lungs (Regnard et al., 1998). Prompt detection of these abnormal groups of cells allows cancer

to be stopped at an early stage decreasing the mortality rate (Menon and Trudgill, 2014; Than et al., 2015).

2.1.1 Identification of colorectal cancer

To prevent cancer in the colon, the digestive system is explored by outside sensing or by introducing sensors inside the body. From the outside, colorectal cancer can be detected by a computed tomography colonoscopy (Mulhall et al., 2005) which is an un-intrusive procedure when other procedures are not available, however, in computed tomography colonoscopy some areas of the large intestine might not be reached. Another procedure from the outside of the patient is a virtual colonoscopy which uses radiation and results are less sensitive than other procedures (Mulhall et al., 2005). Stool DNA test (Imperiale et al., 2014), fecal occult blood test, and fecal immunochemical test (Wilschut et al., 2011) are non-invasive procedures from the outside of the patient which are prone to false positive results. Overall procedures from the outside of the patient are less uncomfortable for the patient but are more prone to false positives with inaccurate location.

From the inside of the GI tract, clinicians perform flexible endoscope, sigmoidoscopy or wireless colon capsule, to identify CRC from the inside of the patients digestive system. Wireless colonoscopy (Iddan et al., 2000) involves patients swallowing a capsule with two cameras at both tips. An external data system records the video streamed by the capsule which surgeons analyze. This method has several downsides, first, it is an expensive method. Secondly, it requires more colon preparation compared to conventional colonoscopy (Spada et al., 2012). Control of the wireless colon capsule for inspecting a specific part of the colon is quite complex and time-consuming because it requires additional time for video analysis (Van Gossum et al., 2009). Furthermore, wireless colon capsules cannot remove polyps or perform biopsies (Martin et al., 2020). Colonoscopy and sigmoidoscopy are similar. If a flexible endoscope is inserted up to the cecum it is called flexible endoscopy, if it only reaches the sigmoid it is called sigmoidoscopy. Sigmoidoscopy uses a smaller and slimmer flexible device and only reaches the descending colon, thus not having the same reach. Flexible colonoscopy remains the primary screening procedure for CRC ahead of other currently available screening procedures due to the ability to identify and locate single or multiple polyps in a single intervention.

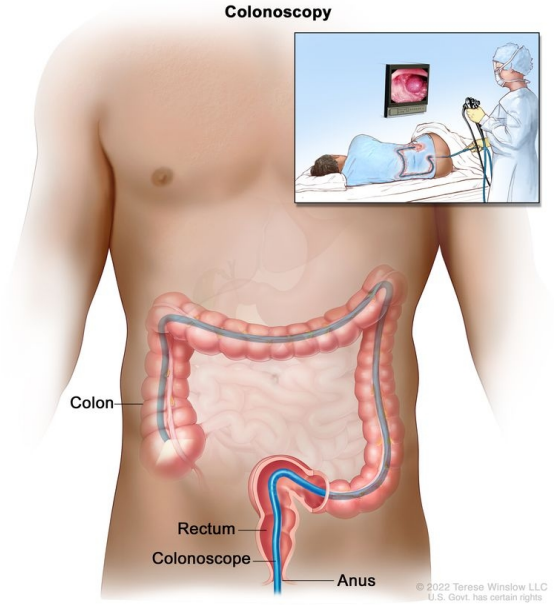


FIGURE 2.1: Colonoscopy overview. Image adapted from (NCI, 2022).

Flexible colonoscopy explores the GI tract from the rectum to the cecum covering the entire digestive system. Furthermore, during a colonoscopy samples could be taken for a laboratory analysis afterwards.

During flexible colonoscopy (depicted in fig. 2.1), flexible endoscopes navigate the digestive system from the anus to the cecum passing by the rectum, sigmoid colon, descending colon, and ascending colon (Bénard et al., 2018). The visual inspection is carried out by endoscopes which are introduced by surgeons and are manually operated. With these manual endoscopes, one surgeon performs insertion while the other performs navigation through tip control. The bendable tip allows the endoscope to steer in the desired direction. The endoscope is controlled to visualise the complete digestive system, which requires re-inserting the endoscope multiple times and involves coordinated cooperation between the two surgeons. The colonoscopy ends when the surgeons have completed the visual inspection. When a tissue is classified as abnormal, a tissue sample is required to perform a biopsy. When the sample is flagged as malignant, it is removed by the surgeon in a separate procedure, effectively stopping colorectal cancer from its early stages and stopping further organ spread, especially if the removal is performed in the early stage.

2.1.2 Removal of colorectal cancer

After the polyp(s) have been detected, scanned, and assessed, appropriate procedures for removing them are performed. Multiple procedures exist for removing polyps, from invasive and risky procedures such as open colectomy (Winslow et al., 2002) to less invasive procedures such as laparoscopic-assisted colectomy (Bonjer et al., 2007). Procedures such as open or laparoscopic-assisted colectomy rely on removing sections of the digestive system that reduce CRC completely if performed correctly. Removing sections of the digestive system can degrade the quality of life of the patient. Nowadays surgeons opt for less invasive procedures such as endoscopic mucosal resection (EMR) (Soetikno et al., 2003) using flexible endoscopes.

During EMR (depicted in fig. 2.2) the tissue is injected with a saline solution that lifts it, easing interactions with the surgical tools. The lifted tissue is pulled by a forceps (in some cases the endoscope has a cap that helps visualise the target[1]) and a snare is put around the polyp to remove it. EMR is complex as the user has only direct control over how much the tool is displaced over the channel of the endoscope, leaving the rest of the control indirectly attached to the endoscope body. Such that if the tool is required to move, the surgeons have to control the endoscope body and the tool simultaneously. The control of the endoscope body and the tools can be performed by the surgeons or the surgeon and an assistant. If an assistant helps in the procedure, the complexity of the tasks increases, as the surgeons have to efficiently communicate the actions they are performing and those required for the assistant. Even highly trained surgeons and assistants can cause bleeding since the task requires specialised equipment and trained personnel making EMR not widely currently available for patients (Kim et al., 2013). To correctly perform EMR surgeons require many training hours to avoid incomplete resection (Oka et al., 2006) and residual local disease (Bhurwal et al., 2016).

A procedure that has been gaining attention in the past few years due to the availability to perform precise dissection of polyps is Endoscopic Submucosal Dissection (ESD) (as depicted in fig. 2.3, showing the procedure after the tissue has been marked). Similarly to EMR, the surgeon has to control the body which has the arm mounted at the tip of the body to perform the ESD. The first step of ESD is marking the tissue with an

[1] the cap at the tip of the endoscope allows to push the tissue away from the camera placed at the tip of the endoscope

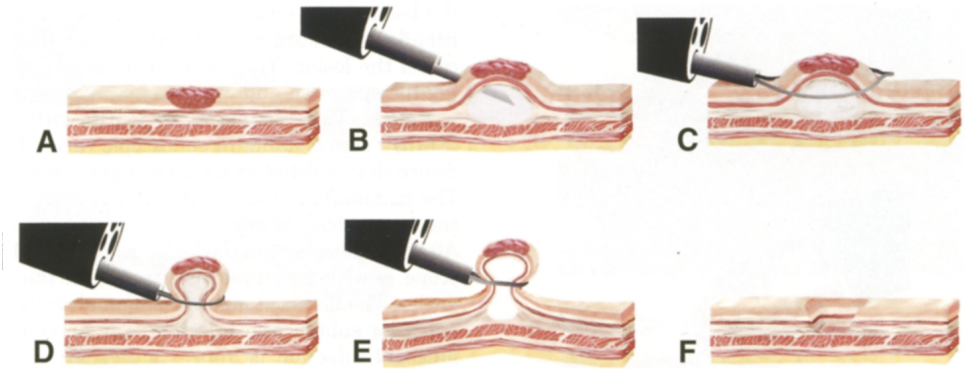


FIGURE 2.2: Overview of endoscopic mucosal resection. Image adapted from Soetikno et al. (2003).

electric knife, the user controls the body and the electric knife to effectively mark polyps of different sizes (Hoteya et al., 2009). Then, the surgeon inserts saline solution to lift the tissue and ease dissection. The flexible endoscope is equipped with a surgical tool –an electric hook or knife– that allows the cutting of the desired tissue until completely removed. Afterwards, the surgeon uses forceps to grab the tissue and retract the complete flexible endoscope. ESD can be more precise than EMR (Cao et al., 2009) because an arm is marking and cutting the tissue instead of a snare. Nevertheless, similarly to EMR, effective communication is necessary between the surgeon and assistant, as well as a high amount of training hours to correctly perform ESD.

Performing procedures to identify and remove polyps from the GI tract, EMR or ESD, is mentally demanding and requires high amounts of training hours. Due to the demanding environment and the requirements of the surgical procedure, dedicated manual endoscopic platforms have been developed, showing clinical improvements (Diana et al., 2013). Multi-user cooperation is, however, a big hurdle for optimizing patient care and efficiency. Single-user FSIR platforms could reduce the crowded operating room and allow surgeons to control the flexible and slender architectures of FSIR to detect and remove CRC from the digestive system. The slender and compliant architecture of FSIR allows them to navigate through the GI tract and act on the tissue since they carry flexible surgical arms at their tip. An overview of the most representative platforms that were developed to aid surgeons in performing such procedures is followed.

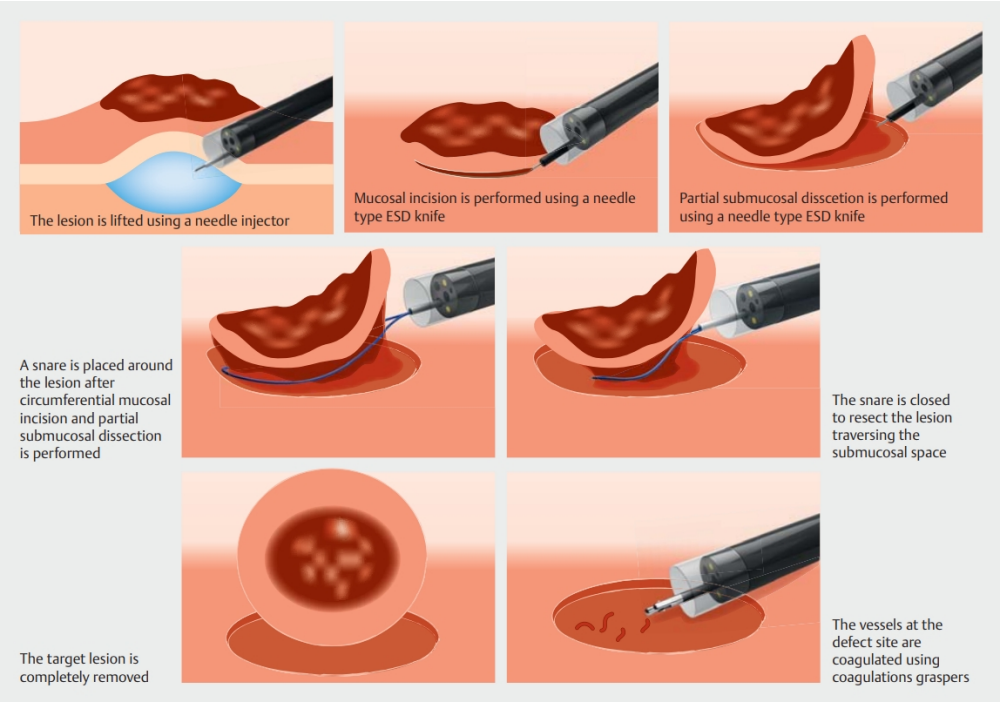


FIGURE 2.3: Overview of endoscopic mucosal resection. Image adapted from McCarty et al. (2020).

2.2 Flexible steerable intraluminal platforms

Multiple FSIR platforms have been developed worldwide to assist surgeons in correctly and safely performing surgical procedures in the patient's body. Flexible steerable intraluminal platforms were initially developed as endoscopes with robotic capabilities, mainly by fitting motors on the handle of regular endoscopes (Spaun et al., 2009; Fuchs and Breithaupt, 2012; Rivera-Serrano et al., 2012), (see fig. 2.4). In practice robotized endoscopes and flexible steerable intraluminal platforms are analogous as they are both composed of a main body with arms at their tip. Generally, the body has 4 DoF and the arms have 3 DoF. Additional actuation of the arms allows the activation of surgical tools, e.g. electric knives, graspers, and snares. The body can be equipped with tools in the central free channels of the body (1 or 2 channels at most) which are not robotized, such that manual control is performed by surgeons. Although the structure of FSIR is generally the same, differences in dimensions and features are present.

Most flexible steerable intraluminal platforms are built with slender flexible bodies, although some exceptions place a flexible body at the end effector of industrial robots (Berthet-Rayne et al., 2018). The bodies of FSIR are cable-actuated and range between

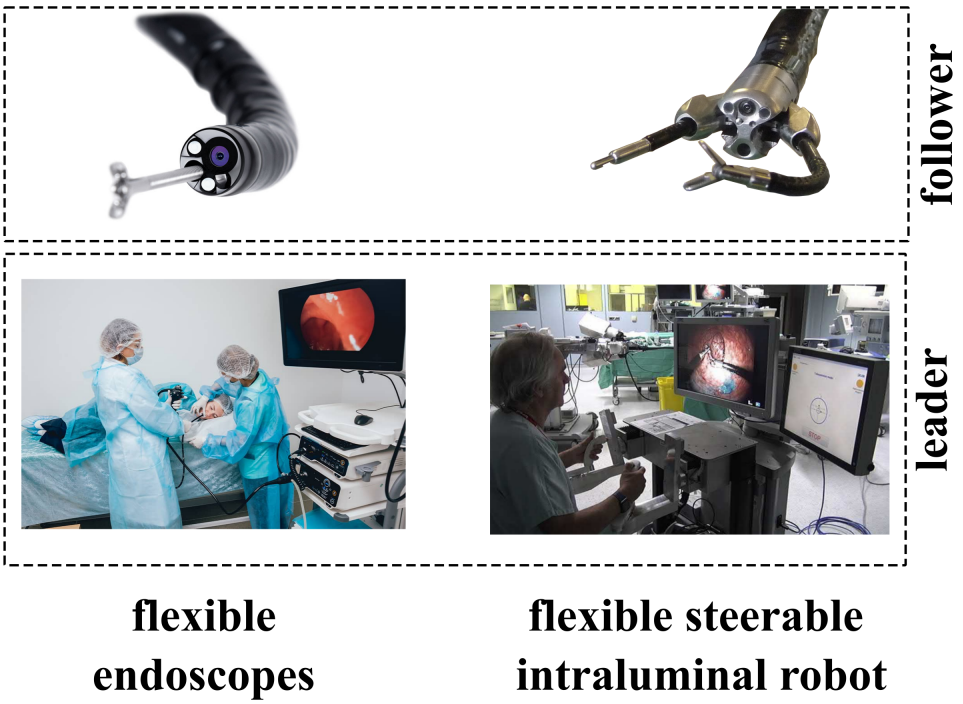


FIGURE 2.4: Flexible endoscopes and flexible steerable intraluminal robots. Leader and follower sides.

7-18 mm in diameter with lengths between 900 and 1100mm, allowing them to be introduced in the narrow colonic walls and cover reach the cecum (Fuchs, 2002; Spaun et al., 2009; Yeung and Gourlay, 2012; Nageotte et al., 2020). Some platforms are only developed for reaching the colon, having less than 500mm of length (Rivera-Serrano et al., 2012). Most FSIR amounts to a total of 10 DoF which are telemanipulated by surgeons using commercially available user interface platforms such as the Omega 7 (Force Dimension, Lausanne, Switzerland) or the Touch (3D systems, Rock Hill, USA). Due to the lack of an intuitive relationship between the leader and follower, proprietary interfaces have been developed (Mitsuishi et al., 2013; Wisanuvej et al., 2017; Ahn et al., 2021). The proprietary interfaces aim to handle an intuitive relationship between the natural movements of the surgeon's arms and the movements of arms and body in the follower side of FSIR. This relationship aims to facilitate arm movements of the arms as they carry most of the surgical tasks. Since movements of the body movements are secondary, their control is enabled by joysticks (Mitsuishi et al., 2013) or by movements of the support of the surgeon's arms (Wisanuvej et al., 2017; Ahn et al., 2021). Overall proprietary interfaces allow surgeons to have complete control of the 10 DoF of most FSIR platforms in a direct way. Although some dual-user FSR platforms exist (Eickhoff

et al., 2006; Phee et al., 2008), they are being replaced by single-user control platforms in order to avoid relying on multi-user coordination (Abbott et al., 2007; Patel et al., 2015; Nageotte et al., 2020).

A large number of FSIR platforms have been developed. One can cite the ViaCath system (Abbott et al., 2007), the MASTER (Yeung and Gourlay, 2012), the k-Flex (Kim et al., 2019) or the STRAS robot (Nageotte et al., 2020). New platforms based on 3D-printed prototypes have also recently been developed (Trauzettel et al., 2022; Mahapatra et al., 2020; Grassmann et al., 2024) with the aim to offer cheap and fast prototypes for specific needs. Some of these FSIR platforms have been tested in preclinical trials (Fuchs and Breithaupt, 2012; Légner et al., 2017; Atallah et al., 2018; Yang et al., 2022) with often positive results on the clinical side but work to be done to facilitate their telemanipulation as to decrease the learning curve.

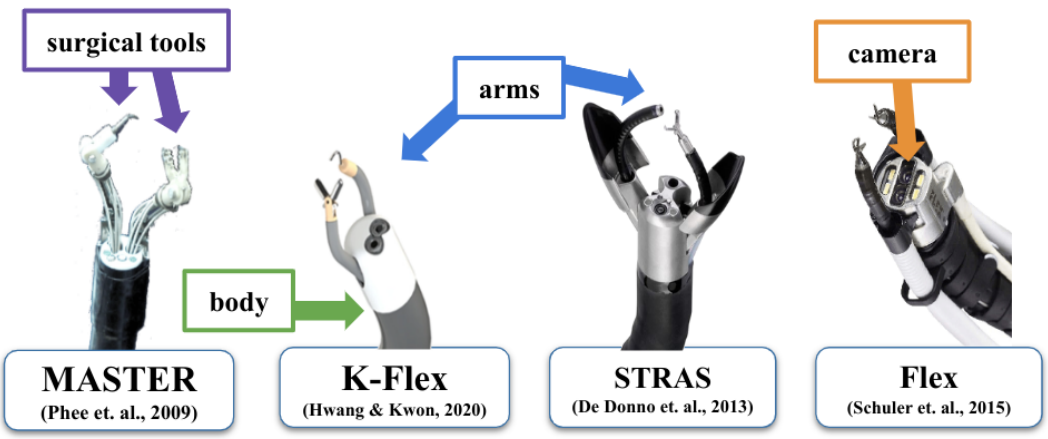


FIGURE 2.5: Examples of flexible steerable intraluminal robots. From left to right MASTER (Phee et al., 2008), K-Flex (Hwang and Kwon, 2020b), STRAS (Nageotte et al., 2020) and FLEX (Schuler et al., 2015). Images adapted from authors.

Interestingly, all the above-mentioned platforms share noticeable similarities. Generally, FSIR platforms follow a leader-follower architecture. The follower side is similar in FSIR platforms (see fig. 2.5): a flexible robotized main body (with at least 4 DoF) carrying a camera at its tip, and surgical tools (also potentially robotized) protruding from the tip of the body (with at least 3 DoF each). This common architecture results in specific challenges, which should be overcome in order to offer a more intuitive and aware –by assistance or autonomy– telemanipulation.

2.3 Current challenges of FSIR platforms

Correct manual telemanipulation of FSIR platforms is challenging, especially during technically difficult and mentally demanding surgical procedures. The surgeon must take into account the correct performance of the surgical procedure and patient safety while moving the leader interface to accuratly control the FSIR platform. This is rendered difficult by two main challenges, which are common for all FSIR described in section 2.2 and arises from their general architecture. First, ***$\mathscr{C}_1$*: coupled architecture** which is inherited from robotized endoscopes enabling robotic telemanipulation in no-scar surgery. This first challenge is related to the fact that the arms are carried by the body, and therefore movements of the body impact the arms, making coordinated movements especially challenging. Second, ***$\mathscr{C}_2$*: high amount of degrees of freedom** which is directly stemming from the fact that FSIR is built to handle complex procedures with bimanual operation. This enables technically difficult surgical procedures that allow the surgical arms to cooperate to grasp, pull and push tissue inside the patient, at the price of increased complexity for the surgeon. The remainder of this section details the two challenges.

***$\mathscr{C}_1$*: coupled architecture** Flexible steerable intraluminal robots are developed based on the design of flexible endoscopes, inheriting their benefits and drawbacks. Surgeons and assistants must cooperate to control the coupled architecture of flexible endoscopes and perform complex surgical tasks. With FSIR, the complete platform –arms and body– is controlled by one user –a surgeon– instead of two users –a surgeon and an assistant, as in flexible endoscopy. Single-user telemanipulation was developed to decrease the number of people in the OR, reducing the workload of a single surgeon. Furthermore, single-user telemanipulation avoids errors due to poor communication (Frankel et al., 2007; Ravindran et al., 2021) between the surgeon and the assistant (Mascagni et al., 2019). During FSIR single-user control, the user must consider the body for performing arm movements and vice-versa. Since the body has a camera at its tip, movements of the body (camera) will impact the movement of the arms, yet arms can be independently controlled. If the surgeon is required to extend the arms' reach, the body is required to move and the arm should compensate for the body's movement. Similarly, when a user needs to adjust the camera's position, it requires moving the FSIR body, which in turn affects the position of their arms. The coupled architecture demands a high

level of awareness of the relationship between the arms and the body. To facilitate the telemanipulation of the coupled architecture, researchers have developed intuitive master consoles that map the leader to the follower. The intuitive mapping facilitates user control. Yet due to the dependency on the movements between the arms and body, and the intricacy of their control challenges to master their control are still present. To deal with the coupled architecture, users perform unnecessary movements originated by the over- or under-compensation of movements of the arms or body. The unnecessary movements (Cao et al., 1996) can increase the operating time and strain surgeons posing safety risks (Zheng et al., 2012).

$\mathscr{C}_2$: high amount of degrees of freedom The coupled architecture of FSIR relies on users controlling the different subsystems simultaneously. The complete set of subsystems of FSIR constitutes a high number of degrees of freedom that the user must consider. Pre-clinical (Légner et al., 2017; Kim et al., 2019) and benchtop experiments with FSIR platforms identified that users struggle to perform movements with high amounts of degrees of freedom. In several FSIR platforms, surgeons have to command up to 10 DoFs. Leader interfaces must consider handles, joysticks and pedals which are 1-to-1 mapped to the follower, maintaining the dimensionality. The common mapping involves mapping the movements of the surgical arms with movements of the arms of the surgeon, such that the leader console allows intuitiveness between the leader and follower sides. The arms are the main actors, therefore, the leader console is focused on intuitive arm movement. The body is of secondary importance, the body movement is governed by joysticks or pedals. Even if the total number of DoF in FSIR platforms is present in the leader console, their control requires surgeons to move their arms and hands –to govern the handles in the leader side which command the surgical arms– and their fingers or feet –governing the pedals on the floor or the joysticks at the tip of the handles to move the body. Thus, users controlling the 10 DoF simultaneously is inconvenient, and it is even harder to perform under mentally demanding procedures, such as during the dissection stage of ESD. To deal with the high amounts of DoF, users avoid simultaneous control of the arms and body of FSIR platforms, effectively reducing the number of DoF to be controlled in a given moment and making the task more manageable yet the number of movements required increases which is undesired due to the patient's safety. Other less demanding surgical steps in ESD, such as tissue marking, require the user to control both the body and one arm (7 DoF), achieved by

first controlling the body (4 DoF) followed by the arm (3 DoF). Overall, the complexity of the surgical step correlates with the number of degrees of freedom (DoF) and the mental demand placed upon the surgeon.

The high amounts of degrees of freedom of the arms allow users to perform complex surgical gestures that require a high level of interaction between the tool and the tissue. The high level of interaction between tools and tissue means the user can manipulate tissue in multiple orientations and directions facilitating grabbing, pushing and/or pulling individually and in cooperation with both arms. The freedom to manipulate the tissue comes from the fact that the arms of FSIR have 3 or more DoFs. Most arms of FSIR can be inserted/retracted, actuated, rotated and bent in one or two directions from the tip of the body –some arms can be overly actuated due to the multiple bending segments. In comparison, the rigid arm of flexible endoscopes can only be translated and rotated. The increased workspace of the arms in FSIR allows to facilitate ESD in small lesions $<$ 10mm (Légner et al., 2017; de Moura et al., 2019; Chiu et al., 2021) and large lesions $>$ 15mm without degradation of precision with a lower recurrence rate, yet correct telemanipulation of the arms is technically difficult for surgeons. For small lesions, performing ESD with manual flexible endoscopy results in procedures prone to recurrence or bleeding (Thompson et al., 2009; Kume, 2009) due to the lack of precision in their movements tied with the body movement and the lower DoF. The lower number of DoF in the surgical arms used in flexible endoscopes requires handling the coupled movements between the rigid tool and the flexible body. To replicate the range of motion of the FSIR arms, rigid tools of flexible endoscopes require users to consider the coupled architecture (fig. 2.4). The user extends the arm workspace by the movements of the flexible endoscope. Performing arm-driven body control in flexible endoscopes is mentally demanding (Légner et al., 2017; de Moura et al., 2019; Chiu et al., 2021) following recurrence or bleeding in the patient's body. Using FSIR complex surgical procedures are reduced in complexity.

Similarly to flexible endoscopy (Ruiter et al., 2013; Gaab, 2013), surgeons require a large number of training hours (Légner et al., 2017; Chiu et al., 2019) to master the control of the FSIR platforms. Even with the development of intuitive leader consoles (Wisanuvej et al., 2017; De Mathelin et al., 2020; Ahn et al., 2021) the users struggle to master the control of FSIR platforms and perform the surgical procedures correctly (Mascagni et al., 2019). To deal with these drawbacks users avoid simultaneous control

of the DoF of the arms and body of FSIR, which can cause interruptions that prolong surgical procedures (Kenngott et al., 2012) or even lead to technical failures (Légner et al., 2017; Meng et al., 2022) due to the extended duration. These problems could be resolved if the user is assisted in handling the two main challenges of FSIR while safely and correctly performing the technically difficult surgical procedures.

The two main identified challenges of FSIR, $\mathscr{C}_1$**: coupled architecture** and $\mathscr{C}_2$**: high amount of degrees of freedom**, must be correctly handled in a robotic assistance setup in order to assist or autonomously perform tasks that aid surgeons in complex surgical procedures. The following presents an analysis of the state of the art of robotic assistance in robotic surgery.

2.4 Robotic assistance during complex surgical procedures inside the body

Robotic assistance in surgery has evolved from the first ideas which envisioned robotic telemanipulated laparoscopy (Alexander III, 1972). Other early works proposed robotic assistance for passing tools to surgeons or holding retractors for assisting brain biopsy (Kwoh et al., 1988; Abdelaal et al., 2020). In current times, the idea of robotic laparoscopy is a reality with multiple platforms available such as the da Vinci from Intuitive Surgical, USA or even single access robots like the Virtuoso by Virtuoso Surgical, USA. Currently, minimally invasive surgery poses a less intrusive procedure in comparison to open body surgery by using small incisions or the natural orifices of patients. This section describes an overview of robotic assistance in minimally invasive surgery .

2.4.1 Robotic assistance in minimally invasive surgery

Assistance via robotic actuation in minimally invasive surgery has been in development since the 1980s. Early works involved simple tasks such as passing tools or holding endoscopes in a static position. Robotic assistance in MIS (Haidegger, 2019; Attanasio et al., 2021; Battaglia et al., 2021) has been classified in 6 main levels, as depicted in fig. 2.6. At level 0, the robotic system has no autonomy, it is all dependent on the

operator, this is commonly referred to as robotic minimally invasive surgery, e.g. laparoscopy by the DaVinci platform or robotic surgery, e.g. knee robotic surgery by industrial robots. At level 1, the robotic assistance level, the operator maintains continuous control of the system while the robot provides certain assistance; the user can be guided to perform a particular task. Providing robotic assistance means the user is guided by the robotic approach or using virtual fixtures (Rosenberg, 1993) during telemanipulation. At level 2 of task autonomy, the system is capable of accomplishing specific surgical tasks autonomously based on specifications given by the operator. During the autonomously executed task, the control switches from the operator to the machine. At level 3, referred to as conditional autonomy, the autonomous task is based on perceptual capabilities to understand the surgical scenario. Tasks are planned and executed, allowing continuous monitoring. The control can switch from user to machine when required. Commonly, levels 1, 2, and 3 are referred to as robotic-assisted surgery, in the case of MIS, robotic-assisted minimally invasive surgery. Level 4 referred to as high autonomy, can interpret pre-operative and intra-operative information to plan the sequence of tasks with the capability of replacing on-the-go if necessary. The last level, level 5 is referred to as full autonomy, machines can perform surgery on their own with no human input. Levels 4 and 5 are commonly referred to as autonomous surgery.

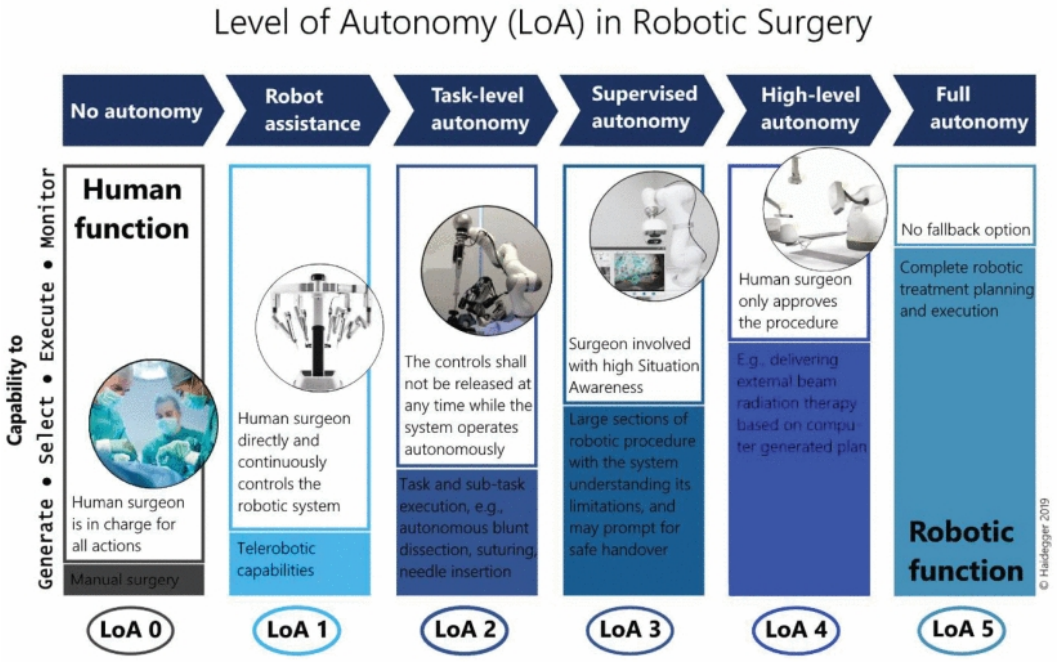


FIGURE 2.6: Commonly proposed levels of autonomy for minimally invasive surgery, as proposed by Haidegger (2019). Image adapted from Haidegger (2019).

The current trends in robotic assistance during MIS for procedures inside the GI tract focus on levels 2 and 3. Most of the focus of research has relied on having the surgeon in the loop while robots assist or conduct –some– autonomous tasks. In this thesis, we will follow the same approach. Indeed, the challenges identified in section 2.3 are closely related to telemanipulation, and it is natural to propose modes of robotic assistance that enable more efficient telemanipulation while keeping the surgeon in the loop. Such approaches are scarce in the literature for FSIR platforms, due to their recent development. Similar problems were however studied in the context of robotic assistance for laparoscopic surgery. We will detail the main approaches related to endoscope control of a single endoscope body in laparoscopic robotic platforms in section 2.4.2. Likewise, approaches that control the arms in the same platforms are presented in section 2.4.3.

2.4.2 Autonomous endoscope control of laparoscopes

In robotic laparoscopy, an endoscope can be controlled in multiple ways. Endoscope control is commonly categorised as reactive, proactive or a combination of the two. Reactive approaches rely on monitoring the task, the surgeon, the environment, and the user commands to react after a triggering event. The reactive approaches are applied after the endoscope has already been moved, rather than being proactive in preventing the need for movement. A comprehensive review of the approaches can be found in (Weede et al., 2013; Pandya et al., 2014; Bihlmaier, 2016; Wagner et al., 2021) covering multiple categorization methodologies expanding on the scope of endoscope camera control for laparoscopes. Furthermore, Ellis et al. (2016) performs an in-depth analysis of how endoscopic cameras are controlled by surgeons and details what steps to consider when conceiving an assisted or autonomous control approach for endoscopes. Other works (Avellino et al., 2020; Ellis et al., 2016) aim to provide a link between semantic understanding of the task and the control approach. In this thesis, approaches for assistance or autonomy in FSIR will be covered.

The first approach that is going to be addressed is the actuation of the endoscope body by user commands. The user can consider its environment, consisting of the surgical site, the surgical tool, the endoscope and the leader console to request an assisted or autonomous movement of the endoscope. Approaches have used user feet (Wang et al., 1998; Huang et al., 2021) or head Yasunaga et al. (2003); Stolzenburg et al. (2011) voice

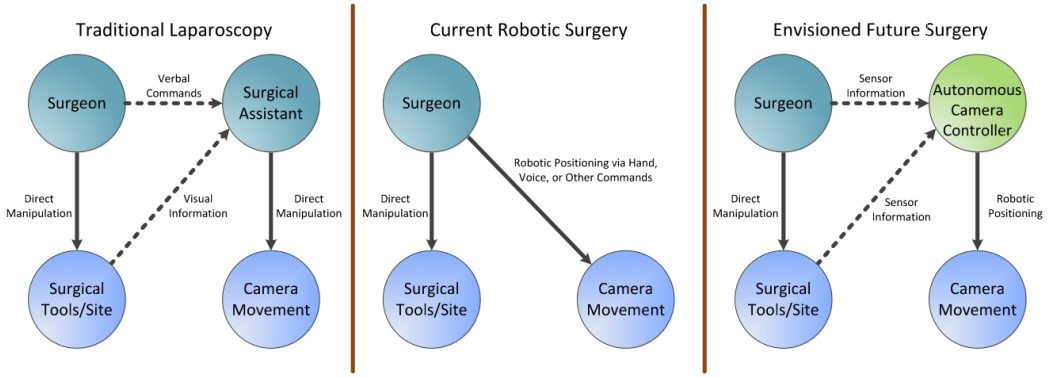


FIGURE 2.7: Overview of the common approaches for autonomous laparoscopy. Image adapted from Pandya et al. (2014).

(Zinchenko et al., 2016), gaze (Latif et al., 2008; Zhu et al., 2011; Fujii et al., 2018) or their whole body (Zuo et al., 2022) to command movements of the body. These approaches directly map commands from the user to the movement of the body in one or two directions. Such approaches require the user to specifically express the desired direction in which they intend to move, and typically function with rate-control, meaning the amount of movement on the body is proportional to the the time, force or position the user sets.

The second approach for controlling the body uses the information from the environment to automatically regulate the position of the endoscope body. Since the camera is positioned at the tip of the endoscope, this type of approach also controls the position of the camera with respect to the environment, e.g. position of the arms relative to an objective, or the position of the camera relative to the environment. Most of these approaches use information about the arms, the surgical task and predefined tasks. This second category of approaches acts after predefined conditions are met, e.g. when the position of the tools is outside of the area of interest in the image of the endoscopic camera (Agustinos et al., 2014; Gruijthuijsen et al., 2022) the body moves. Rivas-Blanco et al. (2014) used the depth of the instruments to control the endoscopic camera. Kinematic tracking of the arms and the position of the camera with respect to the environment (Mudunuri, 2010; Bodenstedt et al., 2012) has also been developed. Other approaches, (Lee et al., 1994) consider surgical landmarks and information about the surgical site embedded in the image to control the body.

Approaches based on direct user commands and those based on the environment are

interesting in formulating autonomous or assisted movements for the endoscopic camera. However, these approaches cannot be directly translated for the control of FSIR platforms. This is because endoscopes in laparoscopic settings are inserted in separate entry points into the body rather than arms. Both are therefore kinematically independent. In other terms, robotic assistance approaches for endoscopic camera control do not tackle $\mathscr{C}_1$**: coupled architecture**, which facilitates their control.

2.4.3 Autonomous control of arms of laparoscopes

Approaches for assistance and autonomous control of the arm of laparoscopes have been developed. In prior works (De Rossi et al., 2021; Attanasio et al., 2020; Pasini et al., 2023), researchers used approaches based on deep learning to detect the gesture and assist the user in correctly performing the detected surgical gesture. The user gesture is detected at the early stages of the arms' movement. Then, the control mode eases carrying out the gesture by helping the user follow a predefined set of movements that will occur based on the detected gesture. The automatic or assisted actions can be overwritten by the surgeon at any time, leaving the user the main actor during this type of approach. In such approaches, the task is required to be known and large amounts of information regarding the task (video, kinematics and sensory information) are required to be conceived.

Another way to train these models with few amount of information is learning-by-demonstration (Schwaner et al., 2021; Pore et al., 2021). In these types of approaches, the user teaches the correct way in which the task should be performed, to allow replication under similar conditions by the agent, which are hard to generalize. (Varier et al., 2020) proposed a similar approach using reinforcement learning, thereby demonstrating robustness to a lack of training scenarios and outlier behaviour. In these types of approaches, the user is aided to control one or both arms into performing tasks that are simple and repetitive for the user. Many approaches still rely on the user for performing the most complex surgical tasks, as fully autonomous execution of such tasks remains unavailable.

Other types of approaches were developed to share control of the arms at specific times. The approach proposed by Bodenstedt et al. (2012) helps surgeons –by sharing control of the arms of laparoscopes– to perform a first-order approximation of a surgical task.

The user performed a path-following task aided by the proposed approach based on Gaussian mixture regressions and 3D reconstructions. Using a 3D shape, researchers validated their approach by facilitating following the desired path.

Other approaches (Daneshgar Rahbar et al., 2021) use visual tracking to command the arms and prevent bleeding. By using an approach based on local entropy, Daneshgar Rahbar et al. avoided abrupt movements of the arms preventing bleeding. They prevented bleeding by anticipating by 0.662s when the arms will eventually lead to bleeding.

Both the arms of FSIR and the arms of laparoscopes enable surgical procedures inside the patient. In FSIR, the arms aim to be flexible and slender with multiple degrees of freedom for controlling positioning and orientation. In laparoscopes, the arms can be articulated, although not as compliant as in FSIR; having a couple of links at most. Furthermore, the key difference is that the arms of FSIR are mounted in the body of FSIR, creating a coupled kinematic architecture. Due to this, approaches developed for laparoscopes are not straightforwardly translated into the arms of FSIR. Challenges in assisting and automatic control of the arms, the body and both at the same time remain.

2.5 Remaining challenges in FSIR and thesis contributions

FSIR platforms pose two main challenges in manual telemanipulation, $\mathscr{C}_1$**: coupled architecture** and $\mathscr{C}_2$**: high amount of degrees of freedom**. To assist and automate the telemanipulation of FSIR their challenges, safety constraints and complex surgical tasks must be equally considered, e.g. dealing with the challenges of FSIR to correctly performing a surgical step with the consideration of the safety constraints. FSIR have not developed any assistance or autonomy approaches, therefore works in laparoscopy and other surgical robots are reviewed. Laparoscopes have been widely developed and multiple approaches for assisting and automating particular tasks have been developed. Due to the similarity to FSIR it is a logical comparison point. Assisted or autonomous control has been developed for laparoscopes with individual control of the body (section 2.4.2) or the arms (section 2.4.3), considering multiple surgical tasks, the environment, safety constraints or a mixture of them. Such approaches offer interesting solutions for assisting the user in handling high-DoF systems, effectively addressing

a challenge similar to $\mathscr{C}_2$**: high amount of degrees of freedom**. However, the fundamental difference between laparoscopic settings and FSIR is the presence of $\mathscr{C}_1$**: coupled architecture** in the latter. Therefore, the approaches developed for laparoscopes cannot be directly translated to FSIR. This task is the main focus of this thesis, in which approaches for assisting FSIR telemanipulation under specific surgical tasks are developed.

This thesis proposes a control framework to handle $\mathscr{C}_1$**: coupled architecture** and the $\mathscr{C}_2$**: high amount of degrees of freedom** to assist users in correctly and safely performing technically difficult and mentally demanding surgical procedures. The control framework developed in this thesis is between levels 2 and 3 of autonomy in robotic surgery, with the surgeon either fully in the loop with robotic assistance (level 2), possibly with some sub-tasks of the operation carried out automatically under direct surgeon supervision (level 3). The proposed control framework is developed and tested under two specific surgical scenarios, the dissection stage of ESD and the colonoscopy and biopsy case. In the remainder of this thesis, developments are carried out on the STRAS system, which is described in details in chapter 3, while chapters 4 and 5 constitute the main contributions of this work. In chapter 4 the problem of telemanipulation assistance in ESD is considered. A suitable control framework is selected after careful analysis of the state of the art in related areas. Clinical objectives are then formulated as optimization objectives and constraints in an optimization-based control formulation. It is then validated with a user study. A similar path is followed in chapter 5, where the problem of robotic assistance for biopsy in colonoscopy is considered. The problem is posed as a visual servoing problem under multiple objectives and constraints, and formulated in the same framework as in chapter 4. Validation is performed in a realistic *ex vivo* scenario, using the follower side of the FSIR platform.

Chapter 3

STRAS system

The STRAS system is selected for the validation of the control framework as it faces similar challenges of most of the FSIR platforms (chapter 2 - section 2.3). A brief overview of the story of the development of the STRAS system is presented in section 3.1. The STRAS system is described in section 3.2, which is followed by the modeling required for the control framework (section 3.3). Section 3.4 describes how the STRAS system is telemanipulated by surgeons and operators at the current state, which has been used in pre-clinical trials (Légner et al., 2017). This chapter finalizes by identifying the problems encountered during the pre-clinical trials and during benchtop experiments. A final statement regarding the platforms problems in the STRAS system is described in section 2.3.

3.1 Development of the STRAS robotic system

Back in 2009 (Bardou et al., 2009) the ANUBIS (see fig. 3.1) project was conceived by Karl Storz and IRCAD for Natural Orifices Transluminal Endo-scopic Surgery (NOTES). This early version of what would become the STRAS system was fully manual and required multiple users to control the arms and the endoscope body. In 2013 (De Donno et al., 2013a) the Anubiscope was motorized and the STRAS robotic system was conceived. The 2013 version allowed telemanipulation of the arms and body by human-interface devices, such as the Omega 7 (Force Dimension, Lausanne, Switzerland) and the Touch (3D systems, Rock Hill, USA). Such controls allowed users to have control of

all DoF of the arms and body (De Donno et al., 2013a). By 2017 (Zorn et al., 2017), proprietary interfaces were developed. These new interfaces allowed a more natural relationship between leader and follower. The new interfaces included pedals and a pair of handles that intuitively mapped the leader and the follower, allowing users to have full control of all DoF. The latest version of the STRAS system is referred to as Endoluminal Assistant for Surgical Endoscopy (EASE) described in fig. 3.1, which improves the leader console with small changes in the way the human-interface devices are set, e.g. improved joysticks and triggers (Nageotte et al., 2020).

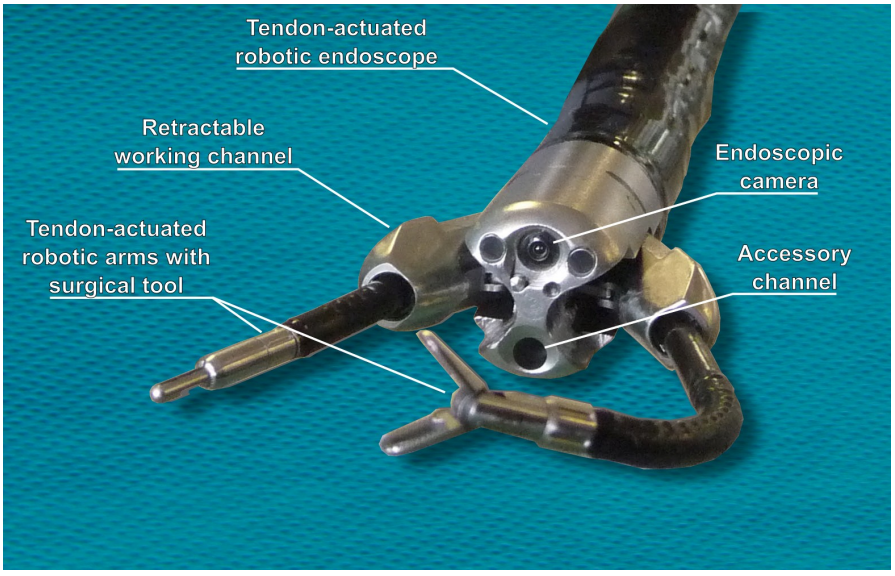


FIGURE 3.1: Anubiscope system developed by Karl Storz. Image adapted from Chikhaoui and Rosa (2022).

The STRAS system has been tested in a pre-clinical trial (Légner et al., 2017) performing multiple ESD procedures in an ex-vivo porcine model. Furthermore, extensive in-silico tests have been performed under multiple surgical training tasks, e.g. peg-transfer and pick-and-place tasks. Tests with novice and expert users have shown that experienced users perform the task in shorter execution times and with specific differences in handgrip forces during manipulations (Batmaz et al., 2017; de Mathelin et al., 2019).

The motorization of the ANUBISCOPE system resulted in the STRAS system, depicted in fig. 3.2. The body and arms are mounted into the cradle, which has the motors used for arm actuation and two DoFs of the body. A cart that carries the cradle that contains the motor controllers, power supply, and EtherCAT bus. The cradle provides another two DoFs of the body and the instruments. The instruments are controlled by motorized modules that actuate the translation rotation and bending of the arms. The cradle and

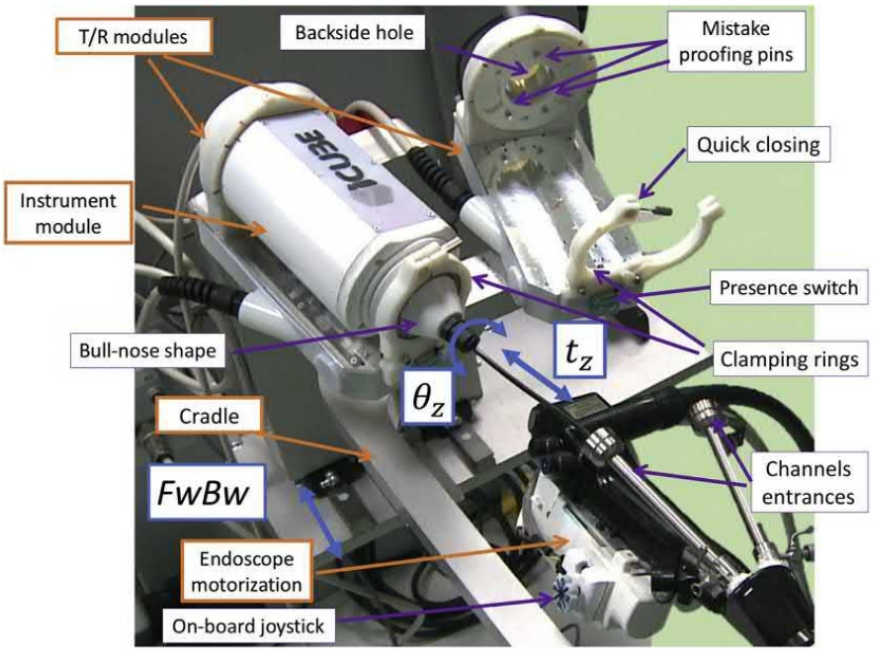


FIGURE 3.2: Motorized modules for actuation of the arms. Image adapted from Nageotte et al. (2020).

cart allow actuation of the body, arms, and arm actuators. The system is dismounted from the cradle and cart in other to be manually inserted into the surgical site.

3.2 STRAS system overview

The STRAS system is controlled by a leader-follower teleoperation architecture (as depicted in fig. 3.3), where the user controls the leader that commands the follower. The follower side follows the leader's references. The leader console is designed to intuitively imitate the arm's natural gestures and easy user control. In this regard, the leader and follower sides are coupled such that movement in the leader constitutes a movement in the follower –each DoF of the follower is paired with a DoF of the leader. The movements of the leader are read by encoders and button signals that are connected to a low-level controller via EtherCAT. The low-level controller is running in a real-time task 1kHz Linux Xenomai PC that governs the motors and actuators. The low-level controller for the follower can handle joint position and velocity control according to the commands of the user through the position and the actuation of the human-interface devices by the handles, joysticks, pedals, and buttons. The STRAS system follower can be completely controlled by the user using the leader side.

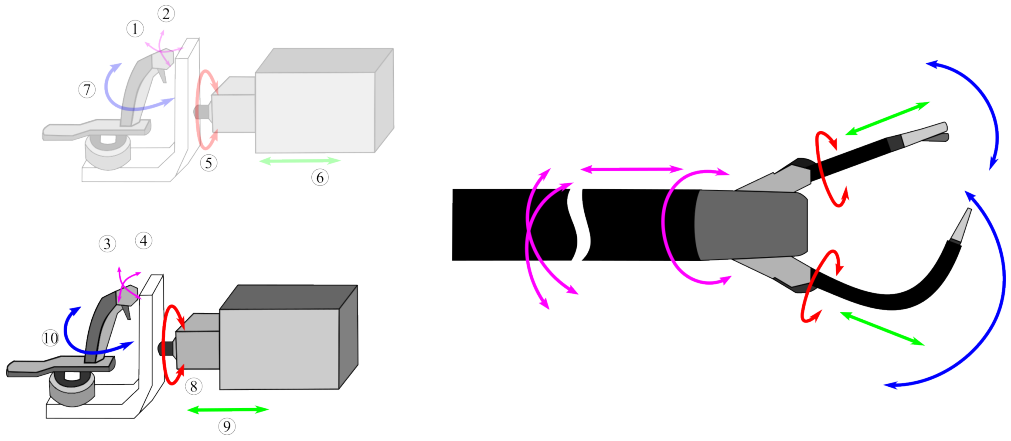


FIGURE 3.3: Overview of the STRAS system. The leader side is on the left side of the image, follower side is on the right.

The STRAS system follower side consists of a main endoscope body that can bend in two orthogonal directions, rotate and translate over its axis, correspondingly to ①-④ in fig. 3.3. An endoscopic camera is positioned at the tip of the body. The body has three working channels that can be populated by three surgical instruments. The common usage scenario has two surgical arms in the two channels. The centre channels also allow fluid management, by using air to increase colon diameter or by allowing injection of saline solution for lifting tissue. For right-handed users, the right arm is the primary arm and the left arm is the secondary arm [1]. The primary arm has an electrocautery tool at its tip while the non-dominant arm holds a grasping tool. In fig. 3.3 the top arm is the left arm, referred to as the non-dominant arm, while the bottom arm which is the right arm is referred to as the dominant arm. The arms can move translate, rotate, and bend, depicted by ⑤-⑦ for the non-dominant arm and ⑧-⑩ for the dominant arm allowing complex surgical gestures as depicted in fig. 3.4.

The STRAS system leader console comprises two handles with joysticks at their tips and a trigger on the back side, as seen in the leader side of fig. 3.3. The joysticks placed at the top of the handles move the body. The usual configuration is as follows. The joystick from the left handle actuates translation and rotation. The joystick from the right handle actuates vertical and horizontal bending. As in the leader side of fig. 3.3, the ①-④ are mapped to the ①-④ in the follower side. The arms are mapped to the handles (see fig. 3.3). The rotation of the handle (yaw-like motion) is mapped to the rotation of the tool, depicted in red arrows for the two arms. Translation of the handle in or outside

[1] for this thesis, the convention of right-handed users is considered, yet the techniques here developed could be mirrored to accommodate left-handed users

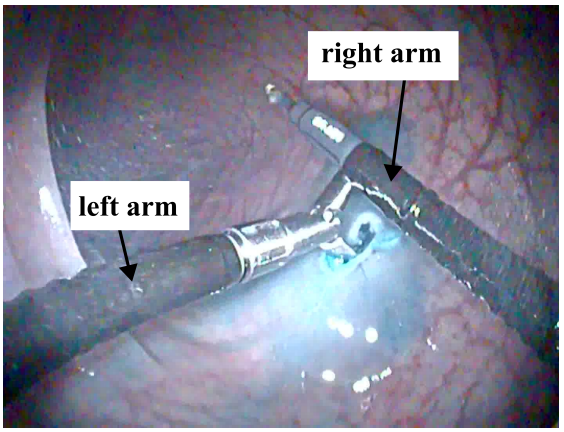


FIGURE 3.4: STRAS arms from the endoscopic camera. Right and left arms for the dominant and non-dominant arms correspondingly (Assuming right-handed users). Image adapted from Légner et al. (2017).

is mapped to insertion or retraction of the arm, by the green arrows. The bending of the handle as depicted by the blue arrow maps to bending of the tool. Similarly to the body, the arms movement is mapped between leader and follower by the ⑤-⑩ DoFs. The leader console allows to approximately match the position of the arms and the instruments (fig. 3.5) which improves intuitiveness and reduces the familiarisation time (De Mathelin et al., 2020).

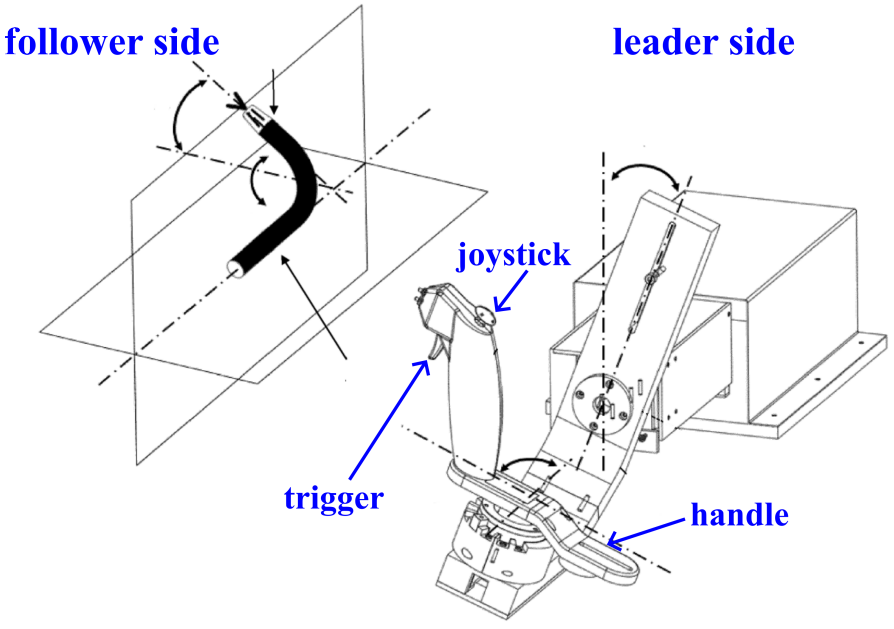


FIGURE 3.5: The handle of the leader side of the STRAS system (right) is made such that the position of the follower arm (left) approximately matches the position of the surgeon's arm. Handles are mounted in an L-shape bracket that translates over their base, with two rotational joints. They also comprise a joystick at the top of the handle, and a trigger at the back-top of the handle.

3.3 Modelling of FSIR platforms applied on the STRAS system

Preface of modelling

Notation During this thesis, frames attached to objects are depicted between curly brakes as $\{m\}$. Matrices are depicted in bold uppercase as $\mathbf{M}$. Vectors are represented in bold lowercase $\boldsymbol{m}$ while scalars use lowercase m.

Example The homogeneous transformation matrix (HTM) depicting the position and orientation of the robot end effector $\{end\}$ relative to the world frame $\{origin\}$ is defined as ${}_{origin}\boldsymbol{T}^{end}$. The homogeneous transformation matrix is expressed as:

$$ {}_{origin}\boldsymbol{T}^{end} = \left[\begin{array}{ccc:c} & {}_{origin}\boldsymbol{R}^{end} & & {}_{origin}\boldsymbol{p}^{end} \\ \hdashline 0 & 0 & 0 & 1 \end{array}\right]. \tag{3.1} $$

constituted by a rotation matrix ${}_{origin}\boldsymbol{R}^{end}$ of the tip of the arm to its base, and ${}_{origin}\boldsymbol{p}^{end}$ is the translational component describing its position.

3.3.1 Introduction

Previous works on the ANUBISCOPE (Ott et al., 2008; Bardou et al., 2010) and the STRAS system (De Donno et al., 2013b; Nageotte et al., 2020) were used to model the STRAS system. The STRAS system is actuated by motors that modify the cable lengths that control the bending of the tendon-actuaded arms, while motorised actuators allow to rotate and insert the arms (as represented in fig. 3.2). The motors define the actuator space ranges and map a theoretical representation of a model in the configuration space. The transformation ($f_{specific}$) between actuator space and configuration space is a matrix between cable lengths, displacements, and rotations according to prior works (De Donno et al., 2013a; Zorn et al., 2019). The following sections describe the configuration space of the constant curvature model (CCM) (described in section 3.3.2) used in the modelling o the arms of FSIR, and therefor on the STRAS system. The

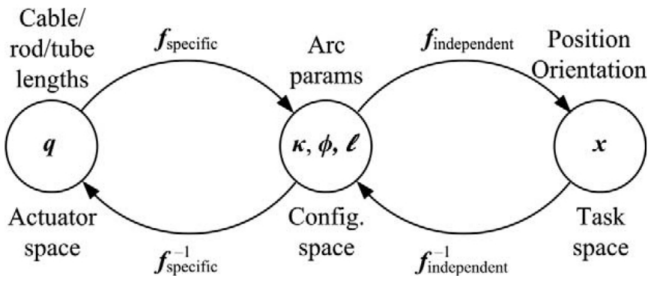


FIGURE 3.6: The three spaces and mappings of the constant curvature model. Image adapted from Webster III and Jones (2010).

specific transformation (as described in fig. 3.6) from the configuration space to the task space ($f_{independent}$) is presented in section 3.3.3, and the inverse relation ($f^{-1}_{independent}$) is described in section 3.3.4.

3.3.2 Constant curvature model

A flexible endoscope and flexible surgical tools of FSIR platforms are a type of hyper-redundant flexible manipulator that can be considered continuum robots. A continuum robot can be modelled as a constant curvature model using Webster's formulation (Webster III and Jones, 2010). The constant curvature model describes the kinematics by the configuration variables:

$$\boldsymbol{q}_{ccm} = [\phi, \kappa, s]^T , \tag{3.2}$$

with ϕ being the rotation about the axis of the continuum robot, κ is the curvature [2] and s is the arc length[3], as depicted in fig. 3.7. Using the configuration variables in eq. (3.2), the forward kinematic model that describes the position and orientation of the tip of a continuum robot is given by:

$$ {}_{base}\boldsymbol{T}^{tip} = \begin{bmatrix} \cos^2\phi(\cos\kappa s) & -\sin\phi & \cos\phi\sin\kappa s & \frac{\cos\phi(1-\cos\kappa s)}{\kappa} \\ \sin^2\phi(\cos\kappa s) & \cos\phi & \sin\phi\sin\kappa s & \frac{\sin\phi(1-\cos\kappa s)}{\kappa} \\ -\sin\kappa s & 0 & \cos\kappa s & \frac{\sin\kappa s}{\kappa} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.3}$$

[2] the curvature can be represented in bending angle by the relationship $\beta = \kappa s$

[3] end effector control is desired, s considers the total length of the bendable part, such that $s = \ell$. s could be modified to know the position of parts of the flexible segment.

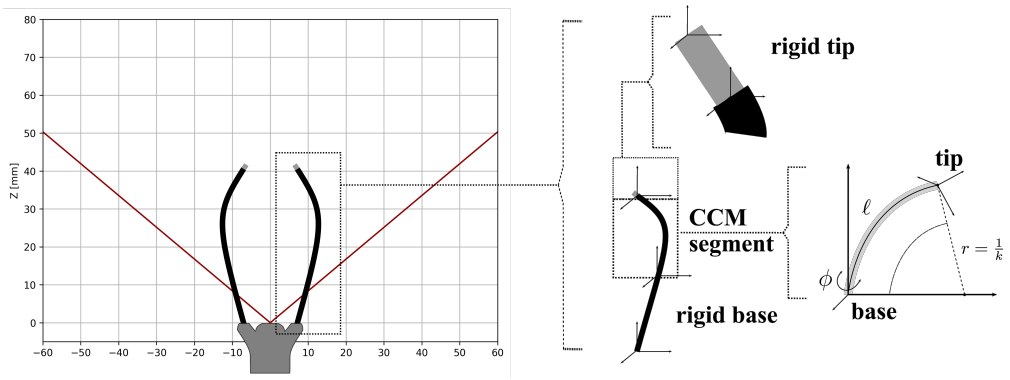


FIGURE 3.7: Overview of the constant curvature model in the STRAS robot. Overview of the constant curvature model used in the right arm. On the left the top view of the arms with red lines depicting the visible region. On the right is a sketch of the segments of the arms and the constant curvature model.

Where ${}_{base}\boldsymbol{T}^{tip}$ is a composition of maps ${}_{base}\boldsymbol{T}^{tip} = T(\phi, \kappa, s)$ and implies that chain rule can be used to compute the differential kinematics as presented in (Webster III and Jones, 2010).

3.3.3 Forward kinematic modelling

As firstly introduced in Bardou et al. (2010), the STRAS robot can be modelled under the assumption of constant curvature. Using this convention, the body and the arms are modelled by a constant curvature segment –the flexible part– followed by a rigid segment at its tip – that constitutes of the surgical tool. For the STRAS system, the rigid segment at their tip measures 10mm, while the constant curvature segment measures 18.5mm followed by a rigid segment adding up to a total of 75mm of length. Figure 3.7 depicts this for the dominant arm, and it is analogous to the non-dominant arm and the body.

The HTM defined in eq. (3.3) is not sufficient to model the arms or body of the STRAS system. A varying term of translation is required to correctly model the insertion and retraction of the arms or the body. Furthermore, the surgical tool at the tip of the arms is also required. Adding the requiring terms, the HTM of the endoscope body is then defined by chaining the HTMs as:

$$ {}_{w}\boldsymbol{T}^{e} = {}_{w}\boldsymbol{T}^{t_z} \cdot {}_{t_z}\boldsymbol{T}^{CCM} \cdot {}_{CCM}\boldsymbol{T}^{d_2} \tag{3.4} $$

where ${}_{w}\boldsymbol{T}^{t_z}$ is an HTM with t_z a constant term in the translation –insertion or retraction. For the body, the term ${}_{t_z}\boldsymbol{T}^{CCM}$ is analogous to eq. (3.3), and ${}_{CCM}\boldsymbol{T}^{d_2}$ is the rigid tip

at the end of the endoscope, such that d_2 is the length of the rigid tip of the endoscope. Therefore, ${}_w\boldsymbol{T}^{t_z}$ and ${}_{CCM}\boldsymbol{T}^{d_2}$ are populated by an identity matrix $\boldsymbol{I}_{3x3}$ and $[0, 0, \triangle]$ where $\triangle$ is defined by t_z or d_2 accordingly. From the FKM defined in eq. (3.3), the position of the tip of the endoscope after considering the transformations defined in eq. (3.4), is expressed as:

$$ {}_w\boldsymbol{p}^e = \begin{bmatrix} \frac{\cos(\theta)\,(d_2\,k\,\sin(\beta) - \cos(\beta) + 1)}{k} \\ \frac{\sin(\theta)\,(d_2\,k\,\sin(\beta) - \cos(\beta) + 1)}{k} \\ t_z + d_2\,\cos(\beta) + \frac{\sin(\beta)}{k} \end{bmatrix} \tag{3.5} $$

which is the translational component of eq. (3.4) of the FKM of the endoscope in the world frame, ${}_w\boldsymbol{T}^e(\boldsymbol{q_e})$. Where $\boldsymbol{q_e}$ is the appropriate term of the configuration variables eq. (3.2), defined as:

$$ \boldsymbol{q_e} = [\beta, \theta, t_z]^T, \tag{3.6} $$

omitting the length of the tip of the rigid endoscope tip, d_2. Furthermore, following the configuration variables defined for eq. (3.2), the term s is omitted for the consideration of tip control, and the term κ is defined by the variable θ due to the linear relationship $\theta = \kappa s$.

Similarly to eq. (3.4), one can describe the position of the tip of the arms. However, instead of considering the world frame $\{w\}$ as a base, consider $\{u\}$ and $\{v\}$ as the bases of the surgical tools. Analogous to eq. (3.4), ${}_u\boldsymbol{T}^r\ (\boldsymbol{q_r})$ and ${}_v\boldsymbol{T}^l\ (\boldsymbol{q_l})$ describe the HTM of the arms to their base $\{r\}$, $\{l\}$. The position of the arms with respect to the camera

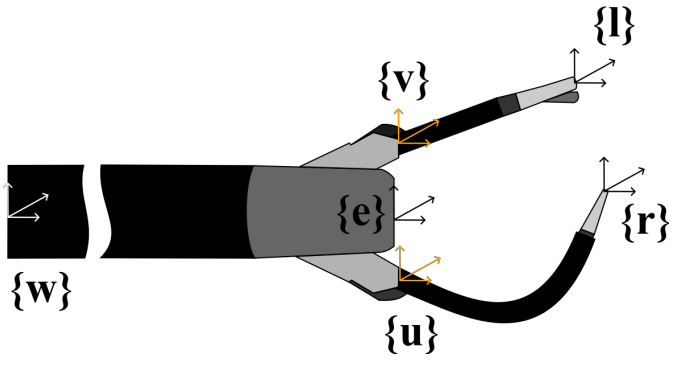


FIGURE 3.8: Frames of interest of the STRAS system.

which is also the body frame $\{e\}$ is of interest. From the base of each arm a constant

transformation is set, such that the HTM for the dominant arm:

$$ {}_e\boldsymbol{T}^r(\boldsymbol{q_r}) = {}_e\boldsymbol{T}^u \cdot {}_u\boldsymbol{T}^r(\boldsymbol{q_r}), \tag{3.7} $$

and similarly for the non-dominant arm in its base $\{v\}$ as:

$$ {}_e\boldsymbol{T}^l(\boldsymbol{q_l}) = {}_e\boldsymbol{T}^v \cdot {}_v\boldsymbol{T}^l(\boldsymbol{q_l}). \tag{3.8} $$

The position of the tip of the arm with respect to the static surgical target ($\{w\}$) is of interest. Equations (3.7) and (3.8) are chained to the HTM of the endoscope, such that the position of the arms in the world frame is defined as:

$$ {}_w\boldsymbol{T}^r(\boldsymbol{q_e}, \boldsymbol{q_r}) = {}_w\boldsymbol{T}^e(\boldsymbol{q_e}) \cdot {}_e\boldsymbol{T}^u \cdot {}_u\boldsymbol{T}^r(\boldsymbol{q_r}) \tag{3.9} $$

$$ {}_w\boldsymbol{T}^l(\boldsymbol{q_e}, \boldsymbol{q_l}) = {}_w\boldsymbol{T}^e(\boldsymbol{q_e}) \cdot {}_e\boldsymbol{T}^v \cdot {}_v\boldsymbol{T}^l(\boldsymbol{q_l}). \tag{3.10} $$

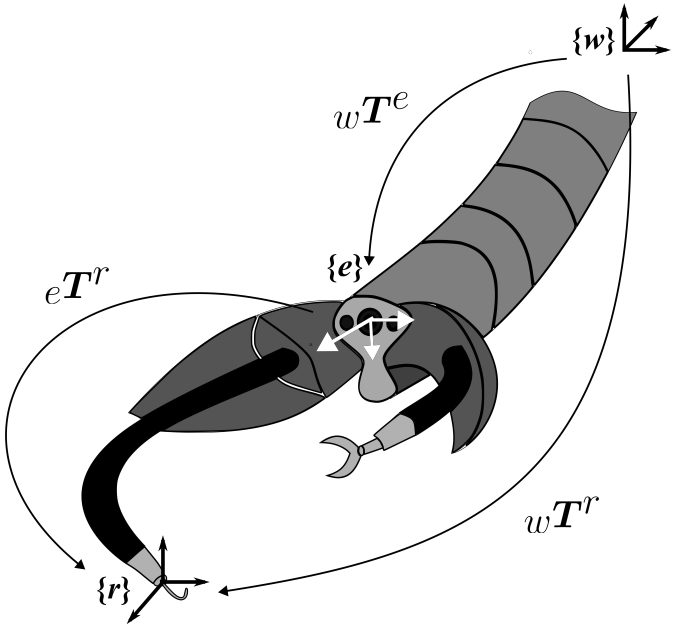


FIGURE 3.9: Tip positions of the arms and body of the STRAS system. HTM describes the position and orientation of the tip of the arms and body.

Equations (3.4) and (3.7) to (3.10) describe all of the frames of interest (see fig. 3.8) in the proposed controlled framework, depicted in fig. 3.9. Controlling the two sub-systems of the STRAS robot, arms and body, is done by controlling $\boldsymbol{q_e}$, $\boldsymbol{q_r}$, and $\boldsymbol{q_l}$ which define

the STRAS robot configuration variable vector as:

$$\boldsymbol{q} = \begin{bmatrix} \boldsymbol{q_r} \\ \boldsymbol{q_l} \\ \boldsymbol{q_e} \end{bmatrix}. \tag{3.11}$$

Finally, the FKM can then be described as:

$$\begin{aligned} \mathscr{M} : \mathbb{R}^{10} &\rightarrow SE(3), SE(3), SE(3) \\ \boldsymbol{q} &\mapsto {}_w\boldsymbol{T}^e(\boldsymbol{q_e}), {}_e\boldsymbol{T}^r(\boldsymbol{q_r}), {}_e\boldsymbol{T}^l(\boldsymbol{q_l}) \end{aligned} \tag{3.12}$$

3.3.4 Differential kinematic modelling

The differential kinematics describe the relations between the displacement (velocity) of the joint space and the motion (linear/angular velocity) of the task pace, which is not as straightforward as the FKM.

In each one of the subsystems, the CCM formulates the relationship between the linear joint velocities of eq. (3.2) and their position eq. (3.3) at its tip from its base by the Jacobian ${}_{base}J_{tip}$ referred as:

$${}_{base}\boldsymbol{J}_{tip} = \begin{bmatrix} \frac{\partial x}{\partial \boldsymbol{q}_{ccm}} \\ \frac{\partial y}{\partial \boldsymbol{q}_{ccm}} \\ \frac{\partial z}{\partial \boldsymbol{q}_{ccm}} \end{bmatrix} \tag{3.13}$$

which considers the appropriate $\boldsymbol{q}_{ccm}$ variables for each one of the subsystems of the STRAS system. Similar to the HTM defined for the FKM, eq. (3.4), the Jacobian of the endoscope can be defined as:

$${}_w\boldsymbol{J}_e = \begin{bmatrix} \frac{\partial X_w}{\partial \boldsymbol{q_e}} \\ \frac{\partial Y_w}{\partial \boldsymbol{q_e}} \\ \frac{\partial Z_w}{\partial \boldsymbol{q_e}} \end{bmatrix} \tag{3.14}$$

Such that ${}_w\boldsymbol{J}_e$ is the Jacobian of the endoscope and relates world displacements of the endoscope tip to $\dot{\boldsymbol{q}}_{\boldsymbol{e}}$.

$${}_w\dot{\boldsymbol{p}}^e = {}_w\boldsymbol{J}_e(\boldsymbol{q_e})\,\dot{\boldsymbol{q}}_{\boldsymbol{e}} \tag{3.15}$$

${}_e\boldsymbol{J}_r$ and ${}_e\boldsymbol{J}_l$ are the Jacobians of the arms in the endoscope frame. Considering the arms Jacobians in eqs. (3.7) and (3.8) as configuration variables that only act on the base of each arm –considering the constant transformations ${}_e\boldsymbol{T}^u$, ${}_e\boldsymbol{T}^v$ to ${}_w\boldsymbol{T}^e$– such that:

$$ {}_e\dot{\boldsymbol{p}}^r = {}_e\boldsymbol{J}_r(\boldsymbol{q_r})\,\dot{\boldsymbol{q}}_{\boldsymbol{r}} \tag{3.16} $$

$$ {}_e\dot{\boldsymbol{p}}^l = {}_e\boldsymbol{J}_l(\boldsymbol{q_l})\,\dot{\boldsymbol{q}}_{\boldsymbol{l}} \tag{3.17} $$

Such that the Jacobians of the arms in the world frame will require chaining ${}_e\boldsymbol{J}_r$ with ${}_w\boldsymbol{J}_e$, similarly for the non-dominant arm: ${}_e\boldsymbol{J}_l$ with ${}_w\boldsymbol{J}_e$. Such that:

$$ {}_w\dot{\boldsymbol{p}}^r = \begin{bmatrix} {}_e\boldsymbol{J}_r & {}_w\boldsymbol{J}_e \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{r}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix} \tag{3.18} $$

$$ {}_w\dot{\boldsymbol{p}}^r = {}_w\boldsymbol{J}_r \begin{bmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{r}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix}, \tag{3.19} $$

and similarly for the non-dominant arm:

$$ {}_w\dot{\boldsymbol{p}}^l = \begin{bmatrix} {}_e\boldsymbol{J}_l & {}_w\boldsymbol{J}_e \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{l}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix} \tag{3.20} $$

$$ {}_w\dot{\boldsymbol{p}}^l = {}_w\boldsymbol{J}_l \begin{bmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{l}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix}. \tag{3.21} $$

Equation (3.15) allows to modify the position of the tip of the body in the world frame by the configuration variables of the body. Equations (3.16) and (3.17) allow to modify the position of the tip of the arms in the endoscope frame by the arms configuration variables. Equations (3.20) and (3.21) allow to modify the position of the tip of the arms in the world frame by configuration variables of the arms and body, correspondingly.

In eqs. (3.18) and (3.20) the corresponding Jacobians ${}_w\boldsymbol{J}_r$ and ${}_w\boldsymbol{J}_l$ define the relationships between the joint values of the $\dot{\boldsymbol{q}}_{\boldsymbol{r}}$, $\dot{\boldsymbol{q}}_{\boldsymbol{l}}$ and $\dot{\boldsymbol{q}}_{\boldsymbol{e}}$. The orientation Jacobian of the arms can be omitted from the eqs. (3.16) and (3.17) since position control is desired, the same cannot be done to eqs. (3.18) and (3.20) since the body affects the position of the arms –posing a lever effect. The body which houses the camera, moves on a sphere whose radius is close to the length of the body (185mm), which remains true even when

the body is bent. Moreover, the amount of lateral movement considered for the body is low in comparison with the length of the body due to the limited space in the surgical environment. Such that for the relatively small movements performed by the body, the orientation effect of ${}_w\boldsymbol{J}_e$ over ${}_w\boldsymbol{J}_r$ and ${}_w\boldsymbol{J}_l$ can be neglected.

Each of the subsystem Jacobians follows the behaviour of the constant curvature model. The Jacobians ${}_e\boldsymbol{J}_r$, ${}_e\boldsymbol{J}_l$, ${}_w\boldsymbol{J}_e$ are populated following the approach used by Nageotte et al. (2020).

3.3.5 Workspace characterization

Each arm workspace consists of a three-dimensional cylinder with a spherical top. The cylinders start from the base of the channel in which the arms start. The arm bending forms a circular path that is rotated over the channel's axis, as to form such surface. The arm translations sweep the surface over the Z-axis, to form the cylinder with a spherical top. The arms forms a theoretical workspace as truncated cylinders of radios 32mm and height of 75mm, fig. 3.10.

Figure 3.11 depicts the workspace of the dominant arm. It shows the lower and upper bounds of the translation reach depicted in red and blue respectively. The singularities are depicted in yellow. The outer blue sphere depicts a position that can be reached in 4 discrete orientations, the red sphere depicts where the position can be reached in one discrete orientation, while the rest of the positions can be reached in at least two opposite orientations.

3.4 STRAS system telemanipulation

The STRAS system is aimed to be manually introduced into the patient's body until the desired area of interest is reached in the digestive system. During manual insertion, the endoscope is covered by a shell that eases navigation through the anus. When the system has reached the desired point of interest in the colon, the shell opens and the system is mounted in the motors to allow telemanipulation. After the surgical procedure has ended, the shell can be closed to facilitate exiting the digestive system. The shell

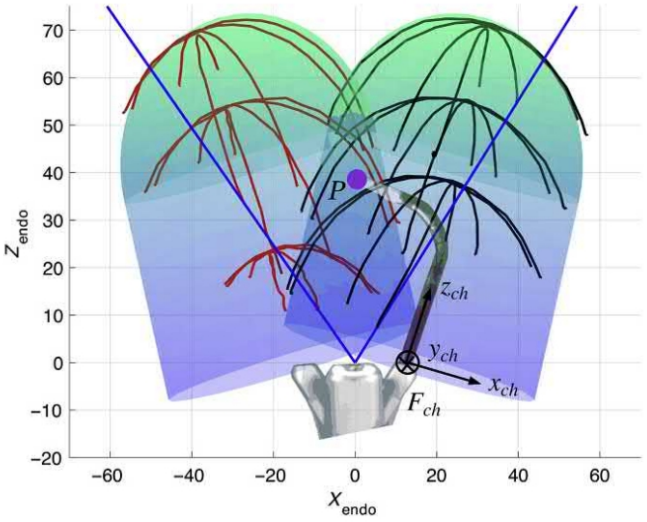


FIGURE 3.10: Theoretical workspace of the arms from the top view. Image adapted from Nageotte et al. (2020).

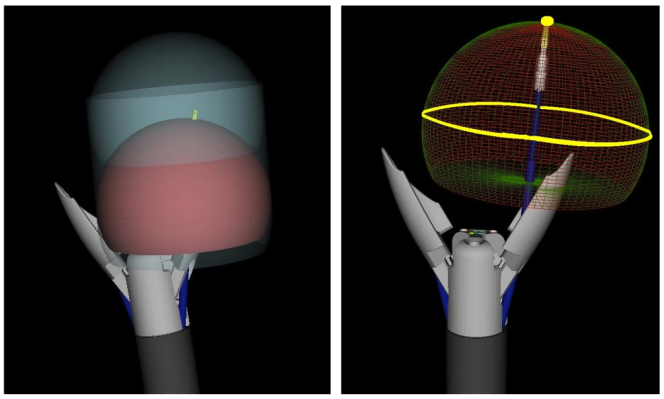


FIGURE 3.11: Workspace of the dominant arm. The workspace for the dominant arm is shown in red for the lower bound of the translation and in blue for the upper bound. Singularities are depicted in yellow. The outer blue sphere depicts a position that can be reached in 4 discrete orientations, the red sphere depicts where the position can be reached in one discrete orientation, while the rest of the positions can be reached in at least two opposite orientations. Image adapted from De Donno et al. (2013a).

stays open to allow grabbing of the tissue and allowing retraction of the tissue for later analysis.

During telemanipulation, the follower is linked to the leader using the kinematic model of the STRAS system. The leader configuration variables are defined as:

$$\boldsymbol{q} = \begin{bmatrix} \boldsymbol{q}_r \\ \boldsymbol{q}_l \\ \boldsymbol{q}_e \end{bmatrix}. \tag{3.22}$$

During manual telemanipulation, the leader configuration variables $\boldsymbol{q}$ are proportionally mapped to the configuration variables of the follower side $\boldsymbol{q}$ by the matrix $\boldsymbol{K}$. The arms are controlled in position at the joint level by:

$$\begin{bmatrix} \boldsymbol{q_r} \\ \boldsymbol{q_l} \end{bmatrix} = \boldsymbol{K}_{arms} \begin{bmatrix} \boldsymbol{q_r} \\ \boldsymbol{q_l} \end{bmatrix}, \tag{3.23}$$

with

$$\boldsymbol{K}_{arms} = \begin{bmatrix} \boldsymbol{K}_r \\ \boldsymbol{K}_l \end{bmatrix} \tag{3.24}$$

being a stack of the two matrices $\boldsymbol{K}_r$ and $\boldsymbol{K}_l$ that are diagonally filled ensuring a correspondence between the range of motion of corresponding leader and follower DoF, as previously defined in Nageotte et al. (2020). The body is rate-controlled via the two joysticks on the handles by a proportionality constant, such that:

$$\dot{\boldsymbol{q}}_{\boldsymbol{e}} = \boldsymbol{K}_e \boldsymbol{q}_{\boldsymbol{e}}. \tag{3.25}$$

With $\boldsymbol{K}_e$ and $\boldsymbol{q}_{\boldsymbol{e}}$ considering both joysticks.

The full model of the STRAS robot to perform telemanipulation is provided by combining eqs. (3.9), (3.10) and (3.12) and eqs. (3.23) and (3.25). During telemanipulation and due to hardware limitations and safety concerns, speeds for the arms (eq. (3.23)) and the body (eq. (3.25)) are bound such that: $\dot{\boldsymbol{q}}^h_{min} < \dot{\boldsymbol{q}} < \dot{\boldsymbol{q}}^h_{max}$ where $\dot{\boldsymbol{q}}^h_{min}$ and $\dot{\boldsymbol{q}}^h_{max}$ are the negative and positive value of the maximum speed.

3.4.1 STRAS system ranges

The leader-side handles (depicted in fig. 3.12) can be rotated in the interval $\boldsymbol{q_r}$, $\boldsymbol{q_l}$ in $[-44, 44]^\circ$ this interval is thus mapped in the bending of the arms in the follower system, in the interval $[-90, 90]^\circ$. Such bending is controlled at the low level by a cable displacement in the range $[-6.5, 6.5]$ mm. The handles can translate $[0, 90]^\circ$, which is mapped to $[0, 72]$ mm of arm insertion. The rotation of the handles ranges between $[-160, 160]^\circ$ that is mapped to $[-270, 360]^\circ$ arm rotation on the follower side.

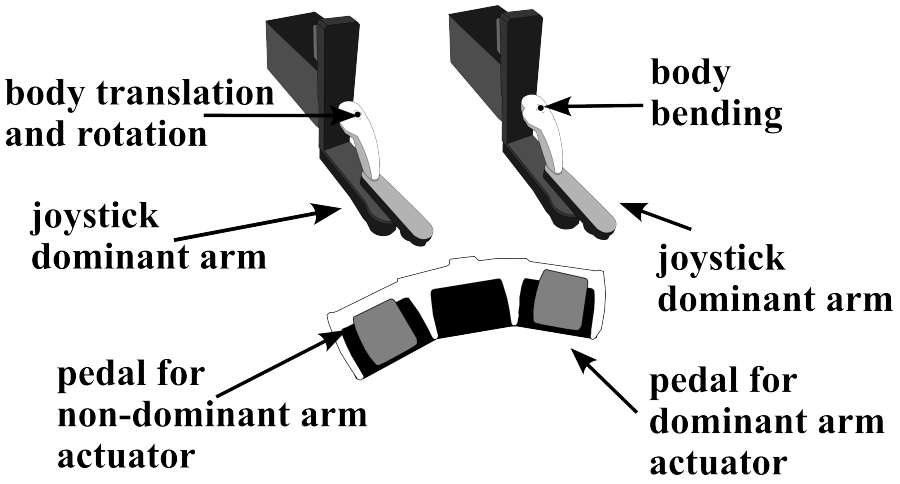


FIGURE 3.12: Leader side of the STRAS system. The main components of the leader system: are handles, joysticks at the tips of the handles, and the pedals.

The joysticks at the tip of the handle are mapped to the movement of the body. The body can translate $[0.5, 90]$ mm and rotate $[-90, 50]°$. The bending in the two orthogonal directions with low-level displacement cable ranges between $[-9.12, 9.12]$mm. The joystick values can be defined by software to apply a constant velocity or to increase the velocity the longer it is being actuated. The grippers at the back of the handle range between $[0, 60]°$ and can be actuated between $[0, 1]$ in the follower, allowing for open, closed and half-open.

The table 3.1 summarizes the range of the platforms and allows the definition of the interpolation required to map the actuator space to the configuration space as depicted in fig. 3.6.

3.5 The STRAS system as validation platform

The STRAS system shares similarities (section 2.2) with other FSIR platforms such as the k-Flex, FLEX, and EndoSamurai. The STRAS system, similarly to other platforms, has a coupled architecture. Telemanipulating flexible steerable intraluminal robots requires users to consider the dependence of the arms and body to correctly control the end effectors to perform surgical gestures. FSIR have a high number of degrees of freedom, that added to the coupled architecture, results in mentally demanding surgical procedures. The similarities of FSIR platforms allow this thesis work to address the

TABLE 3.1: Ranges of the STRAS system leader, follower and configuration space. L is the length of the curvature as in the CCM established in Nageotte et al. (2020). Depicting the values required for the mapping between the three spaces fig. 3.6. Rotation and bending measured in radians, translation measured in mm.

		leader console		config. space		actuator space	
		Min	Max	Min	Max	Min	Max
dominant arm	bending	-44	44	$-\pi/2L$	$\pi/2L$	-6.5	6.5
	rotation	-163	163	$-\pi$	π	$-3/2\pi$	2π
	translation	0	105	2	72	0.5	73
non-dominant arm	bending	-44	44	$-\pi/2L$	$\pi/2L$	-6.5	6.5
	rotation	-163	163	$-\pi$	π	$-3/2\pi$	2π
	translation	0	97	2	72	1	73.5
body	horizontal bending	-1	1	$-\pi/2L$	$\pi/2L$	-9.12	9.12
	vertical bending	-1	1	$-\pi/2L$	$\pi/2$	-9.12	9.12
	rotation	-1	1	$-\pi$	π	$-\pi/2$	$5/18\pi$
	translation	-1	1	2	100	0.5	90

challenges ($\mathscr{C}_1$**: coupled architecture** and $\mathscr{C}_2$**: high amount of degrees of freedom**) generally while considering the specifics of the STRAS system. An approach for assisting the user in correctly performing complex surgical procedures while addressing the challenges of FSIR platforms is proposed in the following chapter.

Chapter 4

Assistance through high-level control in flexible steerable intraluminal platforms

As discussed in the previous chapter, the challenges of FSIR platforms are mainly the coupled architecture ($\mathscr{C}_1$) and the high number of degrees of freedom ($\mathscr{C}_2$). Telemanipulating FSIR requires users to manage these two challenges. Given that flexible steerable intraluminal robots are developed for complex surgical procedures such as EMR or ESD, reducing the complexity of telemanipulating FSIR platforms is desired. To design a system that can assist the user in the control of FSIR platforms during telemanipulation, a number of elements should be accounted for: safety for patients and surgeons, and transparency. In this chapter, a control framework is proposed to implement a controller that enables safe and transparent telemanipulation for the users and patients [1]. The control framework is generic but it is demonstrated on the STRAS robot under a specific surgical task which is the dissection stage during endoscopic submucosal dissection.

This chapter is organised in the following manner. Section 4.1 reviews the endoscopic submucosal dissection performed by telemanipulating a flexible steerable intraluminal

[1] This work has been published as:
Gonzalez-Herrera, F., Nageotte, F., Zanne, P., Borghesan, G., de Mathelin, M., Vander Poorten, E., & Rosa, B. (2024) A semi-autonomous control mode for flexible steerable intraluminal platforms, IEEE Transactions on Medical Robotics and Bionics. In press (12 pages), DOI 10.1109/TMRB.2024.3385990
Herrera, J. F. G., Nageotte, F., Borghesan, G., Vander Poorten, E., & Rosa, B. (2023). Semi-autonomous control for endoluminal robotic platforms. In Computer and Robot-Assisted Surgery (CRAS), Paris, France, September 2023

platform which is then modelled in section 4.2 and described in section 4.3. An overview of the related works to the specific problem is presented in section 4.4. Then, the formalized problem is solved through the proposed approach in section 4.5 and the experiments used to test the approach are described in section 4.6 with the results presented in section 4.7. This chapter ends with section 4.8 which discusses the results and describes the future work.

4.1 Endoscopic submucosal dissection by the FSIR platforms

Complex surgical procedures are characterized by a high mental demand due to the complex and precise surgical gestures the surgical arms must perform in intricate surgical environments. Given the complexity of the procedure, the surgeon is under considerable mental demand.

Before the dissection stage of ESD, the user reaches the desired area of interest in the body. Then, the user commands the non-dominant arm to grab and lift the tissue while maintaining its position, meaning the user establishes *traction* over the tissue. During traction, a force is continuously applied to the tissue to tension it. Having traction (Portolés et al., 2015; Shahkoo and Abin, 2023) on the tissue refers to the act of grabbing and pulling away the tissue. Tissue traction is established by maintaining the position of the arm, grabbing without releasing the tissue. Thanks to the traction, the tissue under traction can be exposed. The dominant arm, equipped with an electric knife, precisely reaches and cuts the boundary of the abnormal tissue. After the dominant arm has finished cutting the exposed tissue at the current body position, the user releases the tissue and moves the body to a different position. In this new body position, the user repeats the procedure of grabbing, lifting –maintaining traction– and cutting until the entire tissue has been removed. The ESD ends with the surgeon retracting the FSIR platform while grabbing the dissected tissue. This set of steps is complex to perform under manual telemanipulation as it involves *coordinated control* of the subsystems of FSIR. During coordinated control, the traction over the tissue should be maintained constant while the body is repositioned. This is difficult for users to conduct due to the coupled architecture and the high number of degrees of freedom.

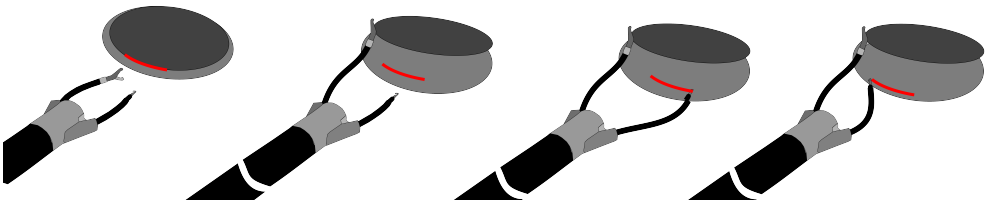


FIGURE 4.1: Overview of the envisioned scenario. A simplification of the tissue is shown in grey. A subsection of the tissue that can be removed at the current body position is shown in red. The body and the arms are shown in four cases, from left to right, free arms and body, in the second the tissue is grabbed by the non-dominant arm, while in the third and fourth cases, the segment is removed by reaching multiple reaching locations.

The inherent complexity of the task causes the user, instead of having a fluid coordinated control of the arms and the endoscope, to move the body and arms sequentially. Individual sequential control of the arms increases the amount of grasping that is required to completely remove the abnormal tissue. Thus, instead of allowing multiple cuttings per one grab, the user is required to grab multiple times and perform multiple cuttings gestures. This inefficient execution approach to the task has been noted during pre-clinical trials (Légner et al., 2017) and benchtop experiments (De Donno et al., 2013a). The user increases the number of times the tissue is being grabbed, by evading coordinated control, which potentially leads to tissue damage (Saito et al., 2014) and increases the chance of bleeding (Lee et al., 2011; Harlow et al., 2020). Dissection could be simplified if the FSIR systems allowed straightforward coordinated control.

Coordinated control can involve one arm and the body or the two arms and the body. The user is required to control 7 DoF if one arm and body are required to be moved, while control of the whole platform involves 10 DoF. The user must control the two arms and the body during dissection, involving the 10 DoF. For the STRAS system controlling the 10 DoF requires coordinating arm-hand-finger movement of the user which translates to arms-body commands. Surgeons using their fingers, arms and hands to control the body and arms of the follower side by the leader console is prone to errors: involuntary movements due to muscle memory caused by prior movements or bias in other surgical robotic platforms, slow movements, and erratic non-smooth movement patterns by the user. The challenges ($\mathscr{C}_1$, $\mathscr{C}_2$) of the control results in inaccurate movements of the body that might require correction actions. Eventually, this strains and frustrates the surgeon, leading to a longer procedure time and impacting the overall surgeon performance (Zheng et al., 2012; Mehta et al., 2014). An approach that reduces the number of DoFs to be

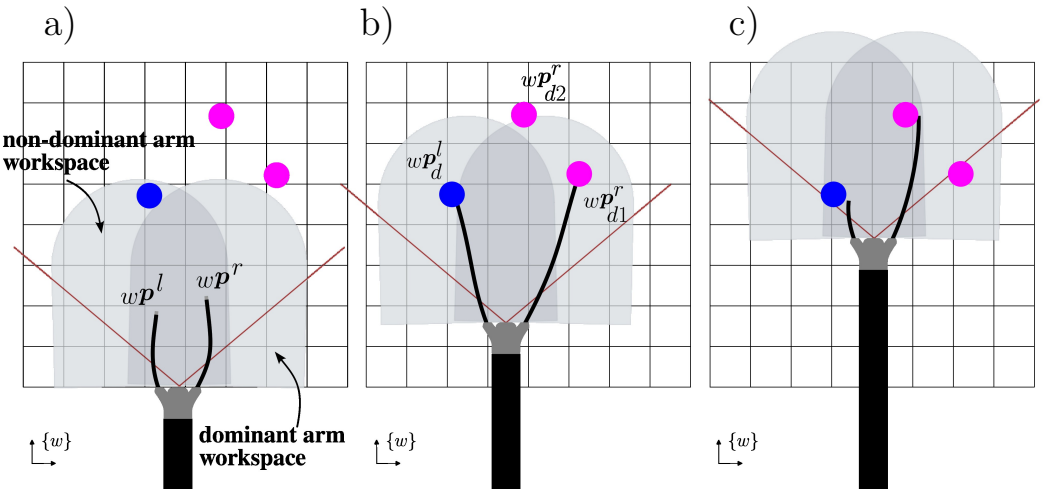


FIGURE 4.2: Overview of the problem. The workspace of the arms is depicted in grey and moves as the body moves, the dark red line depicts the bounds of the camera's field of view. a) The current position of the body allows to reach the target by the non-dominant arm $_w\boldsymbol{p}^l_d$ but not with the dominant arm $_w\boldsymbol{p}^r_{d1}$. b) The body is moved to reach $_w\boldsymbol{p}^l_d$ and the new body position allows to reach $_w\boldsymbol{p}^r_{d1}$ but not $_w\boldsymbol{p}^r_{d2}$. c) While $_w\boldsymbol{p}^l_d$ is not violated, the second target position of the dominant arm $_w\boldsymbol{p}^r_{d2}$ is reached while moving the body.

controlled while not diminishing the performance of the surgical task is needed. Correctly performing the surgical task requires users to maintain traction, perform coordinated control and consider the safety constraints for the patient.

4.2 Problem modelling

Establishing traction while dissecting is modelled by multiple positions the left and right arms must reach. Assuming without loss of generality that the dominant arm is the right arm. The positions for the tip of the arms aim to generalise the gestures required to be performed during dissection. The tip position of the dominant and non-dominant arms are described by $_w\boldsymbol{p}^r$, $_w\boldsymbol{p}^l$ as defined by eqs. (3.8) and (3.9). The envisioned scenario is as follows: the end-effector of the non-dominant arm grabs the tissue and lifts it to a desired position $_w\boldsymbol{p}^l_d$, setting $_w\boldsymbol{p}^l =_w \boldsymbol{p}^l_d$ (fig. 4.2-a). Controlling the position of the non-dominant arm in the world frame, $_w\boldsymbol{p}^l$, involves using the arm and body to control its velocity, e.g. by:

$$_w\dot{\boldsymbol{p}}^l = {}_w\boldsymbol{J}_l \begin{bmatrix} \boldsymbol{K}_l\dot{\boldsymbol{q}}_l \\ \boldsymbol{K}_e\dot{\boldsymbol{q}}_e \end{bmatrix} \tag{4.1}$$

and when the current position the arm has to reach ${}_w\boldsymbol{p}_d^l$ is reachable at the current body position, $\dot{\boldsymbol{q}}_e = 0$, it is simplified as:

$$ {}_e\dot{\boldsymbol{p}}^l = {}_e\boldsymbol{J}_l \left[\boldsymbol{K}_l \dot{\boldsymbol{q}}_l \right] \tag{4.2} $$

The user then cuts the tissue. The tissue cutting positions are modelled as two positions the tip dominant arm must reach, ${}_w\boldsymbol{p}_{d1}^r$ and ${}_w\boldsymbol{p}_{d2}^r$. Reaching ${}_w\boldsymbol{p}_{d1}^r$ or ${}_w\boldsymbol{p}_{d2}^r$ may require moving the body, by controlling its tip position ${}_w\boldsymbol{p}^e$, such that if body movement is required:

$$ {}_w\dot{\boldsymbol{p}}^r = {}_w\boldsymbol{J}_r \begin{bmatrix} \boldsymbol{K}_r \dot{\boldsymbol{q}}_r \\ \boldsymbol{K}_e \dot{\boldsymbol{q}}_e \end{bmatrix} \tag{4.3} $$

else a similar expression to eq. (4.2) is generated with the DoF of the dominant arm. When the dominant arm is required to be moved, coordinated movement of the arms and the body should be performed by the user. Coordinated movement while maintaining traction requires the user to control the body

$$ {}_w\dot{\boldsymbol{p}}^e = {}_w\boldsymbol{J}_e \left[\boldsymbol{K}_e \dot{\boldsymbol{q}}_e \right] \tag{4.4} $$

while accounting for the movements of the dominant arm expressed in eq. (4.3). The coupled architecture affects the position of the non-dominant arm in the global frame, as depicted by eq. (4.1) –about the initial expression eq. (3.20). During such movements, and more generally during the whole cutting phase, the non-dominant arm tip position ${}_w\boldsymbol{p}^l = {}_w\boldsymbol{p}_d^l$ should be kept steady, maintaining traction (fig. 4.2-b,c) requiring coordinated control without impacting the traction of the non-dominant arm since the tissue is being grabbed. Coordinated movement without impact on the non-dominant arm traction is described by:

$$ \text{traction and coordinated control} \begin{cases} {}_w\dot{\boldsymbol{p}}^r = \begin{bmatrix} {}_e\boldsymbol{J}_r \boldsymbol{K}_r \dot{\boldsymbol{q}}_r \\ {}_w\boldsymbol{J}_e \boldsymbol{K}_e \dot{\boldsymbol{q}}_e \end{bmatrix} = 0 \\ {}_w\dot{\boldsymbol{p}}^l = \begin{bmatrix} {}_e\boldsymbol{J}_l \boldsymbol{K}_l \dot{\boldsymbol{q}}_l \\ {}_w\boldsymbol{J}_e \boldsymbol{K}_e \dot{\boldsymbol{q}}_e \end{bmatrix} = 0 \\ {}_w\dot{\boldsymbol{p}}^e = \begin{bmatrix} {}_w\boldsymbol{J}_e \boldsymbol{K}_e \dot{\boldsymbol{q}}_e \end{bmatrix} \end{cases} \tag{4.5} $$

which involves coordinated control of all the elements of the subsystem while moving the body without impacting the position of the non-dominant arm –which is grabbing the tissue– and minimizing the effects on the position of the dominant arm —with the cutting tool, avoiding unsafe positions.

4.3 Overview of the problem

Dealing with the multiple DoFs (fig. 4.3) and the coupled architecture to correctly perform complex surgical gestures requires controlling the multiple subsystems (arms and body). In the specific case of gaining and maintaining traction on the tissue during the dissection stage of ESD, the user must not violate the required positions for the multiple frames. Performing traction by telemanipulation during dissection (ESD) requires the user to control all of the subsystems eq. (4.5) and consider the effects of the coupled architecture:

$$\text{coupled architecture} : \begin{cases} {}_w\boldsymbol{T}^r(\boldsymbol{q_e}, \boldsymbol{q_r}) = {}_w\boldsymbol{T}^e(\boldsymbol{q_e}) \cdot {}_e\boldsymbol{T}^u \cdot {}_u\boldsymbol{T}^r(\boldsymbol{q_r}) \\ {}_w\boldsymbol{T}^l(\boldsymbol{q_e}, \boldsymbol{q_l}) = {}_w\boldsymbol{T}^e(\boldsymbol{q_e}) \cdot {}_e\boldsymbol{T}^v \cdot {}_v\boldsymbol{T}^l(\boldsymbol{q_l}) \\ {}_w\boldsymbol{T}^e(\boldsymbol{q_e}) = {}_w\boldsymbol{T}^e(\boldsymbol{q_e}) \end{cases} \tag{4.6}$$

with a total of 10 DoF to be controlled. The user should perform coordinated control eq. (4.5) which is not easy even for experienced for users.

Supplementary to the surgical task, the user should consider the patient's safety and avoid any abrupt movements, avoiding unnecessary interactions with the digestive system of the patient. There is no official set of safety guidelines for performing procedures with FSIR. Yet, the similarity to flexible endoscopes and laparoscopes allows us to come up with a list of basic rules to ensure the procedure is being carried out safely. Similar to laparoscopes, the arms in FSIR should be kept in sight. Blind control of the surgical tools can pose a risk of bleeding or unnecessary tool-tissue interaction. Thus, the tools should stay inside the field of view of the endoscopic camera placed at the origin of the frame e $\{e\}$, such that arms are kept in the endoscopic camera sight during the semi-autonomous arm-body control mode.

Furthermore, the arms should be restricted to positions that avoid degradation of the surgical site view or reduce depth perception –hinders the position of the surgical target,

e.g. puts the surgical target away from the centre of the screen or the arms are on the limits of the screen. The arm should not be too close to the camera placed at the tip of the body $\{e\}$ to avoid occluding the view of the surgical site. Arms being close to the camera reduces the view of the surgical site. Opposite to the arms being too close to the surgical site, when the arms are too far from the camera the perception of the arms is reduced. A reduced depth perception increases the mental demand for precisely controlling the arms. Arms being too far away also degrades the user's motion capability due to the presence of mechanical singularities on the related joint configuration. In summary, joint configurations where the arm is too close to the boundaries field of view or too close to the joint limits (due to the mechanical singularities) should be avoided.

A semi-autonomous arm-body control could further support safe and simple coordinated control to facilitate maintaining traction during the dissection stage performed by FSIR.

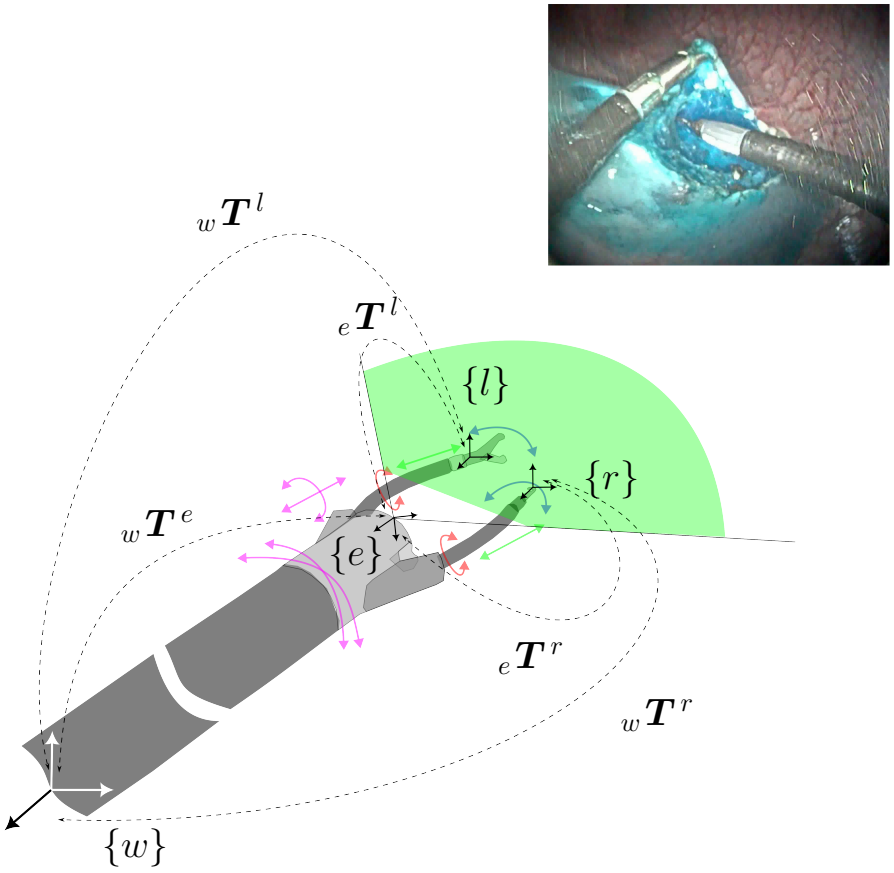


FIGURE 4.3: Overview of the control problem. DoFs of the follower are depicted in a circle. Featured frames to be controlled and their transformation are depicted by a dashed line. Overview of the surgical task from the endoscopic camera view. The safe workspace of the arms is illustrated in green.

The proposed control approach would control the position of the dominant arm in the world $\{w\}$ and endoscope frame $\{e\}$, ${}_{w}\boldsymbol{p}^{r}$ and ${}_{e}\boldsymbol{p}^{r}$ respectively as:

$$\begin{bmatrix} {}_{e}\boldsymbol{p}^{r} \\ {}_{w}\boldsymbol{p}^{r} \\ {}_{w}\boldsymbol{p}^{l} \end{bmatrix}. \tag{4.7}$$

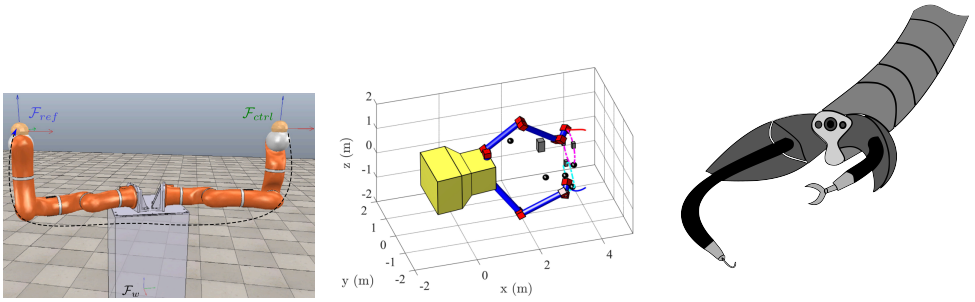
The three positions are selected due to the constraints of arms movement in the world frame ${}_{w}\boldsymbol{p}^{r}$ and ${}_{w}\boldsymbol{p}^{l}$. Furthermore, the position of the dominant arm position at the current position of the body must be controlled since the dominant arm carries the electronic knife. The position of the dominant arm in the body position is constrained to safe positions that are both visible and away from the joint limits to preserve the patients safety.

These three positions are of interest since the arms position with respect to the surgical task The proposed semi-autonomous arm-body control requires considering the multiple DoF of the system eq. (3.11), the coupled architecture, and the desired positions to be controlled eq. (4.7) —as illustrated in fig. 4.3. This means managing the independent DoF for each one of the configuration variables –for each of the subsystems eq. (3.12)– defined in eq. (3.11) while considering the effect on the featured frames that describe the positions eq. (4.7).

4.4 Related works

The problem tackled in this chapter is the coordinated control of dual-arm robotic structures in a mobile base, which has arisen in various areas of robotics, including dual-arm mobile robots (Freddi et al., 2016; Buhl et al., 2019; Que et al., 2022), upper-body humanoid robots (Krüger et al., 2011; Rader et al., 2016; Hoffman et al., 2018), space robots (Wilcox et al., 1989; Yoshida et al., 1991; Flores-Abad et al., 2014), and surgical robots (Mitsuishi et al., 2013; Vandebroek et al., 2019; Nageotte et al., 2020) (see fig. 4.4). In such robotic platforms, a similar challenge to the $\boldsymbol{\mathscr{C}_1}$ challenge is present: the mobile base can aid in fulfilling the tasks of the arms, but it also affects their movement, which may represent a disturbance. Control of the arms and base amount to high number of DoF ($\boldsymbol{\mathscr{C}_2}$) and the effects on each other must be considering

to accomplish the main task. In practice, this requires controlling the position of multiple end-effectors by multiple DoFs while accounting for the effects on each other. Several approaches for handling this intricate relationship have been developed and a summary of the most representative ones is presented now.



(a) (b) (c)

FIGURE 4.4: Examples of similar arm-body robotic structures from the literature. a) dual-arm upper-body humanoid robot (Tarbouriech et al., 2022), b) space-robot simulation (Cai et al., 2022) c) STRAS robot

4.4.1 Control frameworks

There are broadly two families of approaches which tackle multi-DoF arm-body control in the literature.

The first one constructs an analytical formulation in which a task hierarchy is handled, exploiting properties of the Jacobian of the control problem and null-space projections (Liegeois et al., 1977; Slotine and Siciliano, 1991). Null-space projection deals with sequential tasks (Mistry et al., 2007, 2008) in the sense that a main task is solved and then sub-sequential tasks can be solved, without violating the first –or the upper– task. To do so, subsequent tasks are projected in the kernel, or the null-space of the previous one. The idea of null-space projection has been applied in prioritized inverse kinematics (Mistry et al., 2007, 2008), acceleration-based control (Hsu et al., 1989; Flacco et al., 2012) and joint torque control (Khatib, 1987; Sentis and Khatib, 2004a; Sentis et al., 2010). With regard to arm and mobile base robotic structures, the null-space projection has been used in dual-arm mobile robots (Freddi et al., 2016; Di Lillo et al., 2020), space robots (Yan et al., 2018; Wenfu et al., 2019; Yan et al., 2020), humanoid robots (Sentis and Khatib, 2004b; Dietrich et al., 2012, 2015) and personal robots (Philippsen et al., 2011; Geisert et al., 2017; Sakr et al., 2022). In surgical robots with a static base,

null-space projection has been used for single-arm redundant manipulator in a human collaboration task for safe telemanipulation (Su et al., 2019) and a dual-arm system for safe haptic-based shared-control telemanipulation (Selvaggio et al., 2018; Qu et al., 2019).

Multiple formalisms of the null-space projection idea have been developed, allowing to handle equality constraints (Slotine and Siciliano, 1991; De Lasa and Hertzmann, 2009), and inequality constraints by exploiting potential fields (Khatib, 1986; Sentis and Khatib, 2004b; Mansard et al., 2009). Fundamentally, they all rely on the idea that tasks at a lower level are carried out in the null-space of higher-priority tasks. Therefore, a lower-priority task cannot be achieved if it requires even a slight violation of the higher-priority task. Such a strict task hierarchy can be beneficial in specific contexts -for instance, a humanoid robot must ensure that the balance is maintained at all times otherwise it will fall- but can also be a strong limiting factor in many cases, especially in medical contexts where the redundancy of the robotic system is limited by their architecture.

The second big family exploits optimization-based methods. In this context, an important work allowing to design and control the motion of complex, multi-dof robots is the *task function* approach (Samson et al., 1991). The basic idea behind the task function approach is that many robot tasks may be reduced to a problem of positioning and that the control problem reduces into regulating a vector function, known as the task function, which characterizes the task. The task function approach can be used for position and force control (Samson and Espiau, 1990; Espiau et al., 1990; Samson, 1991). To facilitate task-function implementation, frameworks have been developed to enable high-level descriptions of the task which are later transformed into optimization-based expressions. The frameworks enable task-function description of robot tasks to create expressions that can be handled by optimization solvers (Ferreau, 2007; Diamond and Boyd, 2016). Frameworks such as eTaSL (Aertbeliën and De Schutter, 2014), which is inspired by iTASC (De Schutter et al., 2007; Decré et al., 2009), allow the expression of complex robot tasks to allow optimal control. The solver typically handles a quadratic programming (QP) formulation of the task enabling a velocity-resolved solution of positional tasks[2]

[2]Note here that there are other ways of controlling a robot, for instance using acceleration or torque-resolved controllers. The focus here is on velocity-resolved controller because in many medical robots,

Expressing the optimization problem is versatile and allows handling inequality (Khatib, 1986; Sentis and Khatib, 2004a) and equality (Mansard et al., 2009) constraints, which can be defined in the task- and joint-space, possibly combining constraints defined in different frames of references. Similarly to null-space projection, one can stack tasks and constraints in a hierarchical framework. In this case, when a solution for a higher-priority task is obtained, the other tasks solve another QP for a lower-priority task, without increasing the obtained minimum of the previous task objective. The hierarchical quadratic programming (HQP) formulation has been used in whole-body control of humanoid robots (Escande et al., 2014), humanoid upper-body (Hoffman et al., 2018; Tarbouriech et al., 2022), personal robots (Wang et al., 2014, 2018b) and space-robots (Cai et al., 2022), dual-arm robots (Hong et al., 2022) and arm-mounted drones (Cataldi et al., 2019; Imanberdiyev and Kayacan, 2020; Chen et al., 2023). This formulation suffers from the same drawback as null-space projection, i.e. the obligation to define -and comply to- a strict hierarchy of tasks.

It is however possible, with QP formulations, to enable non-strict priorities of the task, allowing a higher level of control. Instead of strictly higher or strictly lower, a non-strict priority parametrization of tasks can be handled by weighting strategies (Liegeois et al., 1977; Abe et al., 2007; Escande et al., 2010; Bouyarmane and Kheddar, 2011; Liu et al., 2011; Salini et al., 2011). Weighted-QP has been used in surgical platforms (Zhang et al., 2021; Colan et al., 2023), space robots (Misra and Bai, 2017; Bouyarmane et al., 2018), dual-arm robots (Wang and Wang, 2020; Yi et al., 2020) and humanoid robots (Escande et al., 2014; Djeha et al., 2023). In QP approaches with weights, all the tasks and constraints are parametrized continuously and solved in one QP problem. Such strategies allow a flexible definition and combination of tasks. The values of the different weights set their relative importance, which allows soft violation of some tasks at the trade-off that the rest of the tasks can also be achieved, depending on the selected weights. The main drawback is that it is not possible to formally guarantee that any task is strictly fulfilled.

In this thesis multiple coupled end-effectors are required to be simultaneously controlled with conflicting tasks, due to the coupled architecture. Null-space projection and HQP approaches are not a viable option. In such a hierarchical approach, it would be natural

and especially flexible endoscopy systems, torques and acceleration signals are either very noisy or inaccessible.

to put the safety task as the highest priory one, meaning the controller should not move the robot if that puts the patient at risk. However, due to the nature of the clinical scenario and the coupled architecture of the system, there is no clear, strict hierarchy between the arm-body and the arm-independent tasks. Putting one in front of the other may create blocking situations due to conflicts between task objectives. For this reason, a weighted QP strategy is selected, to soften the behaviour of the hierarchy by weighted quadratic programming. The weights allow the control framework to establish coordinated control of the arms and body while performing conflicting tasks simultaneously for each of the subsystems to enable the desired assisted behaviour. In the problem of this thesis, the non-strict hierarchy of WQP allows coordinated control of the body and arms of FSIR. Such behaviour is especially required when dealing with the coupled architecture allowing arm-independent and arm-body control in a safe and coordinated manner which is not computationally expensive. In the next section, details of the basic mathematical derivations of the selected approach are described. The formulation of clinical objectives, workflow, and constraints as control objectives within this framework, which constitutes the core contribution of this chapter, will then be detailed in section 4.5.

4.4.2 Quadratic formulation

In order to best introduce the weighted QP approach, let us first consider a simple inverse kinematics problem. The inverse kinematics problem can be formulated as an optimisation problem such that:

$$\min_{\dot{\boldsymbol{q}}} \dot{\boldsymbol{q}}^T \boldsymbol{H} \dot{\boldsymbol{q}} \tag{4.8}$$

$$s.t.\ \boldsymbol{J}\dot{\boldsymbol{q}} = \dot{\boldsymbol{x}} \tag{4.9}$$

with $\boldsymbol{H} = \boldsymbol{J}^T\boldsymbol{J}$ the Hessian matrix, and $\boldsymbol{J}$ the Jacobian of the robot. This simple optimization problem solves for $\dot{\boldsymbol{q}}$ to achieve the desired equality eq. (4.9). The optimization problem eq. (4.8) can be solved through multiple optimization methods (such as Parker et al. (1989); Huang et al. (2012), just to name a few). Consider a problem in which a desired position $\boldsymbol{x}^*$ of the end-effector $\boldsymbol{x}$ of a robot should be reached, which is the objectives are formulated in the following sections of this thesis. One can define the tracking error $\boldsymbol{e}$ as:

$$\boldsymbol{e} = \boldsymbol{x}^* - \boldsymbol{x}. \tag{4.10}$$

Using a simple proportional controller with a gain vector $\boldsymbol{k}$, one can obtain a desired speed $\dot{\boldsymbol{x}}_d$:

$$\dot{\boldsymbol{x}}_d = \boldsymbol{k}\,\boldsymbol{e} = \boldsymbol{k}(\boldsymbol{x}^* - \boldsymbol{x}). \tag{4.11}$$

The eqs. (4.8) and (4.11) can be used to set the best solution of the point task in the least square sense as:

$$\min_{\dot{\boldsymbol{q}}} \|\boldsymbol{J}\dot{\boldsymbol{q}} - \boldsymbol{k}(\boldsymbol{x}^* - \boldsymbol{x})\|_2\,, \tag{4.12}$$

$$\min_{\dot{\boldsymbol{q}}} \|\boldsymbol{J}\dot{\boldsymbol{q}} - \dot{\boldsymbol{x}}_d\|_2\,. \tag{4.13}$$

Similarly to eq. (4.8), eq. (4.13) can be solved as an optimisation problem, via quadratic programming, given by:

$$\|\boldsymbol{J}\dot{\boldsymbol{q}} - \dot{\boldsymbol{x}}_d\|_2 = (\boldsymbol{J}\dot{\boldsymbol{q}} - \dot{\boldsymbol{x}}_d)^T(\boldsymbol{J}\dot{\boldsymbol{q}} - \dot{\boldsymbol{x}}_d) \tag{4.14}$$

$$= \dot{\boldsymbol{q}}^T\boldsymbol{J}^T\boldsymbol{J}\dot{\boldsymbol{q}} - \dot{\boldsymbol{x}}_d\boldsymbol{J}\dot{\boldsymbol{q}} - \dot{\boldsymbol{q}}^T\boldsymbol{J}^T\boldsymbol{J}\dot{\boldsymbol{x}}_d + \dot{\boldsymbol{x}}_d^T\dot{\boldsymbol{x}}_d \tag{4.15}$$

$$= \dot{\boldsymbol{q}}^T(\boldsymbol{J}^T\boldsymbol{J})\dot{\boldsymbol{q}} - 2(\dot{\boldsymbol{x}}_d^T\boldsymbol{J})\dot{\boldsymbol{q}} + \dot{\boldsymbol{x}}_d^T\dot{\boldsymbol{x}}_d \tag{4.16}$$

where $\dot{\boldsymbol{x}}_d^T\dot{\boldsymbol{x}}_d$ does not depend on $\dot{\boldsymbol{q}}$. Such that eq. (4.13) can be represented as:

$$\min_{\dot{\boldsymbol{q}}} \ \frac{1}{2}\dot{\boldsymbol{q}}^T\boldsymbol{H}\dot{\boldsymbol{q}} - \boldsymbol{c}^T\dot{\boldsymbol{q}} \tag{4.17}$$

with:

$$\boldsymbol{H} = \boldsymbol{J}^T\boldsymbol{J} \tag{4.18}$$

$$\boldsymbol{c} = \dot{\boldsymbol{x}}_d^T\boldsymbol{J} \tag{4.19}$$

This formulation is now in a canonical QP form with a Hessian matrix $\boldsymbol{H}$ and a constant vector $\boldsymbol{c}$, and can be solved by standard solvers. Both the Hessian matrix $\boldsymbol{H}$ and the constant vector $\boldsymbol{c}$ can be computed from the task Jacobian $\boldsymbol{J}$, the current robot position $\boldsymbol{x}$, the desired position $\boldsymbol{x}^*$, and the control gain $\boldsymbol{k}$.

If one wants to define several tasks in a weighted QP formulation, one can compute the Hessian matrix $\boldsymbol{H}_i$ and constant vector $\boldsymbol{c}_i$ for each of the $i \in [0, N]$ tasks, and then sum them with a weight γ_i to obtain the full control problem. The complete form of the

control problem is then:

$$\min_{\dot{\boldsymbol{q}}} \frac{1}{2}\dot{\boldsymbol{q}}^T \boldsymbol{H}\dot{\boldsymbol{q}} - \boldsymbol{c}^T\dot{\boldsymbol{q}} \tag{4.20}$$

$$s.t.\ \boldsymbol{A}\dot{\boldsymbol{q}} < \boldsymbol{b} \tag{4.21}$$

$$\boldsymbol{G}\dot{\boldsymbol{q}} = \boldsymbol{b}_{eq} \tag{4.22}$$

$$\boldsymbol{l}_b < \dot{\boldsymbol{q}} < \boldsymbol{u}_b \tag{4.23}$$

where $\boldsymbol{H} = \sum_{i=1}^{N} \gamma_i \boldsymbol{H_i}$ and $\boldsymbol{c} = \sum_{i=1}^{N} \gamma_i \boldsymbol{c_i}$. $\boldsymbol{A}$ and $\boldsymbol{G}$ are matrices defining inequality and equality constraints for $\dot{\boldsymbol{q}}$, while $\boldsymbol{l}_b$ and $\boldsymbol{u}_b$ define bound constraints. Similarly to the single task problem from Equation (4.17), this problem can be solved by standard QP solvers. In the following, translation between clinical objective form section 4.2 into task objectives and constraints in the above-detailed weighted QP framework is described, such that it can be applied in real-time for allowing assisted telemanipulation.

4.5 Semi-autonomous arm-body control approach

The proposed approach involves triggering autonomous motions of the tip positions of the arms and body to provide coordinated control. Using the modelling presented previously (section 3.3), an overview of the proposed approach is presented in section 4.5.1. The problem set as control objectives is described in section 4.5.2. An optimization-based formulation of the problem is presented in section 4.5.3 covering the objectives and constraints.

4.5.1 Overview

To best describe the proposed semi-autonomous arm-body approach, first, consider the situation where no body movement is required. In this case, the user telemanipulates the dominant and non-dominant arms as in standard telemanipulation following the original telemanipulation formulation defined in eq. (3.23), such that:

$${}_e\dot{\boldsymbol{p}}^r = {}_e\boldsymbol{J}_r\, \boldsymbol{K}_r\, \dot{\boldsymbol{q}}_r \tag{4.24}$$

$${}_e\dot{\boldsymbol{p}}^l = {}_e\boldsymbol{J}_l\, \boldsymbol{K}_l\, \dot{\boldsymbol{q}}_l \tag{4.25}$$

with ${}_w\dot{\boldsymbol{p}}^e$ set to 0. In the case of eqs. (4.24) and (4.25), the system –the arms– is then governed by the leader console variables ($\dot{\boldsymbol{q}}_r$ and $\dot{\boldsymbol{q}}_l$) acting on the positions of the arms in the body frame $\{e\}$ while the body is static. For manual and semi-autonomous arm-body modes, controlling the arms is analogous, the user is set to control 6 DoF, thus 3 DoF of each arm.

When body movement is required, the user activates the semi-autonomous arm-body control mode, invalidating the prior relationship: eqs. (4.24) and (4.25).

The semi-autonomous arm-body control is activated by the user pressing the C_1 pedal (see fig. 4.5). The eqs. (4.24) and (4.25) depends on the semi-autonomous arm-body control with independent behaviour on each of the arms. When the C_1 pedal is active, the user moves the body by the speed of the handle which is mapped from the dominant arm (see fig. 4.5). The non-dominant arm has two behaviours depending on the pressing of the C_2 pedal. When the C_2 pedal is not pressed, the non-dominant arm moves with the body, in other words, the body that is indirectly controlled by the dominant arm. When the C_2 pedal is pressed the position of the non-dominant arm position is kept, effectively enabling traction. At the moment in which each of the pedals is de-activated, the behaviour attached to each one is terminated, returning to the relationship previously described in eqs. (4.24) and (4.25).

To control the body by the dominant-arm movements during the semi-autonomous arm-body control the concept of a *virtual* arm is introduced. The virtual arm detaches direct telemanipulation by the user, meaning the movements from the user in the leader console handle do not have a direct effect on the arm but are used for control of the body, such that eq. (4.24) is nullified, and the arm moves according to the semi-autonomous arm-body control mode. Two dominant arms are considered, the real and the virtual. Positions of the real and virtual dominant arms are represented by ${}_e\boldsymbol{p}^r$ and ${}_e\boldsymbol{\rho}^r$ respectively. ${}_e\boldsymbol{p}^r$ is defined by eq. (3.17) and ${}_e\boldsymbol{\rho}^r$ defined by the positional argument $pos(\cdot)$ of the velocity of the virtual arm as:

$$ {}_e\boldsymbol{\rho}^r = \left(pos({}_e\dot{\boldsymbol{\rho}}^r)\right)_t = {}_w\boldsymbol{J}_e\boldsymbol{K}_r\dot{\boldsymbol{q}}_r. \tag{4.26} $$

Pressing transfers the difference between the current position of the arm and the position of the virtual arm, under the assumption that ${}_e\boldsymbol{\rho}^r_i - {}_e\boldsymbol{\rho}^r_0 \neq 0$ for a given t time after

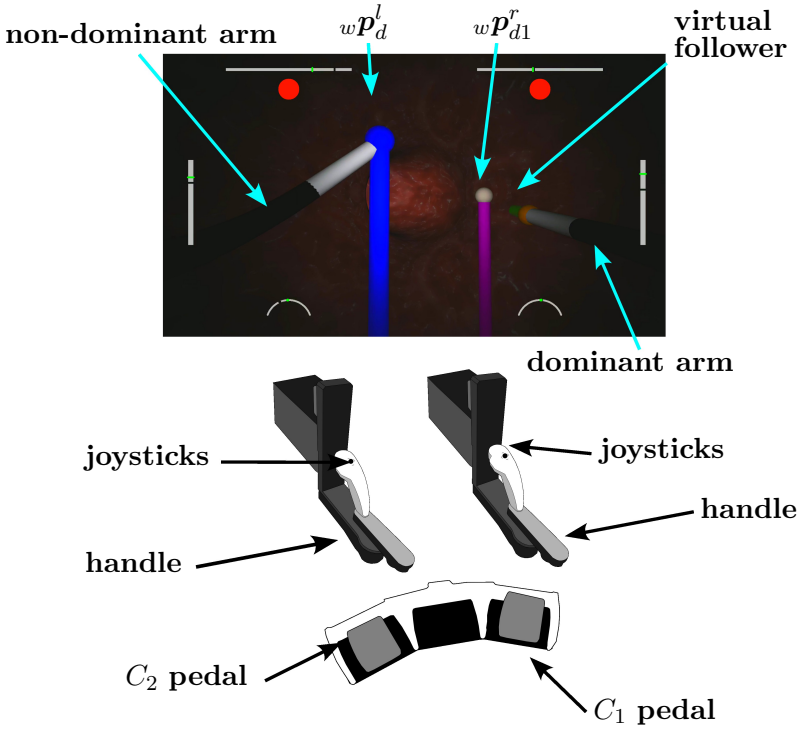


FIGURE 4.5: Overview of the main components of the simulator setup. Reaching targets are depicted in blue and pink for the non-dominant and dominant arms. Follower in black with a grey tip, virtual follower in semi-transparent green.

the initial triggering of the body movement at time 0.

$$\dot{\boldsymbol{q}}_v = \boldsymbol{J}_e^{\dagger}\big(\boldsymbol{K}_e({}_e\boldsymbol{\not{p}}^r - {}_e\boldsymbol{p}^r)\big) \tag{4.27}$$

where k_e is a scalar gain to match manual telemanipulation speeds and $\boldsymbol{J}_e^{\dagger}$ is the Moore-Penrose pseudo inverse of the estimated Jacobian ${}_w\boldsymbol{J}_e$. The endoscope body has 4 DoFs, such that there is a redundancy for the displacement of the body. The virtual arm is shown in semi-transparent green in fig. 4.5 and effectively controls the body movements.

As described in section 4.2, moving the body impacts the position of the tip of the arms with respect to the surgical site which can be detrimental to the surgical procedure. To compensate for this, the three objectives are based on the previous problem statement. First, the dominant arm should not move from its world position $\{w\}$ during body movements, such that ${}_w\dot{\boldsymbol{p}}^r = 0$. In other words, the tip dominant arm position of the leader ${}_e\boldsymbol{\not{p}}^r$ moves the body, while the follower dominant arm ${}_e\boldsymbol{p}^r$ and body ${}_w\boldsymbol{p}^e$ compensate, to keep the world position ${}_w\boldsymbol{p}^r$ still during body movement avoiding interfering with the surgical task. The user can use the semi-autonomous arm-body mode to move the body which does not modify the global positions. Static position with respect to the surgical

site allows the user to continue performing the gestures that were being performed before the body movement.

Second, the non-dominant arm behaviour during body movements depends on the surgical step. In this work, the asumption that the user is performing a dissection task is made, as described in section 4.2. In this task, the user grasps the tissue with the non-dominant arm and lifts it to a desired position ${}_w\boldsymbol{p}^l_d$. In the STRAS system, the C_2 pedal is used to activate the grasping tool of the non-dominant arm. Therefore, if C_2 is not activated, the non-dominant arm can move together with the body since it is not interacting with tissue, reducing constraints on the control system. If pedal C_2 is active a tissue is being grabbed, following the behaviour on the real STRAS system. It is then paramount to maintain the position of the non-dominant arm during body movements, i.e. keeping ${}_w\dot{\boldsymbol{p}}^l = 0$, to keep traction over the grasped tissue. This task is particularly challenging to perform manually as the user will be required to consider 10 DoF —6 DoF of the arms and 4 DoF of the body. This task is carried out automatically by the semi-autonomous arm-body control mode during body movements.

Finally, safety is also considered. The semi-autonomous arm-body controller should not result in the arms being placed in an unsafe position, i.e. outside of the field of view, too close to the body (which holds the camera), or close to joint limits. This comes to similar approaches described previously, sections 2.4.2 and 2.4.3, where the arms are constrained to safe –visible– regions at the current camera –body– position. Similarly to laparoscopes, the arms should not be too far to lose depth perception from the arms neither too close to occlude the view of the surgical site.

Given the described behaviours, the semi-autonomous approach allows to: *i)* have free control of the arms when no body movements are required, *ii)* control the body with the dominant arm, so as to re-position the endoscope with respect to the surgical site –by anchoring the dominant arm position in the world frame, and *iii)* correctly perform the most demanding surgical task –dissection– in which the position of the arms end-effectors should be maintained while the body is moved, to efficiently reach subsequent cutting locations. This three-way behaviour allows the semi-autonomous arm-body control mode to assist the user not only in the most demanding stage of the ESD procedure, which is dissection but also during the other stages while allowing safe movements inside the patients body. During dissection, the proposed approach could also be used to move

the body while marking the tissue and simplifying the task by maintaining the world position of the dominant arm while accounting for safety constraints.

4.5.2 Formulation of the control problems as objectives

The previously described problems (section 4.2) are now presented as high-level objectives. Keeping the absolute position of the dominant arm still, or ${}_w\dot{\boldsymbol{p}}^r = 0$ is expressed by:

$$\mathscr{H}_1 \colon \min \|{}_w\boldsymbol{p}^r_0 - {}_w\boldsymbol{p}^r\| \tag{4.28}$$

where the subscript ${}_0$ refers to the position at the beginning of the body movement (i.e. when the pedal C_1 is pressed). Similarly,

$$\mathscr{H}_2 \colon \min \left\|{}_w\boldsymbol{p}^l_0 - {}_w\boldsymbol{p}^l\right\| \tag{4.29}$$

for when the absolute position of the non-dominant arm is required to be maintained —during dissection after grabbing the tissue. The objective $\mathscr{H}_2$ is active only when the C_2 pedal is active, that is when the non-dominant arm is grabbing –keeping traction over– the tissue.

Secondly, the dominant arm is set to be in a safe position w.r.t. the body frame ${}_e\boldsymbol{p}^r_d$. The translational component $\boldsymbol{q_r}|_{trans}$ of the configuration variables of the dominant arm in eq. (4.28), defines a safe position as:

$$\lambda^-_{trans} \leq \boldsymbol{q_r}|_{trans} \leq \lambda^+_{trans} \tag{4.30}$$

which effectively repels joint limits. Where λ^-_{trans} is the lower bound and λ^+_{trans} is the upper bound. The position of the dominant arm in the image is given by its projection, $img(\cdot)$. The arm defines a safe position as one in which the tip of the dominant arm is visible, in the horizontal component of the image $img({}_e\boldsymbol{p}^r)|_x$, is given by:

$$\lambda^-_x \leq img({}_e\boldsymbol{p}^r)|_x \leq \lambda^+_x \tag{4.31}$$

and in its vertical component $img({}_e\boldsymbol{p}^r)|_y$ as

$$\lambda^-_y \leq img({}_e\boldsymbol{p}^r)|_y \leq \lambda^+_y. \tag{4.32}$$

When the dominant arm leaves the previously defined safe zone in the image or is close to the translational bounds, the arm is returned inside the safe bounds while the body and arms are being controlled by two objectives. A simplified overview of the safe zone in which the dominant arm should be kept is depicted in fig. 4.6. While the body is moving, the dominant arm avoids leaving the safe zone, by eqs. (4.30) to (4.32). Maintaining the arm in a safe position is obtained by setting the following objective:

$$\mathscr{H}_3 \colon \min \| {}_e\boldsymbol{p}_d^r - {}_e\boldsymbol{p}^r \| \tag{4.33}$$

in the Cartesian space, where the goal position ${}_e\boldsymbol{p}_d^r$ is the centroid of the workspace. The arm should also remain inside the field of view of the endoscopic camera, which can be obtained by setting the following objective:

$$\mathscr{H}_4 \colon \min \| img({}_e\boldsymbol{p}_d^r) - img({}_e\boldsymbol{p}^r) \| \tag{4.34}$$

with $img({}_e\boldsymbol{p}_d^r)$ being the image centre. When the arm is in a safe position the objectives $\mathscr{H}_3$, $\mathscr{H}_4$ are deactivated.

The objectives defined previously, eqs. (4.28), (4.29), (4.33) and (4.34), are activated depending on the states of the pedals. When no pedal is active, the body is static and the user has full range control of the arms by moving the leader console handles. Pressing C_1 activates $\mathscr{H}_1$, and if the dominant arm is in an unsafe position $\mathscr{H}_3$ and $\mathscr{H}_4$ become active such that the dominant arm keeps its absolute position, away from the joint limits and in a visible location, while the body moves. When C_2 is pressed, $\mathscr{H}_2$ is active.

4.5.3 Optimization-Based High-Level Control

The body movement eq. (4.27) commanded for the user by the dominant arm handle requires to be compensated in order to maintain the optimization objectives eqs. (4.28), (4.29), (4.33) and (4.34). The objectives are expressed in quadratic form eq. (4.17) such that each one can be solved to compute the desired velocity that reduces in the least square sense eq. (4.13) the difference to the desired position. The different positions to

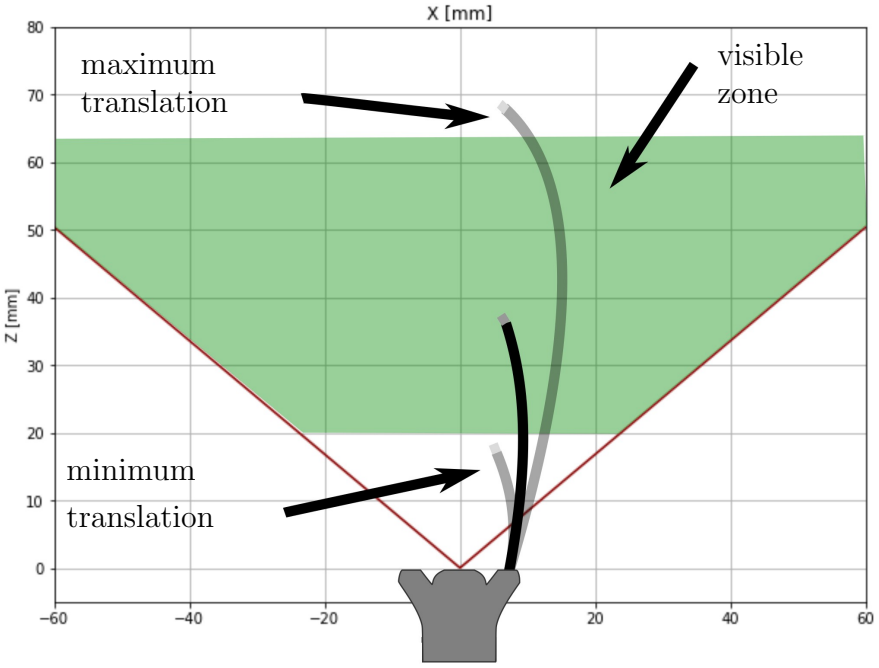


FIGURE 4.6: Overview of the defined safe zones for the dominant arm. The green area is a simplification of the safe zone. This safe zone is visible and is within the desired translation range.

be controlled, are stacked together such that:

$$\boldsymbol{p} = \begin{bmatrix} {}_{e}\boldsymbol{p}^{r} \\ {}_{w}\boldsymbol{p}^{r} \\ {}_{w}\boldsymbol{p}^{l} \end{bmatrix}, \tag{4.35}$$

with the set of configuration variables defined as:

$$\dot{\boldsymbol{q}}_{qp} = \begin{bmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{r}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{e}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{l}} \end{bmatrix} \tag{4.36}$$

and the Jacobian is defined by the appropriate individual Jacobians of each of the positions defined in eq. (4.35). The Jacobians of the dominant arm in the body frame eq. (3.16), the right frame in the world frame eq. (3.18) and the non-dominant arm in the world frame eq. (3.20) are stacked together to match the configuration variables

eq. (4.35), such that[3]:

$$\boldsymbol{J} = \begin{bmatrix} {}_e\boldsymbol{J}_r & 0 & 0 \\ {}_e\boldsymbol{J}_r & {}_w\boldsymbol{J}_e & 0 \\ 0 & {}_w\boldsymbol{J}_e & {}_e\boldsymbol{J}_l \end{bmatrix} \tag{4.37}$$

where the last two components of the first row is filled with zeros due to the lack of relationships between the dominant arm position and the configuration variables of the endoscope and the non-dominant arm. Similarly, for the last element of the middle row for the dominant arm world position and the non-dominant arm configuration variables. Furthermore, in the first component of the last row, there is no link between the dominant arm configuration variables and the non-dominant arm position in the world frame.

The task Jacobian eq. (4.37), the positions to be controlled eq. (4.35) and the configuration variables eq. (4.36) can be combined with the initial body movement commanded by the user eq. (4.27), as:

$$\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}}_{qp} + \begin{bmatrix} \mathbf{0}_{1\times 3} \\ \mathbf{0}_{1\times 3} \\ \dot{\boldsymbol{q}}_v \end{bmatrix}, \tag{4.38}$$

where $\dot{\boldsymbol{q}}_{qp}$ comes from the solution of the quadratic form of each of the goals in $i : \{\mathscr{H}_1, \mathscr{H}_2, \mathscr{H}_3, \mathscr{H}_4\}$ and the subsequent vector is filled from the eq. (4.27). The quadratic form (eqs. (4.20), (4.21) and (4.23)) is then modified according to each one of the objectives present at a given time t as:

$$\min_{\dot{\boldsymbol{q}}} \quad \frac{1}{2}\,\dot{\boldsymbol{q}}^T \boldsymbol{H}_t \dot{\boldsymbol{q}} + \boldsymbol{c}_t^T \dot{\boldsymbol{q}} \tag{4.39}$$

$$s.t. \quad \dot{\boldsymbol{q}}_t^+ \leq \dot{\boldsymbol{q}} \leq \dot{\boldsymbol{q}}_t^- \tag{4.40}$$

where $\boldsymbol{H}_t = \sum \gamma_i \boldsymbol{H}_i$ and $\boldsymbol{c}_t = \sum \gamma_i \boldsymbol{c}_i$ defined accordingly to eqs. (4.18) and (4.19) for each of the i previously defined objectives (eqs. (4.28), (4.29), (4.33) and (4.34)), and γ_i is a given objective weight. At each time t the desired positions are used to obtain the velocities to be applied given the active objectives. The obtained vector $\dot{\boldsymbol{q}}_{qp}$ is obtained by eq. (4.39) and bounded by eq. (4.40). The optimization variables $\boldsymbol{q}$ should stay within

[3]The rotation of the endoscope body is an unusual and possibly disturbing motion for the user because of the resulting rotation of the camera, it is here not actuated, having only the positional Jacobian and not the rotational part.

the bounds $[{}_i\boldsymbol{q}_{min}, {}_i\boldsymbol{q}_{max}]$. This is achieved by introducing lower and upper bounds:

$$\dot{\boldsymbol{q}}_{-}^{\diamond} = \delta({}_i\boldsymbol{q}_{min} - \boldsymbol{q}) \tag{4.41}$$

$$\dot{\boldsymbol{q}}_{+}^{\diamond} = \delta({}_i\boldsymbol{q}_{max} - \boldsymbol{q}) \tag{4.42}$$

depending on the distance of $\boldsymbol{q}$ to ${}_i\boldsymbol{q}_{min}$ and ${}_i\boldsymbol{q}_{max}$, with δ being the convergence factor (set to 0.1) as introduced in (Faverjon and Tournassoud, 1987). Equation (4.40) uses the minimum value for the lower bound between the low-level controller speed limits $\dot{\boldsymbol{q}}_{min}^{h}$ and the joint limits $\dot{\boldsymbol{q}}_{-}^{\diamond}$ presented in eq. (4.41), such that the minimum speed is given by $\dot{\boldsymbol{q}}_{t}^{+} = \min(\dot{\boldsymbol{q}}_{-}^{\diamond}, \dot{\boldsymbol{q}}_{min}^{h})$. The upper bound is obtained similarly as $\dot{\boldsymbol{q}}_{t}^{-} = \min(\dot{\boldsymbol{q}}_{+}^{\diamond}, \dot{\boldsymbol{q}}_{max}^{h})$.

The Jacobian eq. (4.37) linking $\dot{\boldsymbol{q}}_{qp}$ to $\dot{\boldsymbol{p}}$ is therefore populated differently depending on the active objectives at a given time, yielding the corresponding $\boldsymbol{J}_i$: objective $\mathscr{H}_1$ acts on the dominant arm ($\dot{\boldsymbol{q}}_{\boldsymbol{r}}$) and the body ($\dot{\boldsymbol{q}}_{\boldsymbol{e}}$), $\mathscr{H}_3$ and $\mathscr{H}_4$ act only on the dominant arm ($\dot{\boldsymbol{q}}_{\boldsymbol{r}}$) such that, eq. (4.37) is populated until on the last three rows. When C_2 is pressed and $\mathscr{H}_2$ is active, the Jacobian $\boldsymbol{J}$ is populated completely since $\mathscr{H}_2$ acts on $\dot{\boldsymbol{q}}_{\boldsymbol{l}}$.

4.5.4 Summary of behaviours

The semi-autonomous control mode has three main use cases. In case 0, the surgeon freely telemanipulates the arms with no body movements in the system, pedals C_1 and C_2 are inactive. Case 1 is activated by constant pressing of the C_1 pedal to activate the semi-autonomous control mode, with no activation of the C_2 pedal. During case 1, the movement of the physical handle of the dominant arm in the leader console is translated to the body on the follower side. During this movement, the position of the dominant arm is softly anchored (it is kept constant with low priority), meaning that it is being kept at best unless body movements become too large, in which case the dominant arm moves with the body. If the user presses once on C_2, case 2 is active. In this mode, if C_1 is active –constantly pressed– the user controls body movement in the same way as in case 1. The difference is here that the position of the non-dominant arm is hard anchored (i.e. it is kept constant with high priority).

The three cases described above are following the surgical scenario of ESD. Case 0 will be used in normal teleoperation mode when no body movement is required. Therefore, case 0 could be used in all sub-steps of ESD and other surgical procedures. Case 1 would

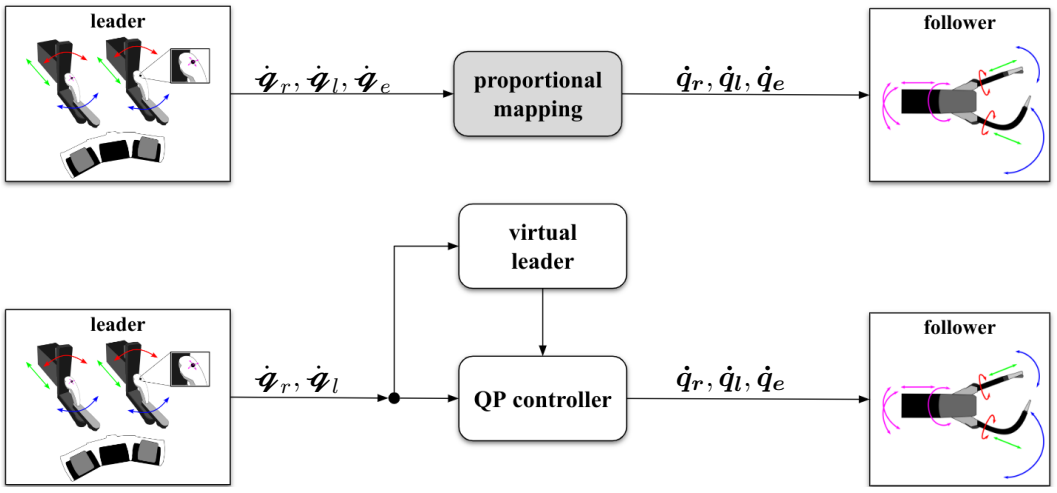


FIGURE 4.7: Overview of the control framework. On top manual telemanipulation (eqs. (3.23) and (3.25)), on the bottom assisted telemanipulation by the semi-autonomous arm-body control mode. When there is no body movement in the semi-autonomous arm-body control mode, case 0, the arms $\dot{q}_r$, $\dot{q}_l$ are governed by the leader console $\dot{q}_r$, $\dot{q}_l$. When cases 1 and 2 are active, the whole platform is controlled by the assisted mode, moving the body indirectly by the arm movements, as per eq. (4.26). In case 1, the non-dominant arm is not anchored, thus having $\dot{q}_r, \dot{q}_e$ as the output of the QP controller. In case 2, both arms will be controlled, anchoring the non-dominant arm as desired by the traction during the dissection stage of ESD while also safely controlling the position of the dominant arm.

be used to reposition the body and arm, without grasping any tissue, especially required during the marking, dissection and retraction part. Case 1 can also be used during multiple surgical procedures that require small re-positioning movements in intricate surgical environments. The soft anchoring of the dominant arm is here used to allow swiftly going back to the surgical task in case of a small body movement, keeping the tip of the dominant arm in a constant relative position with respect to the tissues. If a large body movement is performed, the dominant arm would move along with the body. In case 2, the non-dominant arm is anchored in the world frame, with soft-anchoring on the dominant arm in favor of body movement and safe arm pose. Case 2 would be mostly required during the dissection stage of ESD, and could also be used during suturing and needle passing, where the non-dominant arm is anchored while the dominant arm executes another task. Table 4.1 summarises the expected behaviour based on the pedals active.

In table 4.1 a summary of the behaviours used to control the system during assisted telemanipulation is presented. The behaviour presented follows the convention of a non-dominant arm that executes secondary tasks by a grasper, while the primary arm executes critical tasks by the electric knife. Note that during the retraction phase and

TABLE 4.1: Table of the behaviours of the assisted telemanipulation control mode.

case	description	C_1 pedal	C_2 pedal	objectives active	surgical use case
0	free arm telemanipulation, no body movement	inactive	inactive	-	marking, dissecting
1	body moves and arms compensate, dominant arm anchored	active	inactive	$\mathscr{H}_1$, if required: $\mathscr{H}_3$ $\mathscr{H}_4$	marking, dissecting
2	body moves, arm compensate, arms are anchored	active	active	$\mathscr{H}_1$,$\mathscr{H}_2$, if required: $\mathscr{H}_3$ $\mathscr{H}_4$	dissecting

the injection phase of ESD, arm movement is null in the current FSIR, as the injection is performed by the needle inserted in the centre channel and during retraction a grasper is mounted in the centre channel to ease retrieval from the digestive system.

The table 4.1 is complemented by the diagram of the semi-autonomous arm-body control mode presented in fig. 4.7 which represents the 3 cases. Case 0 is represented by the grey box on the top centre, in which the arms are proportionally controlled by the leader console, following the expression eq. (3.23). This behaviour will be analogous to manual telemanipulation of the arms. The white boxes represent cases 1 and 2. In case 1, the non-dominant arm will not be controlled, thus having $\dot{\boldsymbol{q}}_{\boldsymbol{r}}, \dot{\boldsymbol{q}}_{\boldsymbol{e}}$ as the output of the QP controller instead of the configuration variables of the three subsystems. In case 2, both arms will be controlled, anchoring the non-dominant arm as desired by the traction during the dissection stage of ESD while moving the dominant arm in a safe position while moving w.r.t. the surgical site. In case 2 the output of the QP controller is as depicted in the fig. 4.7.

4.6 Experiments

The proposed semi-autonomous arm-body control mode is tested in a simulated environment as to avoid the perturbations induced by non-linearities of the real STRAS system (Poignonec et al., 2020). The proposed semi-autonomous arm-body control mode is compared to manual telemanipulation in a virtual simulator (see fig. 4.5). The experimental setup used to compare the two control modes is presented and the experimental task is

described. Then, the protocol and metrics are formalised which allows the overview of the results.

4.6.1 Experimental setup

The physical leader side of the STRAS system is paired with a simulator that emulates the main features of the real surgical task. The simulated environment is built in Python using VTK (Schroeder et al., 2006) and the physical leader values are simulated based on the theoretical model of the STRAS system section 3.3.3. The kinematics of the follower implement eqs. (3.9), (3.10) and (3.12), which are then used to render a virtual scene using VTK. The target positions ${}_w\boldsymbol{p}^l_d$, ${}_w\boldsymbol{p}^r_{d1}$ and ${}_w\boldsymbol{p}^r_{d2}$ are displayed as white spheres with a radius of 3mm to simulate the dissection task, as described in section 4.2, and the colon environment is simulated as a textured cylinder around a curved axis. Note that no collisions are implemented in the simulation.

During the user experiments, a Python script reads the leader side positions of the STRAS system using sockets through a wired Ethernet connection. The Python script creates the VTK environment that is presented to the user on a screen on top of the leader side, as when the user telemanipulates the system with the real follower (fig. 4.8). The weights for the objectives γ are set as in table 4.2. When the arm is in an unsafe position the values of γ_3 and γ_4 are increased to 0.15; and γ_1,γ_2 are decreased to 0.4 and 0.3 in order to return the arm into the safe zone, values of $k_{1,2,3,4}$ are unchanged as priority is given by the γ values. Safe image bounds in the image $\lambda_x^{+,-}$ and $\lambda_y^{+,-}$ are set by the normalized coordinates of the screen from 0.1 to 0.9 in both coordinated axes, with a 1920x1080 px resolution. The bounds $\lambda_{trans}^{-,+}$ are set within the 0-72mm range of motion for the arms.

The parameters for the gammas and the gains were selected to better suit continuous telemanipulation and stable control of the body. The parameters allow the user to move the body, while the approach appropriately reduces the difference in the body and world frame. The proper selection of the gammas and gains avoids over- or under-shooting with respect to the body movement so that the user perceives this movement as smooth and steady. Otherwise, the user will compensate for this by the opposite action which might result in longer overall movements.

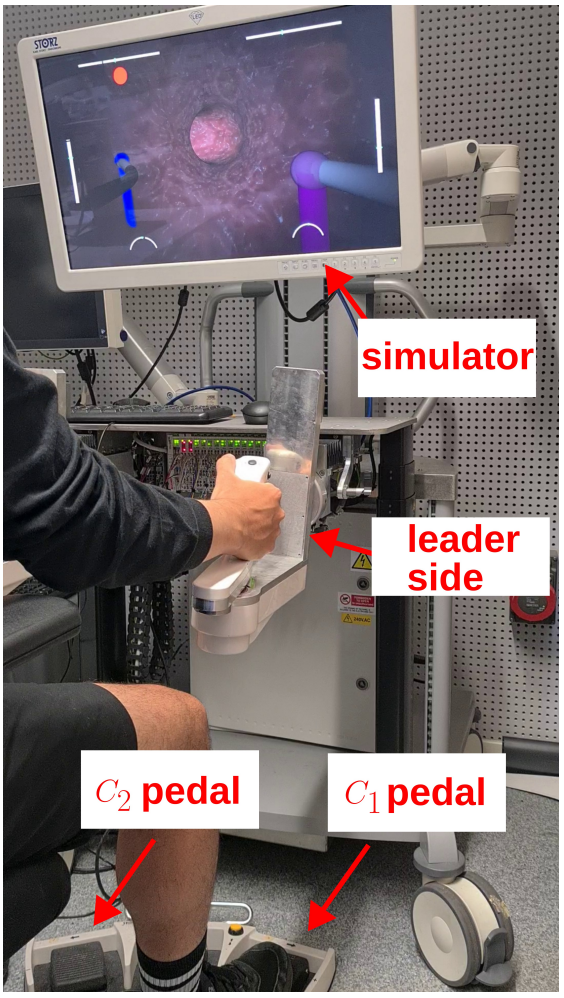


FIGURE 4.8: Experimental setup used. Simulated follower visualization through the VTK environment using the physical leader side of the STRAS robotic system.

TABLE 4.2: Parameters used in the experimental setup. Values of the weights γ and the proportional gains k for each objective active during semi-autonomous control mode.

Variable	γ_1	γ_2	γ_3	γ_4	$k_{1,2,3,4}$
Value	0.5	0.4	0.05	0.05	0.25

4.6.2 Experimental task

During the experiments, a trial consists of the user sequentially reaching the three targets ${}_w\boldsymbol{p}_d^l$, ${}_w\boldsymbol{p}_{d1}^r$ and ${}_w\boldsymbol{p}_{d2}^r$. The position of the targets emulates the dissection stage of ESD, as described in section 4.2.

Each trial starts from a given reference position, then the user reaches the first target ${}_w\boldsymbol{p}_d^l$ with the non-dominant arm. Reaching this first target is referred to as Segment I from now on. Segment I simulates reaching a tissue and grabbing it which is present during the reaching and marking of the stages of ESD (section 4.1) and also the reaching

step during the dissection in ESD. Segment I is completed when the user keeps the non-dominant arm tip within a ball of radius 3 mm centred on ${}_w\boldsymbol{p}_d^l$ for Segment. Deviating too far from ${}_w\boldsymbol{p}_d^l$ resets the timer to 0 and the user is required to telemanipulate the arm to reach again.

The segment II of the experiment simulates the cutting during the dissection stage of ESD. The user is required to reach ${}_w\boldsymbol{p}_{d1}^r$ and ${}_w\boldsymbol{p}_{d2}^r$ with the dominant arm while staying on ${}_w\boldsymbol{p}_d^l$ with the non-dominant arm by a threshold of 2s. Similarly to Segment I, a 2s threshold is used to validate a target being reached for the two objectives of the dominant arm. In practice, ${}_w\boldsymbol{p}_{d1}^r$ is displayed when Segment I is finished, and ${}_w\boldsymbol{p}_{d2}^r$ is displayed only when ${}_w\boldsymbol{p}_{d1}^r$ is validated. If the user deviates significantly (more than 5mm) from ${}_w\boldsymbol{p}_d^l$, the user will need to reach ${}_w\boldsymbol{p}_d^l$ again and stay for two seconds before being able to validate the current target (${}_w\boldsymbol{p}_{d1}^r$ or ${}_w\boldsymbol{p}_{d2}^r$).

Once the two segments have been completed, the two targets for the dominant arm have been reached without leaving the non-dominant arm from its target; the trial is set as finished. New positions of the three targets are then randomly drawn to start a new trial. The positions are uniformly randomly generated such that reaching target ${}_w\boldsymbol{p}_d^l$ by the dominant arm and reaching ${}_w\boldsymbol{p}_{d1}^r$ by the non-dominant arm requires moving the body. Reaching target ${}_w\boldsymbol{p}_{d1}^r$ could be within reach of the dominant arm or require moving the body, depending on the current body position. The positions mimic the intricacy of performing real dissection during an ESD.

4.6.3 Protocol

The experimental protocol is set to last between 30 and 40 minutes. During this time, the user is first introduced to the platform, task, and control modes for 5-10 minutes. Then, the user performs a 15-minute training, during which the user receives instructions regarding the manual teleoperation and semi-autonomous control mode, as well as specific instructions to perform the task correctly. The user receives an explanation of how the task is built and how to correctly perform the task to avoid unnecessary/erroneous movements, avoiding abrupt changes of direction and/or speed that could be detrimental to the patient in the real surgical scenario. Furthermore, the user is instructed to avoid placing the tools out of sight or too close to the camera. These instructions apply for both control modes, although they are expected to be harder to perform with

manual telemanipulation. The user then performs as many trials as possible during the 15-minute training time, with the possibility to ask questions to the investigator.

During the experiments and the training, the user alternates 5 trials with the semi-autonomous arm-body control mode and 5 trials with the manual telemanipulation mode (using the same target positions in both) until the time (10 minutes) runs out, without any interference from the investigator.

After the experiment is finished, the user fills out a NASA TLX (Hart, 2006) questionnaire. The NASA TLX questionnaire evaluates the subjective opinion of users for both control modes. Afterwards, the user is asked for a general opinion of both control modes and is allowed to give feedback for further improvement of the setup or the proposed control mode.

4.6.4 Quantitative evaluation metrics

Three categories of performance metrics are considered: time, kinematics and surgical task. Metrics are adapted from similar approaches in laparoscopic surgery and other metrics for flexible endoscopy (Da Col et al., 2020; Jarc and Curet, 2017), considering the architecture of the FSIR platforms, especially the STRAS system. To describe the different kinematics under the different segments, the starting time at each trial is t_0, and the start of segment II (end of segment I) is depicted as t_1. The end of the trial is set as t_2.

4.6.4.1 Duration metrics

Time performance metrics measure the raw duration of Segment I t_{sI}, of Segment II t_{sII}, and the complete trial t_s. Time performance metrics can be of interest when validating how easy is the task to perform in comparison to the time spent learning or mastering the control interface or method. In the case of the study, the duration metrics show how easily the user can master the semi-autonomous arm-body control mode and have on-par performance with manual telemanipulation.

4.6.4.2 Kinematics metrics

Kinematic metrics measure how the STRAS system is controlled. For the kinematics category, the first set of metrics relates to body movements. A body movement starts when the speed of the body $\|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\|$ goes from 0 to a positive value and ends when $\|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\| < \epsilon$, with ϵ denoting a small body motion threshold. a_b measures the number of independent movements of the body per trial,

$$a_b = \sum_{t_0}^{t_2} \{\|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\| > 0, \|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\| < \epsilon\}. \tag{4.43}$$

Then,

$$m_{b,sI} = \int_{t_0}^{t_1} t \; dt \quad \text{for} \begin{cases} \text{if } \|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\| > 0 \wedge \|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\| < \epsilon \\ \text{else } 0 \end{cases} \tag{4.44}$$

measures the total duration of body movements in Segment I (respectively $m_{b,sII}$ for Segment II) and m_b during the complete trial. The total travelled distance by the body for the complete trial is measured by c_b.

$$c_b = \int_{t_0}^{t_2} \|_w\dot{\boldsymbol{p}}^e\| \, dt \tag{4.45}$$

Finally,

$$f_b = a_b / m_b \tag{4.46}$$

is the frequency of body movements (i.e. number of body movements per second) during a given trial. i_b is the amount of time the body is idle

$$i_b = \int_0^{t2} t \; dt \quad \text{for} \begin{cases} \text{if } \|\dot{\boldsymbol{q}}_{\boldsymbol{e}}\| < \epsilon \\ \text{else } 0 \end{cases} \tag{4.47}$$

Similarly to eq. (4.47), c_l and f_l measure the total distance travelled and frequency of movements of the non-dominant arm. Such metrics are not computed for the dominant arm, because the triggering of body movements during semi-autonomous arm-body control mode would interfere and not provide meaningful output.

A second set of kinematic metrics relates to the smoothness of the movements. The spectral arc length $SPARC$ is used to quantify the smoothness (Balasubramanian et al.,

2015) of the movements, which was shown to be correlated with surgical skill and smooth operation (Estrada et al., 2016) and was used in several studies with similar design (Caravaca-Mora et al., 2023).

$$SPARC \triangleq -\sum_{k=1}^{K_c-1} \sqrt{\left(\frac{1}{K_c-1}\right)^2 + \left(\Delta \hat{V}[k]\right)^2} \tag{4.48}$$

The smoothness measured by eq. (4.48) is a modified version of the spectral arc length initially defined in Balasubramanian et al. (2011) for given movement speed profile $v(t)$ and considers a given threshold $\hat{V}$, with an upper-bound w_c^{max} for the cut-off frequency. Calculating the smoothness requires first segmenting the speed profile, and then computing the Fourier magnitude spectrum which is then normalized.

Smoothness is computed, eq. (4.48), w_e of the body (resp. $w_{l,sII}$ of the non-dominant arm) during Segment II. Note here that the numbers computed are the opposite of the number computed classically and defined in Balasubramanian et al. (2015), negating the scalar obtained in eq. (4.48). This operation is done to obtain a positive number which increases as the smoothness diminishes (i.e. the lower the better, as for most of the other metrics, aside from the i_b and a_b). Again, metrics relating to the dominant arm are not considered for the reasons exposed above.

Overall the kinematic metrics measure how efficiently and how well the movements of the arms or body are carried out during manual telemanipulation or the semi-autonomous arm-body control mode.

4.6.4.3 Surgical performance metrics

The surgical performance metrics measure how well the surgical task was performed. The distance between the non-dominant arm and its objective ${}_w\boldsymbol{p}_d^l$ at a given time t is given by:

$$\boldsymbol{d}_l\,(t) = \left\| {}_w\boldsymbol{p}_d^l - {}_w\,\boldsymbol{p}^l(t) \right\|_2 \tag{4.49}$$

and should be maintained to the minimum during Segment II. Since deviating too much from ${}_w\boldsymbol{p}_d^l$ will require performing extra surgical gestures, the maximum recorded distance

for each trial is of interest. Maximum recorded distance is set as:

$$d_{max} = \max_{\{t_1,t_2\}} \boldsymbol{d}_l(t), \tag{4.50}$$

where t_1 and t_2 are the starting and ending times of Segment II for a given trial, respectively. t_v measures the time extent during which the non-dominant arm is deviating from ${}_w\boldsymbol{p}_d^l$ during Segment II:

$$t_v = \int_{t_1}^{t_2} t \; dt \quad \text{for} \begin{cases} \text{if } \boldsymbol{d}_l\,(t) > d_{thr} \\ \text{else } 0 \end{cases} \tag{4.51}$$

where d_{thr} is a distance threshold set to 5 mm. Finally, the measurement of the time during which the arms are outside of the field of view during a trial by $t_{r,o}$ and $t_{l,o}$ for the dominant and non-dominant arms respectively.

The surgical performance metrics are the most important type of metrics as they allow to corroborate that the surgical task is being performed correctly. Erroneous performance of the task can be detrimental to the patients and extend the duration of the surgical procedure with strain on the user which is undesired.

4.7 Results

After obtaining the ethical approval of the Comité d'Ethique de la Recherche from the University of Strasbourg, agreement CER 2022-56 the experimental protocol was tested with a group of eight users. All users were novice users with null training hours in the STRAS system or any other FSIR system. From the group of eight users, two users had clinical background.

To compare the two control modes statistical analysis is carried out. Preliminary data was found not to be normally distributed, therefore non-parametric tests and metrics are reported. The median and interquartile range (IQR) are reported in table 4.3, which represents the central and dispersion of the data. Statistical analysis by the one-sided Wilcoxon signed-rank test allows to compare the data from the two control modes evaluating if one is better than the other or if there is no difference between them. A p-value

TABLE 4.3: Summary of the results of the experiments with the semi-autonomous arm-body and the manual telemanipulation mode. The experiments compare the manual (**MAN**) and semi-autonomous (**SEM**) modes in a multiple-reaching task. Metrics are separated by dashed lines, first group of metrics are time-related, followed by kinematic and surgical task metrics. Statistically significant differences where semi-autonomous arm-body control mode outperforms manual mode are marked with '*' ($p \leq 0.05$). p-values marked with $^{\triangleleft}$ ($p \geq 0.95$) indicate that manual mode outperforms semi-autonomous arm-body control mode.

metric	A_w	p-**value**	**median MAN**	**IQR MAN**	**median SEM**	**IQR SEM**
t_s	small	$1.0000^{\triangleleft}$	25.23	17.22	31.00	21.69
t_{sI}	small	$0.9924^{\triangleleft}$	16.78	8.40	18.18	12.34
t_{sII}	small	$1.0000^{\triangleleft}$	7.58	5.45	11.43	13.56
a_b	small	0.0000*	2.00	6.50	2.00	2.00
c_b	small	$1.0000^{\triangleleft}$	46.36	29.60	57.71	39.89
f_b	small	0.0000*	0.19	0.33	0.00	0.16
m_b	small	0.9260	4.01	5.81	4.12	6.14
$m_{b,sI}$	small	0.4977	1.47	3.26	0.00	3.36
$m_{b,sII}$	small	0.9334	1.92	2.42	2.12	2.64
i_b	small	$1.0000^{\triangleleft}$	22.04	14.60	25.02	21.85
c_l	small	$1.0000^{\triangleleft}$	12.37	8.57	14.88	10.87
f_l	small	0.9016	0.83	0.85	1.03	0.99
m_l	large	0.0000*	40.00	34.50	4.00	6.25
$w_{l,sII}$	mod.	0.0078*	8.46	15.69	4.59	7.95
w_e	mod.	0.0091*	5.54	12.39	4.03	3.18
t_v	large	0.0000*	3.30	5.25	0.00	0.00
d_{max}	large	0.0000*	7.17	9.45	1.65	1.34
$t_{r,o}$	small	0.1435	0.00	0.10	0.00	0.00
$t_{l,o}$	small	0.8435	0.00	0.00	0.00	0.00

lower of 0.05 indicates a statistically significant advantage of the semi-autonomous arm-body control mode for a given metric (i.e. smaller value in the proposed approach). Conversely, since pair-wise comparison is used with an alternative hypothesis, a p-value higher than 0.95 depicts a statistically significant advantage for the manual mode. Finally, a p-value between 0.05 and 0.95 describes no statistical significance in any control mode.

It is important to note that the p-value is an indicator of statistically significant differences between two distributions but does not say anything about the size of the difference. In fact, any two distributions can yield statistically significant p-values with a large enough sample size (Tomczak and Tomczak, 2014; Ferguson, 2016). Therefore,

an effect size metric, which effectively estimates the magnitude of the difference between two distributions (Kelley and Preacher, 2012), is used in our analysis. Due to the non-normal distribution of the data, the non-parametric estimator for common language effect size A_w (Ruscio, 2008) is used (further details can be found in appendix B). A_w is a dimensionless number between 0 and 1. Values of between 0.56 and 0.64, between 0.56 and 0.64, or > 0.71 depict small, moderate and large effect sizes, respectively (Li, 2016). In summary, to assess a better performance of any control mode statistical significance and effect size need to be evaluated simultaneously (Sawilowsky, 2009; Sullivan and Feinn, 2012; Tomczak and Tomczak, 2014; Ferguson, 2016), with the expectation of statistical significance and at least moderate effect size.

Table 4.3 presents an overview of the metrics. Results show that using manual mode, users take less time to perform the overall task and segments I and II, as evidenced by the median values of t_s, t_{sI} and t_{sII} respectively. Note that the IQR of these metrics is lower during manual teleoperation which suggests a more homogeneous spread of user performance in terms of timing than in semi-autonomous arm-body control mode. The effect size is however small.

Results on the kinematic performance metrics are mixed. Users trigger fewer numbers of body movements a_b and with a smaller frequency f_b in semi-autonomous arm-body control mode, but travel a larger distance with the body, as shown by the values of c_b. On the other hand, the idle body time i_b was shorter and less distance was travelled by the non-dominant arm c_l in manual teleoperation. Yet, the effect size associated with those metrics is small, less than 0.56. The duration of non-dominant arm movements (m_l), and the smoothness of both body and non-dominant arm movements ($w_{l,sII}$ and w_e) were statistically significantly better in semi-autonomous arm-body control mode with large and moderate effect sizes. The smoothness metric in particular shows an important decrease in the IQR, showing a more homogeneous performance of the users in that mode.

Finally, the semi-autonomous arm-body control mode outperforms the manual teleoperation mode regarding the two main surgical performance metrics. The duration of violating the reaching task of the non-dominant arm t_v has a median of 0 for the semi-autonomous arm-body control mode and of 3.30 in the manual teleoperation mode, with statistically significant p-values, and large effect size. In the same manner, the

max recorded distance d_{max} between the non-dominant arm and ${}_w\boldsymbol{p}_d^l$ is lower during the semi-autonomous arm-body control mode, with a median of 1.65 compared to 7.17 for manual mode, again with statistically significant p-value and a large effect size. The last two metrics report the amount of time the dominant (respectively non-dominant) arm spends outside of the field of view $t_{r,o}$ (respectively $t_{l,o}$) and do not give significant insights with values close to zero in all cases.

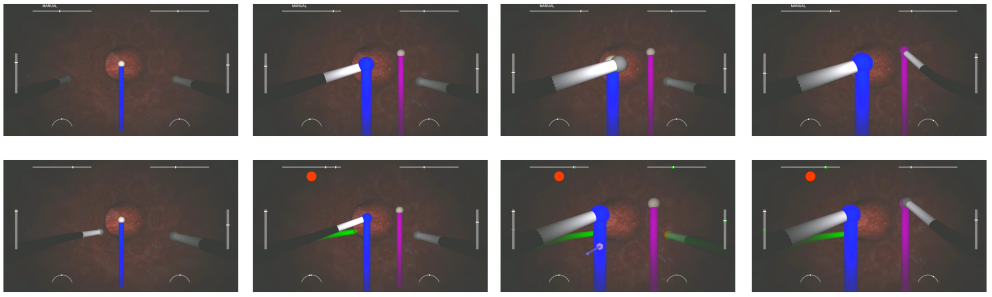


FIGURE 4.9: Overview of manual and semi-autonomous mode. A sequence of images taken from the simulator during manual mode is shown at the top row, while the bottom row shows semi-autonomous mode.

The user's opinion is measured by the NASA TLX, fig. 4.10, measuring the qualitative metrics; where lower values are preferred. The purple bars represent the users' opinions under manual teleoperation mode and the blue bars represent the semi-autonomous arm-body control mode opinion. Overall, the users rank both control modes similarly. Although, there is a notable exception of the temporal demand. Users felt more pressured during the semi-autonomous arm-body control mode due to the indirect body control through the dominant arm, which also affected their frustration. It is expected that the pressure diminishes over time for users and becomes unnoticeable by users in comparison to manual telemanipulation (Kenngott et al., 2012; Zhu et al., 2022).

4.8 Discussion and future work on semi-autonomous polyp removal

This section presents and discusses the major outcomes of the study. The dissection stage in ESD requires users to consider the relationship between the arms and body of FSIR platforms to correctly perform the surgical task. For this reason, a semi-autonomous control method that assists users is proposed. The proposed approach assists users in managing the arm-body relationship allowing them to focus on the task at hand. During

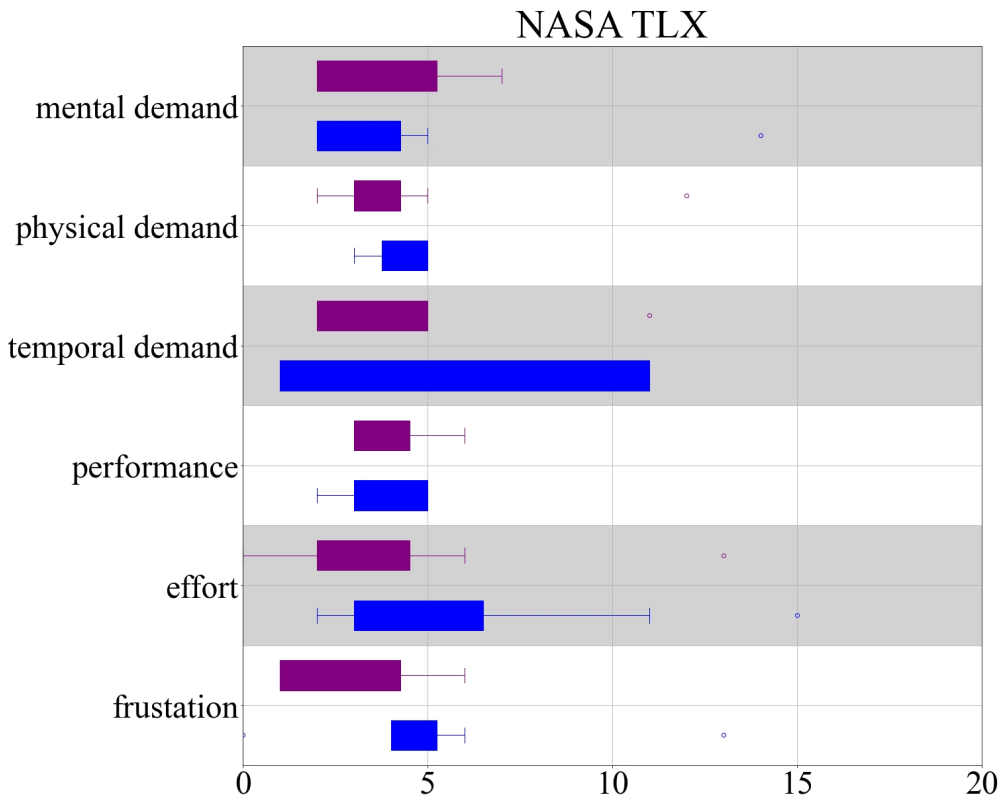


FIGURE 4.10: NASA TLX results from the manual and semi-autonomous arm-body control mode. Purple depicts manual control. Semi-autonomous control mode is depicted in blue.

the dissection stage of ESD, the user is required to grab a tissue and maintain it in a position that eases cutting while controlling the body. The user maintains traction with the non-dominant arm while moving the body which is difficult to perform while telemanipulating FSIR platforms (Meng et al., 2022) —the intricacy of dealing with the coupled architecture is also present in manual flexible endoscopes (Kume, 2009; Thompson et al., 2009). In manual telemanipulation, commanding both arms and the body is a task that involves controlling 10 DoF –while controlling one arm and the body requires 7DoF. During the semi-autonomous control mode, the arms are directly telemanipulated by the handles of the leader console and body control is set by the user commanding a pedal and a leader handle. The proposed approach reduces, the above-mentioned 10 DoF to be controlled to correctly perform dissection, to 3 DoF –those of the dominant arm handle and actuation of the pedal. The user focuses on body control through the handle while the arms can be locked into place.

The user study validates that the semi-autonomous arm-body control allows performing the task in a better and safer way than with manual telemanipulation. This is evidenced

by the task-related metrics performing better with the semi-autonomous arm-body control mode at a time cost with an increase in the temporal demand. Six of the users do not have any clinical background and no evidence of having a clinical background betters or worsens the performance. The group without clinical background did not report any special training or abilities that would invalidate the statement, which further validates the ease of use of the semi-autonomous arm-body control mode. As for many other robotic assistance systems, the assisted mode does not improve the duration performance (Kenngott et al., 2012) as depicted by the duration metrics, t_s, t_{sI} and t_{sII}. This conclusion is supported by the NASA TLX questionnaire outputs depicted in fig. 4.10, for which there is a larger spread of the *temporal demand* responses by the users in semi-autonomous arm-body control mode.

Figure 4.11-a,b shows the distances of the tip of the arms and the target position for two different users. Body movement is depicted in grey highlights. The semi-autonomous arm-body control allows users to move the body while not impacting the distance $\boldsymbol{d}_l$, which is the distance between the non-dominant arm and its target point. With manual telemanipulation mode, there is a clear influence of body movements on $\boldsymbol{d}_l$ when reaching for the dominant arm targets and moving the body, large deviations from ${}_w\boldsymbol{p}_d^l$ which potentially leads to tissue damage and/or loss of traction and will require re-grabbing (red arrowheads on fig. 4.11-a,b). This is because handling coordinated multi-arm movements with an FSIR platform, such as the STRAS, is extremely complex in manual teleoperation. The spikes in the movement of the non-dominant arm in the manual mode are also the possible reason why smoother movements were performed by the non-dominant arm and body in the semi-autonomous arm-body control mode, as evidenced by the values and statistical analysis of $w_{l,sII}$ and w_e.

The proposed approach shows promising results regarding the performance of the task, by easing traction –here modelled as the positional task for the non-dominant arm. Moreover, the semi-autonomous arm-body control mode avoids unnecessary movements as demonstrated by the kinematics metrics performing better than manual telemanipulation. The study has however several limitations which could be addressed in future work.

The intricacy of the dissection stage is modelled using static sequentially presented positions. This first-order approximation is a major simplification in two aspects. The

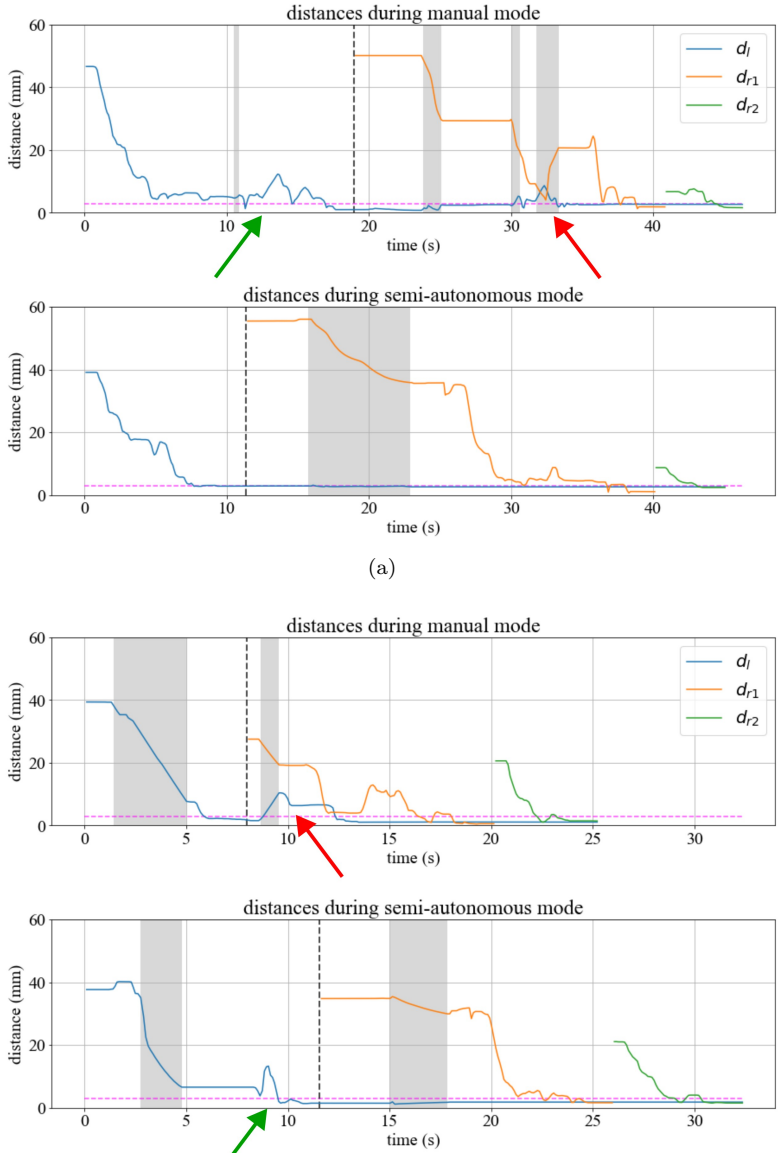


FIGURE 4.11: Example of distances of the arms to the objectives in manual (upper) and semi-autonomous arm-body control mode (lower). a) and b) are trials done by two different users. Body movements are depicted with grey highlights. $\boldsymbol{d}_l$ is the distance between the non-dominant arm and its objective ${}_w\boldsymbol{p}^l_d$, $\boldsymbol{d}_{r1}$ (resp. $\boldsymbol{d}_{r2}$) is the distance between the dominant arm and its first objective ${}_w\boldsymbol{p}^r_{d1}$ (resp. second objective ${}_w\boldsymbol{p}^r_{d2}$). Distances are plotted only for the relevant stages. The thresholds for fulfilling the reaching tasks are represented by the horizontal dashed pink line, of 3 mm. The vertical dashed line represents the transition between stage I and stage II of the dissection. Red arrowheads represent the violation of the surgical task after the ${}_w\boldsymbol{p}^l_d$ objective has been reached (after the black dashed vertical line). Green arrowhead shows an increase in the distance of ${}_w\boldsymbol{p}^l_d$ but is before the reaching objective by the non-dominant arm has been attained.

first one is that during this study the environment is considered to be perfectly known and static. In practice, the environment would be moving with physiological movement. Therefore, estimation algorithms would be needed to estimate the current stage of the

operation (marking, dissection, etc.) as well as the positions to be reached and maintained. Such estimations would add uncertainty to the control system which would need to be tackled in future work. The second simplification is that having traction over the tissue of the non-dominant arm is modelled as maintaining a given position. Even with minimal physiological movement (which can be achieved by insufflation of the colon during ESD, yet can be uncomfortable for the patient post-operation), the desired task is more akin to keeping a desired tension. A desired tension will be required to control the stiffness of the arm and have an accurate state of the robot arms. Full shape estimation and contact detection are needed to allow the desired tension to be fused with the visual feedback that detects the remaining tissue to be dissected. Tension would evolve as the underlying tissue is being dissected and changes with respect to the location. A more complex formulation of the control problem accounting for the evolution of the desired position ${}_w\boldsymbol{p}_d^l$ during the tissue dissection would therefore be needed.

Second, arms telemanipulation is equivalent to manual telemanipulation and with semi-autonomous arm-body control. The target position of the arms is provided by the user and it is not regulated by the semi-autonomous arm-body approach. Depending on the current state of the robot $\boldsymbol{q}$, some goal positions at the current or subsequent positions of the body might not be reachable, leading to a violation of the non-dominant arm tissue-grabbing task. This event happened on a few occasions during the experiments and required the user to stop the semi-autonomous control to release the position of the non-dominant arm, then move the body to a more favourable position, and perform a re-grabbing. These cases were left on the experiments to avoid filtering the total number of experiments, hindering the performance of manual vs semi-autonomous mode. If these results are removed from the total population, the hypothesis is that the performance of the semi-autonomous mode will improve. Such behaviour leads to a considerable time penalty and is likely the source of increased temporal demand for the users (see fig. 4.10), as well as increased time in semi-autonomous arm-body control mode. The longest trials were recorded over the semi-autonomous mode with the occurrence of such events. Coupling our proposed approach with a robust motion planning algorithm for anticipating such events would be a possible solution to mitigate this problem, potentially leading to a more intuitive and effective semi-autonomous control mode.

A third limitation of the study is that it was not tested on the physical follower of the STRAS system. This choice was made to evaluate the effect of the proposed control

approach, without undue disturbance from the non-linearities of the physical system (Poignonec et al., 2020). Translating the proposed approach on the physical system suffers from one major hurdle: state estimation of the robot follower, which is impacted by non-linear frictional and backlash behaviour as well as by forces on the arms when interacting with the tissue. Such a problem could be tackled using endoscopic vision as a sensing input. When coupled with the model of the robotic system, it was shown to be sufficient for state/shape detection using marker-based (Cabras et al., 2017; Zhang et al., 2021) or markerless (Sestini et al., 2021) solutions. To enhance the robustness of the controller, coupling such vision-based approaches with our proposed semi-autonomous control approaches would likely require reformulating some of the objectives in the image space, using visual servo control.

Fourth, duration metrics are not improved during the semi-autonomous arm-body control mode. Yet, kinematic (smoothness of movements) and task-related metrics (performing the reaching tasks with the dominant arm without violating the positional task –traction– on the non-dominant arm) are outperformed by the proposed semi-autonomous arm-body control mode. This increase in duration is potentially linked to the limited duration of the experiments. Users spend 40 minutes –at most– to get familiar with the platform, receive instructions, and perform the training and the experiments. A short duration was selected due to the availability of users and to avoid stress and/or fatigue. During this short period reaching a plateau in the learning curve (Siau et al., 2018) is challenging (Jiménez-Rodríguez et al., 2016), and duration is not expected to be heavily improved –similarly to experiments in laparoscopy (Leijte et al., 2020). Learning curves in laparoscopic robotic surgery are reported to be between 20 and 40 cases (Wong and Crowe, 2022; Choi et al., 2022). Future work could involve a longitudinal study (Zhu et al., 2022) with more users under the simulated scenario and the usage of the physical STRAS system, to validate the gained skills and the fidelity of the proposed simulation with novice users. It would also be interesting to evaluate the proposed control approach with both novice and expert users.

Finally, a minor limitation is that the proposed approach is specifically focused on the STRAS system for practical reasons (the leader and follower systems being available for testing in our lab). The tree-like architecture of the STRAS system is, however, common to many FSIR platforms (Abbott et al., 2007; Hwang and Kwon, 2020a; Nageotte et al., 2020). Therefore, results are very likely to transfer easily to other FSIR platforms.

Dissection being the most difficult part of one of the main procedures targeted by FSIR platforms, it is natural to focus on it. The proposed control formulation, however, could be extended beyond ESD and dissection, since many surgical scenarios feature grasping (Shahkoo and Abin, 2023), tissue tensioning (Portolés et al., 2015), and body or camera movements. Future work could therefore consider extending our proposed approach to multi-arm continuum robotic structures for urology (Hendrick et al., 2014), fetal (Vandebroek et al., 2019), or brain surgery (Price et al., 2023). The proposed approach could allow dealing with the coupled multi-DoF architecture of FSIR platforms and other surgical robotic platforms.

Chapter 5

Assisted colonoscopy in flexible steerable intraluminal robots

In the previous chapter, a method for assistance during the dissection stage of ESD was proposed and validated in a virtual simulator with the physical leader of an FSIR platform, the STRAS. The current chapter expands on the control formulations presented in chapters 3 and 4 and proposes an approach for assisting the user during a colonoscopy and a biopsy. Colonoscopy through flexible endoscopes is the gold standard for colorectal cancer screening due to its low invasiveness and affordability. Considering the similar architecture between manual flexible endoscopes and flexible steerable intraluminal robots, it is a logical surgical procedure to focus on, as the benefits of FSIR could facilitate colonoscopy through FSIR. This chapter proposes a control framework for assisting the user in correctly performing the colonoscopy and biopsy procedure utilising the physical follower of an FSIR platform[1].

During a colonoscopy and a biopsy with FSIR, the users visually inspect the surgical environment while telemanipulating the multi-degree of freedom coupled architecture of flexible steerable intraluminal robotic platforms. Due to the challenges of FSIR ($\mathscr{C}_1$**: coupled architecture** and $\mathscr{C}_2$**: high amount of degrees of freedom**), correctly

[1]The work presented in this chapter has been published in the following two publications:
Herrera, J. F. G., Pore, A., Sestini, L., Sahu, S. K., Liao, G., Zanne, P., ... & Gora, M. (2022). Autonomous image-guided control of endoscopic orientation for OCT scanning. In Computer and Robot-Assisted Surgery (CRAS), Naples, Italy, April 2022
Liao, G., **Herrera, F. G.**, Zhang, Z., Pore, A., Sestini, L., Sahu, S. K., ... & Gora, M. J. (2022, May). Autonomous OCT volumetric scanning with robotic endoscope. In Clinical Biophotonics II (p. PC1214602), SPIE

performing a colonoscopy and a biopsy by FSIR requires investing high amounts of training hours (Nasseri et al., 2021; Harji et al., 2023). Enabling easy telemanipulation of FSIR could reduce the number of hours required to master complex surgical procedures such as colonoscopy and biopsy. Enabling easy FSIR telemanipulation under complex surgical procedures is of special interest as FSIR increments the capabilities of surgeons inside the digestive system. The capabilities of surgeons are increased by novel sensors and actuators mounted on the tip of the arms of FSIR. The emergence of innovative scanning methodologies, such as Optical Coherence Tomography (OCT), enables surgeons to conduct in-situ analysis of abnormal tissue. Although it is not commonly present in manual endoscopic of FSIR platforms today, OCT has the potential to streamline the diagnostic process, reducing the need for multiple visits (Caravaca-Mora et al., 2020, 2023). However, the telemanipulation of OCT tools is complex due to the flexible architecture of their bodies and sensors. Facilitating smooth telemanipulation of the sensory arms and the body is essential during screening and in-situ scanning procedures. A control framework, similar to the one developed in chapter 4, could allow assistance in the colonoscopy and biopsy procedure.

This chapter formulates a QP-based control framework, following the work developed in chapter 4 by the initial formulations of chapter 3, to assist users in correctly performing the colonoscopy and biopsy procedure which handles the visual servoing (VS) for the coordinated control of the subsystems of FSIR. This chapter considers coordinated control of the OCT-capable arm and the body to enable safe and correct colonoscopy and in-situ biopsy. The control framework handles colonoscopy and reaching to perform the biopsy in a synthetic colon model with visual and surface features. The control framework assists users, allowing them to oversee the procedure with minimal intervention.

The chapter is structured as follows. Firstly, the chapter states the problem in section 5.1 with an introduction to the main concepts regarding visual servo control, the navigation of the GI tract and the in-situ inspection of the tissue. Then, section 5.2 covers the related works towards assisted –and autonomous– navigation and in-situ scanning of tissue in the GI tract. Section 5.3 describes the proposed approach and formulates the control framework for visual servoing with the task-function approach. Section 5.4 covers the experiments performed and the results are described in section 5.5. This chapter ends with section 5.6 summarizing the contribution and describing future work.

Note for the reader: This chapter uses the formulation of the Jacobians, HTM and telemanipulation formulation defined in chapter 3 but redefines some of the terms formalised in chapter 4. For the sake of readability, this thesis avoids the usage of subscripts or superscripts and instead redefines the variables, especially present in the control formulation.

5.1 Problem statement

5.1.1 Current state of colonoscopy and biopsy

Colonoscopy (fig. 5.1) through flexible endoscopes requires users to command the control handle and the flexible endoscopic bodies. Flexible endoscopy entails proficiency in manipulating both the control handle and manually manoeuvring the flexible endoscopic bodies for insertion, rotation, and bending within the body (fig. 5.1). While the endoscopic body navigates the digestive system the surgeon visually inspects the colonic walls screening for visual abnormalities. Due to the high number of degrees of freedom and the multiple tasks to consider (fig. 5.2), colonoscopy through flexible endoscopes (Fanelli, 2018) is a two-person procedure, a surgeon and an assistant. The surgeon and the assistant share 5-6 degrees of freedom, those of the body (4DoF) and the tool (1-2DoF). In the most common convention, the surgeon controls the body while the assistant controls the tool. Other conventions have the surgeon controlling the deflections of the body and the tool while the assistant controls the translation and rotation. Independent of the convention, the successful execution of a colonoscopy hinges on effective and continuous communication between the surgeon and the assistant. This collaborative approach minimises the likelihood of overlooking lesions, alleviates patient discomfort, and mitigates associated risks.

A correct colonoscopy visually inspects the complete digestive system (fig. 5.1) from the colon to the cecum and locates the suspected lesion(s). To correctly perform a colonoscopy, users must align the centre of the endoscope body to the centre of the lumen, reducing unwanted interaction between the digestive system and the tool and/or body. Minimised contact with the colonic wall reduces patient discomfort and the risk of bleeding (Rabeneck et al., 2008), infection (Fisher et al., 2011) or perforated intestines (Lohsiriwat, 2010). In navigation of the digestive system, lumen and body alignment

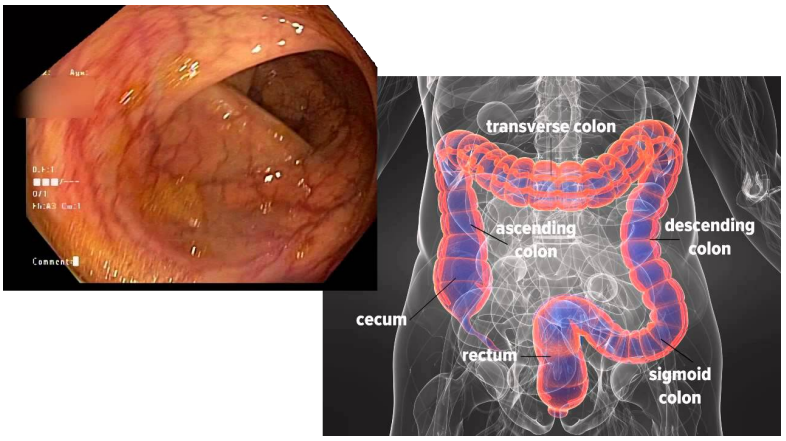


FIGURE 5.1: Diggestive system, as is explored during colonoscopy. Image adapted from Johns Hopkins Medicine (2024)

maximise the field of view from the endoscopic camera view and the workspace of the body, which amounts to 4DoF while surgeons visually inspect the gastrointestinal tract. After a suspect tissue has been detected by visual inspection of the endoscopic images, a biopsy is performed by taking a sample of the tissue. To take a biopsy, a forceps is inserted through one of the endoscope channels. The sample is sent to the pathological lab for analysis. Once evaluated, the patient revisits the clinician to receive the diagnosis and, if necessary, schedule any required interventions. If multiple abnormal tissues are detected, multiple samples are taken. If multiple tissues are detected as abnormal the surgeon must track and locate them in the digestive system and later perform adequate procedures regarding each of the samples' diagnosis.

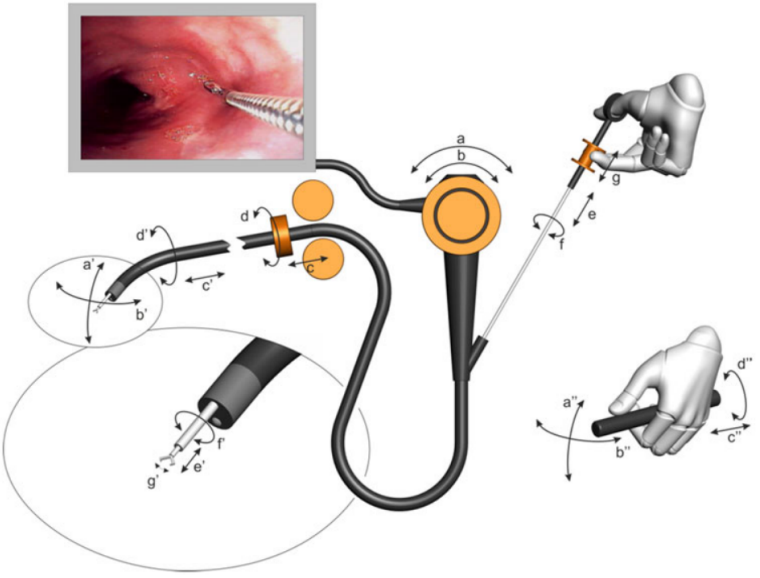


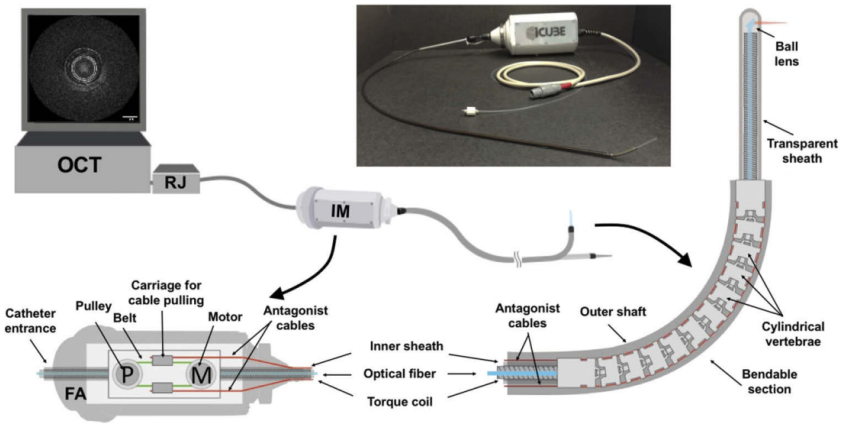
FIGURE 5.2: Degrees of freedom in manual flexible colonoscopes. Image adapted from Ruiter et al. (2013).

5.1.2 Optical biopsies using OCT

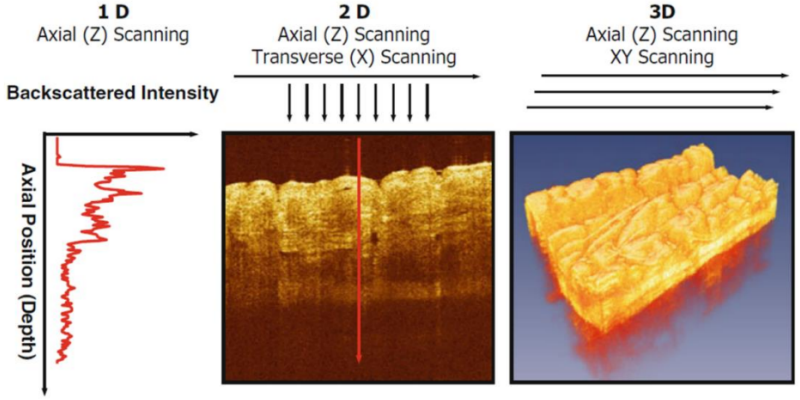
To correctly perform a biopsy, surgeons and assistants are required to coordinate their movements, ensuring the forceps tip is accurately positioned at the centre of the abnormal tissue. Care must be taken to extract a sample from the tissue while minimising inadvertent interaction with the digestive system without impacting the accuracy of the sample. The surgeon and the assistant have to control the 4DoF of the body to indirectly control the position and orientation of the tool since the tools can only rotate and translate (1 or 2 DoF at most). Colonoscopy through flexible endoscopes is a procedure that requires high amounts of training hours (Marshall, 1995; McCashland et al., 2000) to deal with the multiple DoF and the challenging architecture. Furthermore, when abnormal tissue is detected, the patient must perform multiple visits (Wexner et al., 2001; Lieberman et al., 2012). Surgeons and assistants must spend several hours in the operating room just to get familiar with the procedure. Then, they can assist in the procedure to build experience which leads them to eventually be able to perform procedures as the main surgeon. Overall it is a long training procedure to build the required skills (Vassiliou et al., 2010; Haycock et al., 2010) necessary to perform colonoscopy and biopsy.

After a biopsy has been taken, the analysis of the tissue is performed in a pathology lab after several preparation steps which depend on the tissue type and suspected lesion. Such steps can take a long time, meaning that the biopsy result is not typically known during the colonoscopy. This time delay can lead to the surgeon over-sampling the tissue, i.e. taking unnecessary biopsies in order to evaluate suspicious tissue, which harms the patient (Wang and Van Dam, 2004). It can also be the source of erroneous diagnosis if the original biopsy was not taken at the correct place. Finally, even in the event of a correct diagnosis, if the outcome is deemed malignant the time delay means that the patient will have to go through a second operation. As a result, researchers have looked at alternatives for assessing the tissue *in situ*, in order to establish a diagnosis in real-time and remove the tissue directly if necessary.

A first possibility is to use endoscopic images in order to detect abnormal tissue and guide the surgeon. Various approaches have been developed in endoscopic settings, in order to visually identify normal/abnormal tissue (Tajbakhsh et al., 2015; Kominami et al., 2016; Zhang et al., 2016; Ribeiro et al., 2016; Yeung et al., 2021; Jha et al., 2021), or segment



(a) Schematic drawing and picture of the OCT steerable catheter. Image adapted from Caravaca-Mora et al. (2020).



(b) Optical Coherence Tomography. Cross-sectional images are generated by performing 1D axial scans to obtain depth information (Z), transversal plus axial (ZX) scans for 2D information (B-scan), and axial scanning plus XY scanning for three-dimensional imaging. Image adapted from Drexler and Fujimoto (2015).

FIGURE 5.3: OCT schematic and OCT images overview.

the different parts of the image, e.g. segmentation of surgical robotic arms (Rosa et al., 2019; Sestini et al., 2021), the environment (Kang and Gwak, 2019; Jha et al., 2021; Al-Bander et al., 2022) and potential abnormal tissue (Fernández-Esparrach et al., 2016; Wang et al., 2019). After such information has been segmented, multiple approaches for exploiting it have been developed based on their applications, e.g. centroid detection of lumen segmented area for correct navigation in the lumen. Such methods can be useful for guiding the surgical gesture but do not allow a clear diagnosis.

An alternative way to diagnose tissues *in situ* is to deploy advanced imaging tools. Novel scanning modalities enable accurate assessment of tissue by their micro-scale features instead of just image data. Many different imaging modalities have been proposed,

for instance, intravascular ultrasound (Sathyanarayana, 2006; Gatta et al., 2009; Garcìa-Garcìa et al., 2011), confocal endomicroscopy (Goetz et al., 2011) or OCT (Huang et al., 1991; Feldchtein et al., 1998). This thesis focuses on OCT, which offers a good compromise between resolution, field of view, and depth penetration into the tissue (Drexler and Fujimoto, 2015). An OCT tool (Huang et al., 1991; Gora et al., 2017) performs in-situ scanning by measuring the "echo" time delay of the reflected light in the tissue by the phase difference between the backscattered light from the sample and that from a reference beam. The depth information from the tissue, referred to as A-line, is obtained allowing to perform sweeps on the tissue surface, referred to as B scan mode. The A-line is one dimensional while the B-scan is a 2D image, which enables 3D information (fig. 5.3(b)) of the tissue by moving the plane of imaging of the B-scan. The tissue can be imaged in situ and in real-time with resolutions of one or two orders of magnitude higher than ultrasound (Drexler and Fujimoto, 2015) by the multiple scans. Multiple scans across the tissue of interest enable complete cover of the tissue and an accurate assessment in situ of healthy and abnormal tissue (depicted in fig. 5.4).

Recent studies have shown that miniaturizing OCT sensors is possible for bringing the probe inside the patient (Gora et al., 2013, 2017), which also makes it possible to mount OCT sensors at the tip of the surgical arms in endoscopic and/or FSIR platforms (Caravaca-Mora et al., 2020). Surgeons can then manipulate the arm equipped with the OCT probe to regulate both its position and orientation, thereby controlling the movement and alignment of the OCT probe (Caravaca-Mora et al., 2023). Scanning by OCT pullback in FSIR platforms has been covered recently by Caravaca-Mora et al. (2020); Zhang et al. (2021); Caravaca-Mora et al. (2023) with positive results in real-time in-situ assessment of lesions in surface-realistic colon model.

5.1.3 Problem overview

This chapter focuses on the scenario of a colonoscopy with in-situ scanning by an OCT arm. Even if recent studies have shown an interest, performing such a clinical gesture is still challenging. It involves aiming the imaging plane of the OCT probe at a polyp while taking care of the multi-DoF cooperation between the arm and the body of the endoscope. This problem involves acting based on the visual data while performing

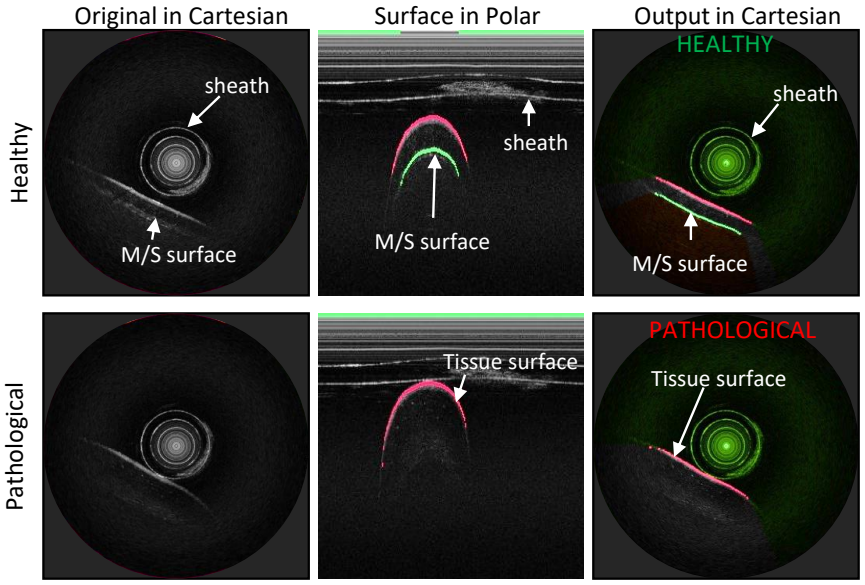


FIGURE 5.4: Rotational scanning by the OCT with healthy and abnormal tissue. Image adapted from Caravaca-Mora et al. (2020).

coordinated control in FSIR to scan the suspicious tissue. Visual data must be considered while performing coordinated control of the arm and body in an intricate and unstructured environment. To achieve this, flexible steerable intraluminal robotic platforms are telemanipulated robotically by users from outside the patient's body, instead of manual manipulation of endoscope bodies and surgical tools. Similar to flexible endoscopes, colonoscopy and biopsy through FSIR requires users to perform two main tasks: *i)* telemanipulate the arms and body (deal with the high amounts of DoF and the coupled architecture), and *ii)* visually inspect the different elements of the surgical environment. The first task requires considering the 7 DoF of the arm and body, with the coupled architecture that depends on the feedback of the visual inspection. The second task requires the user to perform real-time hand-eye coordination to correctly translate their commands through the leader console (using their fingers, arms and feet in the joysticks, handles and pedals) of FSIR and inspect the gastrointestinal walls of the patient's digestive system. Individually, the two tasks are difficult to perform, simultaneously performing both tasks is both difficult and can be detrimental for the patient (Liberman et al., 2005; Fritscher-Ravens et al., 2009). The increment in difficulty poses a risk to the patients due to the high number of DoF which can be actuated erroneously or extend the surgical procedure duration. A control framework that considers the multiple DoF and the visual information of the surgical task to correctly perform a biopsy has the potential to reduce the complexity for users to allow the correct performance of

colonoscopies. An overview of the most relevant works regarding assisted colonoscopy and robotically assisted in-situ biopsy is presented next.

5.2 Related works

In a standard colonoscopy procedure, with or without a biopsy, surgeons mainly rely on the feedback from the camera at the tip of the endoscope (or the FSIR body). In a robotically assisted setting, it is also natural to exploit information from the images in order to close the loop and assist the surgeon's gestures. In robotics, this is the task of visual servoing, acting on the robot based on visual data. In the following, a brief introduction of the main concepts in visual servoing (section 5.2.1) is presented, before exploring the main works in the literature regarding visual servoing-based assistance in endoscopic settings, in navigation (section 5.2.2) and arm control (section 5.2.3) tasks. Finally, section 5.2.4 will finally describe the remaining challenges and the contributions of this chapter.

5.2.1 Visual servoing

Visual servoing (Hutchinson et al., 1996; Chaumette et al., 2016) relies on exploiting *visual data* to control the motion of a robot. Visual data is obtained by imaging devices which can be very diverse: cameras (Azizian et al., 2014), depth sensors, ultrasound probes (Nadeau and Krupa, 2013; Nadeau et al., 2014), tomographic imaging devices (Azizian et al., 2015), etc. The general principle is to extract information from such visual data in order to derive a control law for the robot.

Two main approaches for visual servoing have been developed, depending on the nature of the information considered. The first one, Position-Based Visual Servoing (PBVS), consists of using 3-D parameters estimated from image measurements to control the robot movements (Yasunaga et al., 2003; Chitrakaran et al., 2007). For instance, estimating the 3D position of an object based on a calibrated pair of cameras, and using this information to control the robot movement. In many contexts, and especially in medical settings where environments are unstructured and imaging sensors are not always perfectly calibrated, doing such 3D estimations is complex. An alternative is to perform Image-Based Visual Servoing (IBVS). In IBVS, the control law is directly formulated

in the image space, e.g. aligning pixel positions to environment points depicting 3D positions in the image (Wu et al., 2014; Zhang et al., 2021).

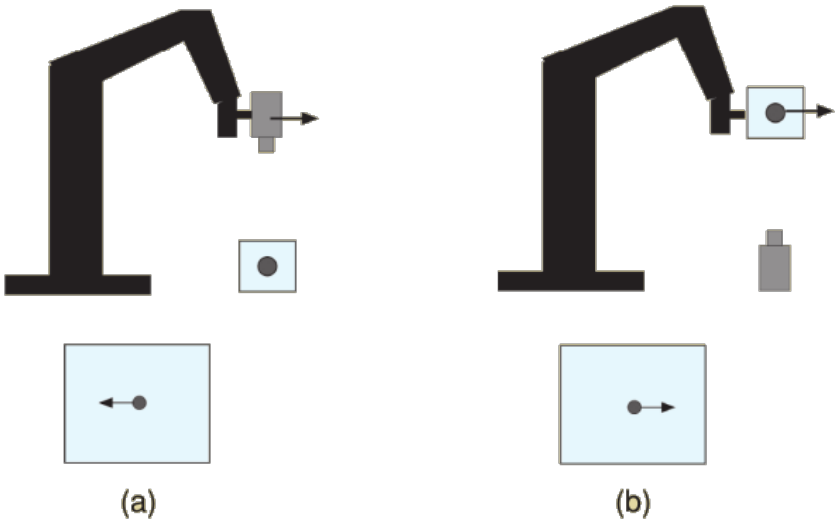


FIGURE 5.5: Visual servoing configurations. a) eye-in-hand; b) eye-to-hand. Image adapted from Hutchinson et al. (1996).

In parallel to the type of information considered, visual servoing approaches can also be classified by the configuration of the imaging sensor and the controlled robot (fig. 5.5). In the *eye in hand* configuration, the imaging device is mounted on the end-effector tip (Allen et al., 1993; Marchand et al., 2002). Conversely, in the eye-to-hand configuration, the imaging devices observe the scene and the end-effector together. Taking the example of the arm-body structure of the STRAS robot described in chapter 3, controlling the movements of the body based on visual input from the endoscopic camera would be an eye-in-hand scheme, while controlling the movements of the arms based on visual data from the same camera would be eye-to-hand. Along the same lines, if one arm is equipped with an OCT sensor, controlling the movements of said arm with OCT data would, again, be an eye-in-hand task.

For any type of approach and configuration, information must first be extracted from the imaging sensors in order to be fed back to the control law. This can be performed by classical image vision tools such as edge detection (Li et al., 2005; Wang et al., 2018a), colour segmentation, (Wei et al., 1997; Kragic and Christensen, 2001; Doignon et al., 2005) or by visual codes, e.g. QR codes (Shi et al., 2018; Zhang et al., 2021). More recently, deep learning based approaches such as CNN have shown promising results

for estimating various features of interest and exploiting them in visual servo control schemes(Saxena et al., 2017; Tokuda et al., 2021; Lazo et al., 2022) or even directly learning the whole visual servo control scheme Bateux et al. (2017, 2018).

In surgical images, robustly extracting visual data is complex due to the dynamic nature of the surgical environment. For this reason, many dedicated approaches have been developed. In the next section, an overview of the main works related to visual servoing for endoscope control (i.e. navigation of endoscopes inside the lumen is presented).

5.2.2 Assisted navigation of the digestive system

Manual navigation of an endoscope inside the lumen of an organ (colon, esophagus, etc ...) requires the user to perform real-time detection and tracking of each of the main actors of the surgical procedure: colonic walls, lumen, abnormal tissue and surgical tools (possibly mounted with the OCT probe) while accounting for their unique visual and physical features. This task presents a complex challenge because users must navigate multiple degrees of freedom within a demanding surgical context, and has therefore been the subject of robotic assistance research in the past.

Assisted manual navigation of flexible endoscopes has been previously proposed for gastroscopy (Reilink et al., 2010) which is analogous to colonoscopy as the user has to navigate the digestive system while aligning the body to the centre of the lumen. Reilink et al. (2010) motorised the handle of a gastroscope, in order to automatically steer the endoscope tip towards the lumen (detected using image thresholding operations) in an IBVS, eye-in-hand setting. On the contrary, insertion/retraction DOFs are handled by the clinician. Using a similar actuation scheme, Ott et al. (2008, 2011) proposed to assist the users during flexible endoscopy by automatically compensating physiological movements such as breathing, with validation in an in-vivo porcine model.

More recently, deep learning-based visual servoing approaches for flexible steerable robots and endoscopes have been developed, with first results in simulated environments (Saxena et al., 2017; Athiniotis et al., 2019; Pore et al., 2022; Corsi et al., 2023) and, more recently, on physical platforms (Lazo et al., 2022).

All the above-mentioned approaches share a similarity in the way the visual servoing problem is posed: it is an eye-in-hand IBVS scheme in which the information is extracted

from endoscopic cameras images. Similar schemes have in fact been exploited for other types of continuum robots in the literature, as detailed in Nazari et al. (2022). All the above-mentioned approaches however, cannot be applied for the *tissue identification* step in which the OCT probe needs to scan the tissue. An overview of the approaches for robotically assisting this step is presented now.

5.2.3 Assisted in-situ scanning for tissue identification

The second step of the clinical scenario is the assessment of the suspected tissue. In the considered scenario, the OCT scanning sensor is mounted at the tip of the surgical arms. In this context, Caravaca-Mora et al. (2023) presented comparisons between manual, telemanipulated and automatic scanning in a benchtop scenario, showing a clear advantage of the automatic scanning trajectory with a smoother and more accurate trajectory. In this study, however, the trajectory was predefined for the arm.

Again, visual servoing is a common solution in the literature in such cases. In rigid endoscopy settings, several studies have exploited visual servoing for controlling the endoscope so that the arms stay in the center of the field of view (Voros et al., 2007; Gruijthuijsen et al., 2022). Krupa et al. (2003) directly control the robotic arms to obtain specific 3D poses in an eye-to-hand PBVS setting.

Closer to our problem, Zhang et al. (2021) presented an approach for automatically reaching a specific location in a controlled environment using QR codes. The view from the camera at the tip of the body is used to guide the movements of the arm of the STRAS robot using a quadratic programming approach with markers. Zhang et al. (2021) implemented the two VS configurations to automatically scan a surface. The body was controlled by the visual data provided planar board filled with QR codes. The arm was controlled by visual servoing with coloured markers at the tip of the arm. Jointly the arm and body were controlled to primarily maintain certain orientation and image position with contact, with the secondary task being body displacement. The approach relies on an optimal control approach using a constrained QP formulation, which is similar to the formulation of chapter 4. However, the approach developed by Zhang et al. (2021) was deployed in a benchtop scenario with an unrealistic environment relying on markers and QR boards which are not translatable to the real surgical case.

5.2.4 Conclusion and contributions

In this section an overview of the most relevant state of the art regarding robot-assisted navigation and tissue identification for colonoscopy and biopsy, particularly by exploiting visual servoing. As one can see, the two-step procedure considered in this chapter is not present in other related works.

The most representative work that considers navigation of the GI tract with in-situ scanning is Zhang et al. (2021). This work describes a dual eye-in-hand and eye-to-hand approach for contact and scanning in a 2D-defined trajectory (i.e. IBVS). Although developed and tested in a very controlled, unrealistic environment, it illustrates very well the fact that the task-function approach is well suited for the case of colonoscopy and biopsy.

Indeed, this case has multiple constraints that must be considered, such as keeping a safe arm position and/or surgical targets in the screen. Additionally, the control of the arm and body requires dealing with multiple frames of reference as well as with the coupled architecture of the FSIR. Furthermore, to facilitate the body navigation lumen must be kept as much as possible in the center of the image. Finally, when the objective for the arm is unreachable at the current body position, the arm and body must collaborate. Some tasks can conflict with each other yet the FSIR platform should perform the task of colonoscopy and biopsy.

The objective of this chapter is to assist the user in correct and safe identification and in-situ assessment of suspicious tissue through eye-in-hand and eye-to-hand IBVS for controlling the body and the OCT arm. In the following, the contribution of this chapter, which is formulating the clinical objectives and constraints of the two-step surgical procedure in the task-function approach based on a visual servoing scheme, is presented.

5.3 Colonoscopy and biopsy performed by flexible steerable intraluminal robotic platforms

A colonoscopy through flexible steerable intraluminal robotic platforms starts by manually inserting the body of FSIR inside the patient's digestive system through the anus.

Then, the user can command the system inside the gastrointestinal tract allowing navigation by telemanipulation of the leader console. While the body is navigating the digestive system, the body must be centred on the lumen –as initially described in section 5.1. The user simultaneously navigates and visually inspects the tissue of the digestive system. The user can move forward and backward while bending and rotating, covering the completeness of the digestive system, and screening from the rectum to the cecum. When the user finds an abnormal tissue, it stops navigation and approaches the tissue. While approaching the tissue, the user keeps the body and lumen centred, minimising the complexity of the problem by maximising the workspace of the body and the visibility. Depending on the location of the lesion, maintaining the lumen-centred could be ignored. Then, the body aligns towards the proximities of the tissue until the position of the lesion is within the workspace of the OCT arm at the current body position, ignoring lumen and body alignment. The multiple targets to track and the multi DoF of FSIR make colonoscopy through telemanipulation a difficult task. The user has to control the body, visually inspect the tissue, and track and centre the lumen to perform a colonoscopy. The OCT probe at the tip of the arm the body position and the surgical environment required to be considered to correctly perform a colonoscopy and a biopsy. A formulation of the problem (fig. 5.6) is presented now.

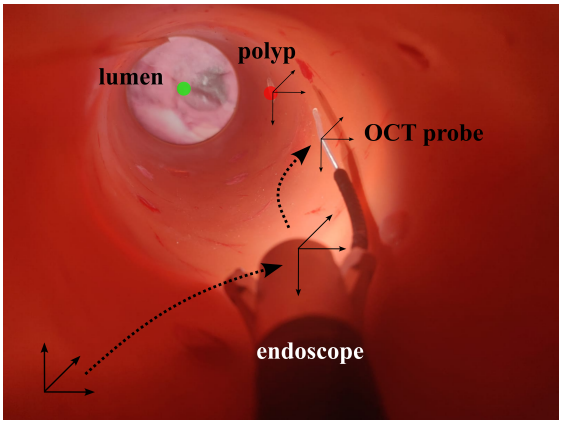


FIGURE 5.6: Overview of the colonoscopy and biopsy.

5.3.1 Problem statement

For the sake of simplicity, the term "lesion" is used, referring to abnormalities in the colon model which can be detected by the visual detection module. The lesion will be later classified as a specific type of polyp, cancer, benign tissue or normal lesion

in the digestive system. An overview of the steps the user is required to perform by telemanipulating the FSIR in colonoscopy and biopsy is presented now.

After the surgeon has introduced the FSIR into the anus of the patient and initiated the navigation of the digestive system, the user has to control the 4DoF of the body while visually inspecting the tissue. The user must manually inspect and track the tissue and the lumen. When the OCT arm has been deployed the user has to control 7DoF, 4DoF of the arm and 3 DoF of the tool –in comparison to the 5-6DoF to be manually controlled during flexible endoscopy. While controlling 4 or 7DoF the user is required to view the screen and scan and track the OCT tool, the lumen and the entire surgical environment to find the largest lesion and assess it. The large number of DoF, the coupled architecture, the eye-to-hand coordination and safety requirements make the tasks mentally demanding. After a lesion has been detected, the body of the FSIR approaches the lesion and OCT scanning is performed.

Approaching the lesion requires the user to control the arm, 3 DoF. When the lesion is outside of the workspace of the arm, the body needs to move. If coordinated control (3 and 4DoF, of the arm and body) was easy to perform by the user, the body and arm would move in coordination to move the arm with the body, such that the new body position allows to reach the lesion. Since coordinated control is difficult, the user controls the body and arms separately and over- or under-compensates the effect on the body or arms. The abrupt or slow movements of the arm and body cause the incorrect compensation of the actions of the arm or body and vice-versa which may increase the duration of the procedure or place the arm or body in an unsafe position. To correctly reach the lesion, coordinated movement must be performed by the user while tracking and estimating the position of the targets on the endoscopic camera view, performing visual servoing. After the lesion is within reach of the arm, automatic OCT movements can be performed to correctly perform a diagnosis of the lesion.

A colonoscopy with a biopsy can be divided into three steps: navigating, reaching and scanning. This chapter proposes a technique to assist during the navigation and the reaching.

5.3.2 Problem modelling

The control approach considers assisted colonoscopy and biopsy –excluding OCT pullback– by two main tasks: *i)* aligning the body to the lumen, and *ii)* aligning the scanning tool to a region of interest in the 3D task space. The *i)* tasks controls the FSIR body while task *ii)* requires performing coordinated control between the arm and the body. To fulfil the task considered, the control approach must perform visual servoing while considering the coupled multi-DoF architecture. Eye-in-hand visual servoing is required to align the centre of the body to the centre of the lumen. Eye-to-hand visual servoing is needed to align the scanning tool to the lesion image space or to a 3D position, constrained by the image position after approaching. The visual servoing uses the detected image positions (fig. 5.8) of the objects of interest —lumen, OCT tool and abnormal tissues— and acts on the required subsystem of the FSIR –arm or body. Furthermore, and similarly to the proposed control framework presented in chapter 4, safety concerns by joint limits are considered. Overall, the conceived tasks and the safety concerns jointly allow for assisted colonoscopy in FSIR, which is difficult to perform by manual telemanipulation.

For the first task, the body end-effector must be aligned to the centre of the lumen, controlling the body as:

$$ {}_w\dot{\boldsymbol{p}}^e = {}_w\boldsymbol{J}_e\dot{\boldsymbol{q}}_{\boldsymbol{e}} \tag{5.1} $$

to match the image centre, which is the central axis of the body of FSIR, to the centre of the detected lumen. The second task requires controlling the arm in the current body position:

$$ {}_e\dot{\boldsymbol{p}}^r = {}_e\boldsymbol{J}_r\dot{\boldsymbol{q}}_{\boldsymbol{r}} \tag{5.2} $$

matching the position of the arm in the image and the detected largest lesion.

Then, coordinated control in colonoscopy and biopsy is especially required when reaching the lesion, just before the OCT performs the precise sweeps to assess the tissue. The arm and body (eqs. (5.1) and (5.2)) should be controlled to reach the proximities of the tissue without losing alignment of the centre of the lumen and the body. Due to the coupled architecture and to extend the workspace of the arm safely joint limits for the arm are considered. Then, body movements can be performed while the arm compensates to stay within a safe workspace, specified only by the arm translation. For simplicity and to due the fact that the joint values for the translation of the arm mostly translate into

the modification of the Z-axis end effector position, the arm position is bound to:

$$trans_{min} \leq_w \dot{\boldsymbol{p}}^r \leq trans_{max} \tag{5.3}$$

where the $trans_{min}$ and $trans_{max}$ are defined as safe translation value in the taskspace. The arm and body coordinated control is required, as:

$$\text{coordinated control in navigation and biopsy} \begin{cases} {}_w\dot{\boldsymbol{p}}^r = \begin{bmatrix} {}_e\boldsymbol{J}_r\dot{\boldsymbol{q}}_{\boldsymbol{r}} & {}_w\boldsymbol{J}_e\dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix} \\ {}_w\dot{\boldsymbol{p}}^e = \begin{bmatrix} {}_w\boldsymbol{J}_e\dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix} \end{cases} \tag{5.4}$$

to perform visual servoing with the arm, and the body to enable correct navigation –by extending the arm workspace and maintaining the arm in a safe workspace. Where the first row of eq. (5.4) expresses eqs. (5.2) and (5.3) and the second row eq. (5.1)

5.3.3 Overview of the approach

Similarly to the control framework developed in the previous chapter, the multiple tasks are expressed by the tracking errors, to then take an optimisation form. The first task, aligning the body to the lumen is expressed using the lumen centre $\boldsymbol{m}$ and the image centre $\boldsymbol{n}$:

$$\mathscr{B}_1 : \ \min \|\boldsymbol{n} - \boldsymbol{m}\| . \tag{5.5}$$

Where the $\boldsymbol{m}$ and $\boldsymbol{n}$ in are eq. (5.5) are two dimensional vectors. For the second task, the position of the lesion in the image space ($\boldsymbol{s}$) is matched with the image position of the tip of the OCT arm, ($\boldsymbol{p}_O$), similarly, as objective:

$$\mathscr{B}_2 : \ \min \|\boldsymbol{s} - \boldsymbol{p}_O\| . \tag{5.6}$$

The position of the lesion might not be within the reach of the workspace of the arm. When this occurs, the body must compensate for the lack of reach in the arm and apply some translation. A safe range of translation should be maintained eq. (5.3) in the form of:

$$\mathscr{B}_3 : \ \min \|(\boldsymbol{p}_O^*|_z) - (\boldsymbol{p}_O|_z)\| . \tag{5.7}$$

The three objectives require controlling the arm, and the body and simultaneous control during colonoscopy while considering the safety constraints.

The positions of the prior objectives are obtained by a visual detection module. The vision module previously developed was initially used during (Gonzalez Herrera et al., 2022) and it is based on a U-Net convolutional neural network. The U-Net network uses the endoscopic camera frames and segments its main components (arm, lumen and lesions). The centroid of the segmented area of each actor (arm, lumen and lesions) is taken in the current endoscopic image. The lumen, $\boldsymbol{m}$, is defined as the detected lumen in the current image position. Note that the lumen refers to the inner space of the digestive tube where food and digestive fluids move along. In the case of navigation inside the patient, the lumen is referred to as the centre of the detected outer bounds of the GI tract, thus maintaining the body close to the lumen avoids the colonic walls. The biggest lesion is described as $\boldsymbol{s}$, and it is registered at the initial body position. While $img(\boldsymbol{p}_O)$ is the image position of the tip of the OCT probe which moves as the body moves. An overview of the positions is presented in figs. 5.6 and 5.7. The values of the arm $\boldsymbol{p}_O$, lumen $\boldsymbol{m}$ and image centre $\boldsymbol{n}$ are defined in the image space. The positional values are used for the control problem, as previously introduced in section 5.3.2. Furthermore, automatic OCT pullback is performed by another module, the OCT controller. The OCT controller manages correct control of the OCT probe and performs the OCT pullback that extends the workspace of the OCT probe.

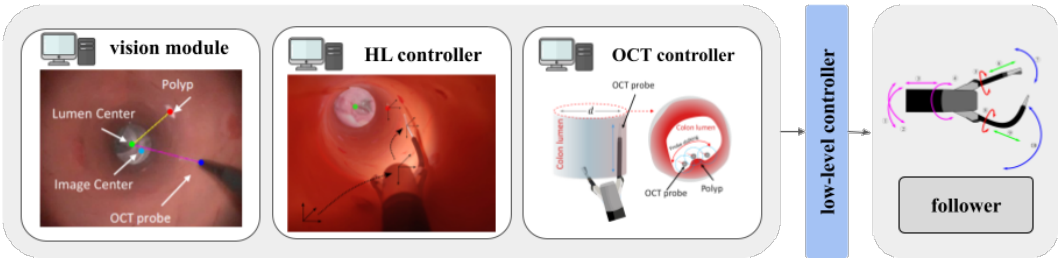


FIGURE 5.7: Overview of the proposed solution for assisted colonoscopy and biopsy. The vision module detects the position of the lumen, lesion, and the OCT tool. A high-level (HL) controller acts on the desired objectives to assist the colonoscopy while the OCT controller governs the OCT probe. The commands are performed by the STRAS follower side by the STRAS low-level controller.

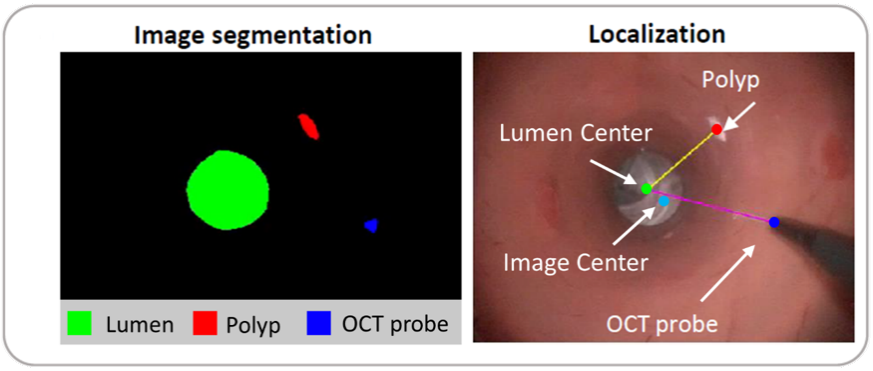


FIGURE 5.8: Overview of the visual detection module outputs. On the left side: Output from the image segmentation module. On the right side: real-time endoscopic image segmentation for lumen, polyp and OCT probe localization.

Similarly to the approach in chapter 4, the optimization problem that is solved by QP solver is set as:

$$\min_{\dot{\boldsymbol{q}}} \ \frac{1}{2}\,\dot{\boldsymbol{q}}^T \boldsymbol{H}_t \dot{\boldsymbol{q}} + \boldsymbol{c}_t^T \dot{\boldsymbol{q}} \tag{5.8}$$

$$s.t. \ \ \dot{\boldsymbol{q}}_t^+ \leq \dot{\boldsymbol{q}} \leq \dot{\boldsymbol{q}}_t^- \tag{5.9}$$

where $\boldsymbol{H}_t = \sum \gamma_i \boldsymbol{H}_i$ and $\boldsymbol{c}_t = \sum \gamma_i \boldsymbol{c}_i$ defined accordingly to the formulation in the previous chapter for each of the i previously defined objectives (as the expressions $\mathscr{B}_1$, $\mathscr{B}_2$ and $\mathscr{B}_3$), and γ_i is a given object's weight. Solving eq. (5.8) defines a set of $\dot{\boldsymbol{x}}$ that reduces the tracking error of each of the tasks to allow for assisted colonoscopy.

Given that only the primary arm –the arm equipped with the OCT scanning tooltip– and the body are actuated, the appropriate vectors of controlled positions are defined as:

$$\boldsymbol{p} = \begin{bmatrix} {}_e\boldsymbol{p}^r \\ {}_w\boldsymbol{p}^r \end{bmatrix}, \tag{5.10}$$

with the configuration variables defined as:

$$\dot{\boldsymbol{q}}_{qp} = \begin{bmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{r}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{e}} \end{bmatrix} \tag{5.11}$$

and a Jacobian that correctly maps the positions with the configuration variables is defined as:

$$\boldsymbol{J} = \begin{bmatrix} {}_e\boldsymbol{J}_r & 0 \\ {}_e\boldsymbol{J}_r & {}_w\boldsymbol{J}_e \end{bmatrix}. \tag{5.12}$$

With the forward kinematic model that links configuration space variables to task space as:

$$\dot{\boldsymbol{p}} = \boldsymbol{J}\,\dot{\boldsymbol{q}}. \tag{5.13}$$

For $\mathscr{B}_1$, the eq. (5.10) is defined by the speed formulation of $\left[0,\ {}_w\boldsymbol{p}_d^e\right]^T$ with a Jacobian eq. (5.12) non-zero in the last row and last column, ${}_w\boldsymbol{J}_e$ only being active. Similarly, for objective $\mathscr{B}_2$ the task Jacobian is formulated with the first two elements populated. For $\mathscr{B}_3$, the Jacobian populated only in the last row, as $\left[0,{}_w\,\boldsymbol{p}_d^e\right]^T$ with a task Jacobian that is only populated in the last row, containing ${}_e\boldsymbol{J}_r$ and ${}_w\boldsymbol{J}_e$. The objective is defined as a desired position z_d in the Z component, such as $[0, 0, z_d]$, compensating the translation of the arm. When all of the objectives are active, the lesion and tool are active in the image position in the local frame of the body, such that $\mathscr{B}_2$ affects the global frame, actuating the body and arm to maintain a desired translation. Jointly, the three objectives that are activated in the different stages allow for assisted colonoscopy.

The multiple objects of interest in the screen (actors: lumen, OCT arm and lesion) have unstable locations due to the online detection. A moving average filter is applied to smooth the noisy image detection. Since multiple lesions can be detected at the starting point, to simplify the problem, the approach detects and reaches the largest lesion. A visual tracker (Babenko et al., 2009) follows the largest lesion during the entire procedure. A small time window, referred to as the training phase is performed at the beginning of the experiment, to have online registration of the current positions of the actors. During the training phase the visual tracker and the average filter record stable measurements. Since estimating the 3D position on lesions with variable size is difficult, the experiments consider the largest lesion to be one that also is the largest on the image. The filtered output is used by the control formulation allowing to perform each of the described tasks.

5.4 Experiments

The control formulation is tested in a controlled environment. The surgical environment is emulated through a colon model that has the visual and surface geometry features of different types of tissue in the gastrointestinal walls. The STRAS robot is placed in the bounds of the colon model, as to require moving towards the lesion. At the start of

the experiments, the OCT arm is placed inside the channel of the body of the surgical arms in the STRAS robots. When required the OCT arm is deployed to allow tracking and reaching the desired lesion. The orientation and position of the colon model are modified to simulate different clinical situations. The trial starts by aligning the lumen to then reach a desired image position. The trial ends when a condition is met, in some of the experiments is an image position based on the largest abnormal tissue or an image position that depends on the position of the body.

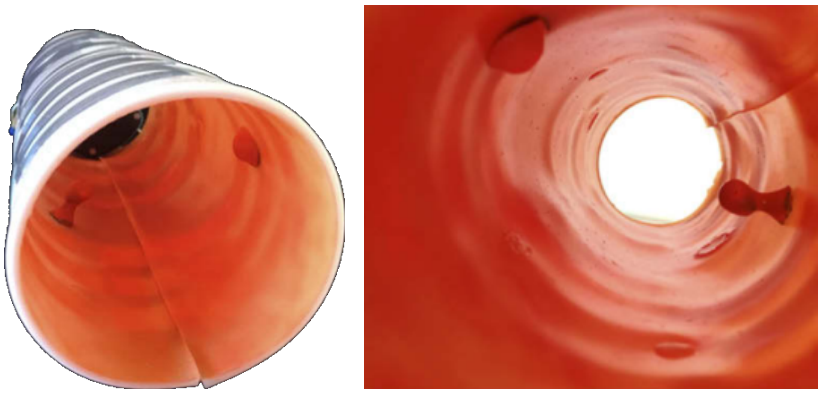
5.4.1 Experimental setup

The experiments used the STRAS robot follower side and a colon model. The STRAS robot is placed inside the colon model fig. 5.9 and performs the assisted colonoscopy.

The colon model used during the experiments was developed by Zulina et al. (2021). The physical dimensions of the colon emulate a colon swollen by air pressure achieved by surgeons inside the digestive system. The colon model is made of a PDMS/TiO2/Dragon Skin (Smooth-On Inc) and has polyps that are visually detectable as well as in-tissue lesions with mixed properties (visual features, surface geometry). The colon model used considers lesions with visual surface features and two types of lesions: hyperplastic polyps or adenoma sessile polyps. In the first type, the are visual features that allow suspicion from surgeons, which are evaluated by the OCT scanning tool. In the second type, the surface of the polyp is elevated and the lesion requires to be scanned and analysed to have a diagnosis. Further details of the colon model and the polyps can be found in Zulina et al. (2021).

The colon model allows geometry modification to evaluate multiple scenarios, e.g. using different orientations of the colon model. The different orientations allow the targeting of multiple lesions at different locations. A lesion has visual features that could allow visual identification by the surgeon but require a biopsy to assess the tissue. In the case of the considered experiments, multiple OCT scans can assess the tissue. The properties of the colon allow to have both visual and in-tissue unique properties properties that could be evaluated.

The three objectives defined in section 5.3.3 are validated individually and simultaneously. The colon model is modified in its orientation to evaluate multiple locations of



(a) (b)

FIGURE 5.9: Overview of the colon model. (a) shows the view from the outside of the surface oriented to build the colon model (b) inside view of the model with the different lesions (normal lesions and polyps).

the largest abnormal tissue, with up to 6 unique poses of the polyp. Placing the body of the STRAS at different locations allows for increasing the 6 unique poses, allowing for at least 2 different starting positions, creating 12 unique colons to be analysed. This allows for unique cases presented in each of the experiments reported in this chapter.

During the experiments (as depicted in fig. 5.7), a computer running a real-time Linux operating system handles the low-level control. The low-level controller PC governs the follower side of the STRAS system by commands of a high-level controller PC using RJ45 cable by a TCP/IP socket. Another PC handles the visual module. The analogue signal of the endoscopic camera is grabbed via a PCI grabber card. The image is segmented by a Python script running the local model with the U-Net outputting the positions of the arm $\boldsymbol{p}_O$, lumen $\boldsymbol{m}$ and lesion $\boldsymbol{s}$. The outputted positions are then passed through the average filter and the visual tracker during the training period at the start of the experiments. The positions are sent to the high-level controller via ROS (Quigley et al., 2009) which uses the QP controller and new positions are commanded to the follower. During the experiments, when the translation of the follower's arm is within the desired value of translation, the values of the gammas and k are set as table 5.1. When the translation is outside the desired valued, $\mathscr{B}_3$ is active, the γ_1, γ_2 and γ_3 are modified to 0.4, 0.3, 0.3 respectively.

TABLE 5.1: Parameters used in the experimental setup for assisted colonoscopy. Values of the weights γ and the proportional gains k for each objective active during semi-autonomous control mode.

Variable	γ_1	γ_2	γ_3	$k_{1,2,3}$
Value	0.5	0.5	0	0.1

5.5 Results

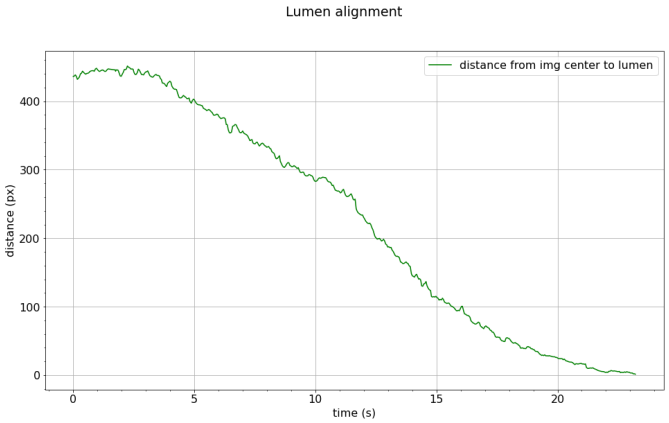
The proposed control framework is tested using the synthetic colon model. The results of two main experiments conducted are reported, one by individually testing the objectives, while the other tests the completed surgical scenario.

5.5.1 Individual validation of each of the tasks

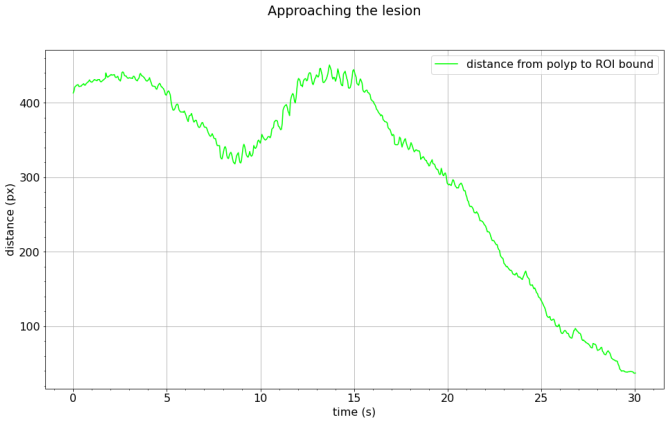
Each of the tasks is evaluated individually, fig. 5.10. The tasks are evaluated individually to validate that constraints of the physical control are handled by the control approach in real-time, due to the slack and non-linearities present in the control of FSIR. Objectives $\mathscr{B}_1$, $\mathscr{B}_2$ (eqs. (5.5) and (5.6)) are validated individually in fig. 5.10, while $\mathscr{B}_3$ (eq. (5.7)) is validated in the next subsection. The lumen alignment $\mathscr{B}_1$ is performed without the arm being in the scene. Multiple abnormal tissues are present on the screen and detected by the visual detection module. The image centre is aligned to the lumen centre, from the initial detected positions until a threshold of 7 pixels is fulfilled. In Figure 5.10(a), the distances of the different objectives are presented for one experiment. The distance between the image centre and the detected lumen centre is shown. In Figure 5.10(b) the tip of the OCT arm is aligned to a desired position in the image space. The distance between the desired potion and the tip of the arm is shown. Jointly the two images of fig. 5.10 present individual experiments on two cases for visual servoing in eye-in-hand and eye-to-hand IBVS.

5.5.2 Lumen alignment and arm in largest abnormal tissue

As described in section 5.4.1, the body of the STRAS system is placed at the beginning or the middle of the colon model $\mathscr{B}_2$, such that reaching the lesion requires approaching. The result of an experiment is shown in fig. 5.11. The green lime line depicts the distance between the OCT arm tooltip and the lesion. The green olive line depicts the distance



(a) Lumen alignment.



(b) Lesion reach.

FIGURE 5.10: Individual validation of the $\mathscr{B}_1$ and $\mathscr{B}_2$ objectives.

between the image and the lumen centres. The normalized translation of the body and arm are depicted in red and blue correspondingly. In fig. 5.11, the cyan highlight depicts when the arm set with the OCT scanning tool is deployed after lumen alignment, while the state compensating for arm translation is depicted in orange. Reaching the lesion is reported in fig. 5.11 as depicted by the orange highlight.

As shown in the individual results section 5.5.1, the separate control of the arm or body is performed correctly during the segments that are not highlighted. During the state that compensates for the translation of the arm and the body, it is shown that when the endoscope (red line) has a positive increment the arm (blue line) has a negative increment. A compensation in translations allows to effectively remain at the same distance from the surgical target, while incrementing the available workspace for the arm, effectively avoiding joint limits.

Figure 5.12 shows the results of another experiment performed with multiple objectives

active. During the OCT deployment, it is especially noted the issues for state estimation of the arms. During the state that compensates for translation, it is shown that the body translation is activated when the state is active, furthermore, the arm is retracted. After such movements, the arm translation is not activated as heavily during the rest of the experiments, while the distance towards the target is reduced, showing that rotation and bending are used to reach the target.

In the experiments reported on Figures 5.11 and 5.12 it is shown that the multiple targets ($\mathscr{B}_1$, $\mathscr{B}_2$ and $\mathscr{B}_3$) are active and collaborate to fulfil the expected behaviour during assisted colonoscopy. Two experiments are reported in figs. 5.11 and 5.12. The experiments show an average duration of 4.98s for the translation compensation which presents several increments in the two errors in both figures. In fig. 5.12 this increments of the distances depicted by the green-coloured lines in the orange highlight. The control framework acts on the arm and body which increases the errors to then be slowly reduced until the desired threshold is reached. Despite the instabilities of the visual module and the robot state, the errors performed during the translation compensation are handled correctly since backslash and non-linearities are present on the physical follower side.

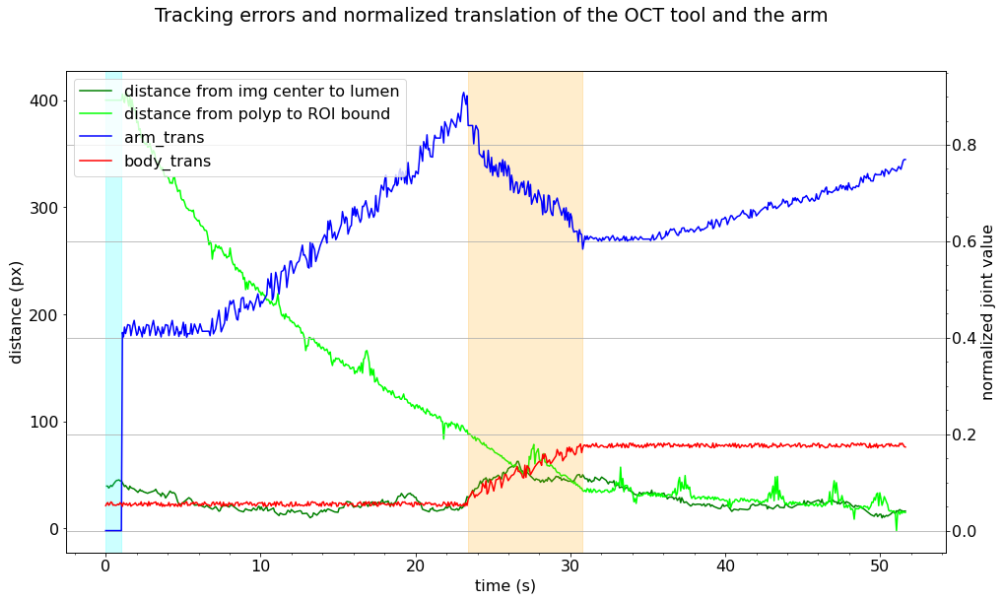


FIGURE 5.11: Experiment 1: Tracking errors during the experiment. In lime: distance of lumen to the image centre. In green: distance from the OCT tool to the centre of the region of interest. Orange highlight compensates for the translation in the arm (after reaching the maximum translation). Cyan highlight then the OCT is taken out of the channel to reach the desired image position.

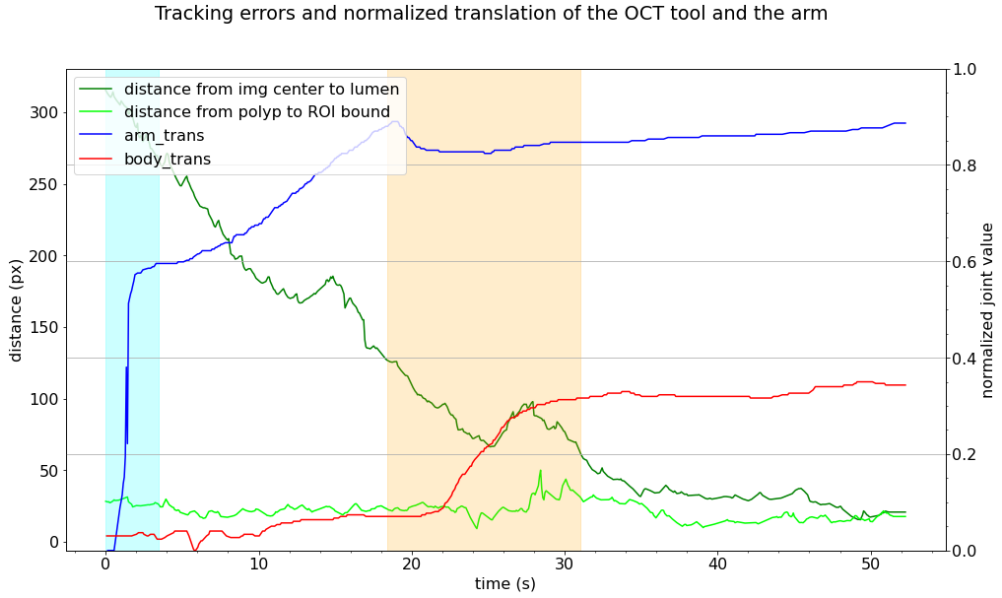


FIGURE 5.12: Experiment 2: Tracking errors during the experiment. In lime: distance of lumen to the image centre. In green: distance from the OCT tool to the centre of the region of interest. Orange highlight compensates for the translation in the arm (after reaching the maximum translation). Cyan highlight then the OCT is taken out of the channel to reach the desired image position.

5.6 Discussions and future work

The proposed approach developed in this chapter validates the control framework proposed in chapter 4 which enables coordinated control for flexible steerable intraluminal robots during the colonoscopy and biopsy to screen, detect and analyse the tissue in situ. The control framework is validated on the physical follower side of an FSIR platform, the STRAS system. The control framework in chapter 4 was extended to enable visual servoing for navigation and in-situ scanning in a colon model. This section presents and discusses the major outcomes of the study.

The control framework developed in chapter 4 formulates a QP controller that enables coordinated control in the dissection stage of ESD. In this chapter, the QP controller is reformulated to enable coordinated control during colonoscopy and biopsy, enabling simultaneous simultaneous eye-in-hand and eye-to-hand visual servoing. Visual servoing for the two modes is validated individually and simultaneously to enable coordinated control. Integration of the visual detection module and the QP controller enables the user to correctly reach the lesion that will be then analysed by OCT scanning. The control framework takes command of the 7DoFs required to perform coordinated control during colonoscopy and biopsy to require only user supervision.

Due to known non-linearities in the control of FSIR, some of the limitations of the approach are considered.

First, the control framework is evaluated under free space. The visual serving approach is tested only in free space, as tissue contact is complex to perform during visual serving due to the inherent challenges of estimating the robot's state. The limitations of this study are also present in other related works. Scanning via OCT has been performed previously performing automatic pull-back movements that ensure covering the tissue multiple times due to the automatic movements performed by the arm. The prior works have used downgraded scenarios to account for the complex real scenarios –physiological movement, visual obstructions and unwanted artefacts– and the challenges of the physical follower side –non-linear friction, backlash behaviour and interaction forces from arm-tissue interaction. Caravaca-Mora et al. (2020) tested a real-time OCT scanning tool mounted in the arm of a flexible steerable intraluminal robot. Zhang et al. (2020) used markers to track the position in the image space integrating the OCT scanning tool in one of the arms of FSIR to perform scanning while considering visual and orientation constraints. Furthermore, Caravaca-Mora et al. (2023) explores the capabilities of the scanning OCT tool developed for the STRAS system. Overall, the control framework proposed and tested uses a more realistic scenario because of the realistic colon model than other approaches with limited validation due to the difficult state estimation. The experiments, partially validate the control framework initially proposed in chapter 4 and chapter 5, the control framework proposed in chapter 4 is used with IBVS and allows assisted coordinated control. In the experiments, limits of the translation are avoided while maintaining the arm between a range of desired translations. This range is set by the maximum translation allowed and a lower bound, effectively setting a constraint in the arm translation. Jointly with the rest of the experiments, this approach allows to assist in the majority of the colonoscopies.

A second limitation of this study is that colonoscopy is not assisted during the completeness of the colonoscopy. The last step of the colonoscopy, assessing the tissue –benign or not– depends on an accurate estimation of the state of the OCT arm. The current approach lacks validation of this step due to the poor model estimation. Current approaches that used markers (Zhang et al., 2020) or that consider a controlled environment with a static background and no robot movement (Caravaca-Mora et al., 2023) have been published in the past. Prior works (Poignonec et al., 2020, 2023; Porto

et al., 2019) have an extensible look into precise control of the flexible structures of FSIR. Current works aim to use the visual features of arms (Sestini et al., 2021; Cabras et al., 2017) to accurately obtain the state of the arm of FSIR, improving control by high-level approaches as the one presented in chapters 4 and 5. The current state estimation provided by the visual detection module is not reliable enough to provide a stable input for the controller, and an updated state estimation provided by works such as Sestini et al. (2021) could allow using the proposed control framework to assist during the completeness of the colonoscopy. Automatic visual detection in surgical images is a complex task due to the ever-changing surgical environment originated from the insulation (Church, 2002; Yoshida et al., 2014; Dik et al., 2014) and the breathing (Fennerty et al., 1990; Waye et al., 1996; Saltzman et al., 2015). The current control framework integrates information from the multiple sensors available. The control framework allows assistance during complex procedures in colonoscopy and would be beneficial when a stable and reliable robot state can be obtained. By performing OCT pull-back with the current proposed approach the entire colonoscopy procedure will be carried out by flexible steerable intraluminal robots.

The third limitation of this approach is the static environment. Similar to the statement developed in section 4.8, the surgical environment is dynamic and will introduce unwanted interaction between the arms and body with the tissue. The visual detection of a module will require stable detection of the different elements of the control framework. A correct segmentation allows correct detection of the different actors, and correct tracking of the objects depends on the correct segmentation. Alongside this, accurate tracking allows for correctly performing the assisted colonoscopy. In the current experiments, the environment is controlled yet the accurate correct performance of the complete colonoscopy is difficult due to the high dependency on the multiple steps. Experiments not reported in this thesis include several cases where the detection was lost and handled by the filters and the control framework. This drawback is required to be considered in the future to provide stable output from the vision mode. As initially described in section 4.8, state estimation of the arm could improve the current kinematics and enable accurate control of the arm, yet tool-tissue interaction must be considered. In the case of complex surgical procedures such as ESD, the interaction might require more robust state estimation, e.g. electromagnetic tracking (Song et al., 2015) or multi-core Fiber Bragg Grating (FBG) (Ha et al., 2022). Providing a reliable

and accurate state for the arms and body will allow reliable and robust assistance during complex surgical procedures.

A fourth limitation of the approach is regarding the implementation. A computer handles the low-level control of the follower, while the vision module is handled by another computer and the high-level control is performed by another computer. Everything is orchestrated by ROS and the low-level controller. The control output is only active when the last command output has been reached or when the control output is above a given threshold, to not overflow the implemented position control in the low-level PC. This limitation was made due to taking extra precautions in the OCT due to the limited availability. The limitation could be solved by implementing a low-level controller in the high-level PC that manages the logic.

Another limitation is that the current experiments do not consider extended navigation. As navigation is present in just a small portion of the digestive system. The movements performed by the endoscopy body are only considered as re-position movements. This is due to two factors, one being the mobility capabilities of the STRAS system. The STRAS system sets a length of 900m with only 100mm of motorised translation, to extend this range the system is required to be manually inserted, to be mounted in the robotised cradle. This procedure is going to be solved in a new version with increased capabilities and improved dimensions. The second factor is the difficulty of performing navigation in the colon due to the complex environment and the multiple intricate direction changes. The colon changes direction after the rectum, the descending colon and the transverse colon, these changes in direction require users to master fluid management to properly navigate the GI tract. Correctly performing navigation of the colon through FSIR requires reducing unnecessary interaction between the bodies of FSIR and the surgical environment while improving available FSIR devices.

Chapter 6

Conclusion

Telemanipulating flexible steerable intraluminal robots is a mentally demanding task that requires users to deal with the $\mathscr{C}_1$**: coupled architecture** and $\mathscr{C}_2$**: high amount of degrees of freedom**. Due to the complexity of the FSIR systems, users avoid coordinated control of the arms and body, preferring single sub-system control. This means that instead of performing parallel control of the subsystems, the user executes sequential actions, which breaks the workflow of the surgical procedure. The sequential control of FSIR has several detrimental aspects. Sequential control may reduce awareness of the subsystems that are not being controlled at the moment, limiting patients' safety by unwanted contact between the arms and body with the lumen. Furthermore, sequential control increases the number of movements the body and arms have to perform due to the under- or over-shooting caused by unwanted movements due to the coupled architecture of the FSIR, which in turn calls the user's manual compensation. Sequential control hinders the correct performance of the surgical task by requiring users to perform extra arms and body movements, which are unnecessary and increase the amount of times tools interact with the tissue, increasing the risks associated with high amounts of tool-tissue interactions. The detrimental aspects of sequential control should be avoided by coordinated control. Coordinated control is especially required during complex surgical procedures increasing the cognitive load for surgeons due to the challenges of FSIR control ($\mathscr{C}_1$ and $\mathscr{C}_2$). Surgeons' performance suffers under high mental demand, leading to longer procedure times and increased error risk due to stress and fatigue-induced impairment affecting control precision. As such, control algorithms for the FSIR surgical system should be designed to avoid hindering surgeon performance or

diverting their attention to more important tasks instead of struggling with the system control.

In this thesis, the research question (initially defined in section 1.5, together with the four main objectives $\mathcal{O}_2$, $\mathcal{O}_1$, $\mathcal{O}_3$, $\mathcal{O}_4$) is formulated as: **How can a control framework assist users to manage the multiple degrees of freedom, coupled architecture, and complex tool-tissue interaction to enable accurate execution of demanding surgical procedures?**. This question is answered by modelling the problem ($\mathcal{O}_2$) based on a QP-based controller that assists ($\mathcal{O}_1$) users during technically difficult surgical procedures which involves multiple coupled end-effectors. The approach governs the movements of an FSIR platform in a safe manner that correctly ($\mathcal{O}_3$) and safely ($\mathcal{O}_4$) performs the required surgical gestures during specific times. The control framework proposed in this thesis focuses on coordinated control during the *dissection stage of ESD* and during *colonoscopy and biopsy* primarily (although it could also be useful for less complex surgical tasks, e.g. marking during ESD or OCT pullback which require coordinated control involving multiple DoFs). Similarly to Zhang et al. (2021), the control framework developed in this thesis is designed using the task-function approach using a weighted QP formulation, which is an efficient solution for handling multiple tasks and constraints in various frames of reference. The main contributions of this thesis are then to model the clinical features of each of the two problems and formulate them as optimization objectives and constraints in the proposed framework. The proposed solution is then implemented on the STRAS robot (described in chapter 3) to validate the proposed approach. For the dissection stage of ESD (chapter 4), this is done through a user study in simulation, with a thorough statistical analysis showing an advantage of the proposed control scheme over manual telemanipulation. In the colonoscopy and biospy case (chapter 5), validations are performed in a realistic colon model and automatic execution of the task using visual servoing, again formulated using a QP-based task function approach[1].

The control framework developed for the first surgical scenario, the dissection stage of ESD, considers multiple reaching tasks for all of the subsystems that are satisfied in parallel —which will be mentally demanding and require high numbers of training hours to be performed by manual telemanipulation. During the dissection stage of ESD, coordinated control of the 10 DoF is required and thanks to the developed control

[1] a list of this thesis publications can be found in chapter 7

algorithm, the user needs only to take care of 3 DoFs out of the 10 DoFs, while the others are regulated automatically to accommodate the task. In comparison to manual telemanipulation, the control framework allows users to correctly perform the dissection stage of ESD which is represented by multiple reaching objectives for the two arms.

In order to evaluate the two control modes, non-assisted teleoperation and assisted control mode in a controlled environment, a simulator of an FSIR platform, the STRAS system, is developed. The virtual simulator simulates the colon model and the STRAS system. Both control modes are tested by non-experts that carry out the experiments, after a short training time in which users get familiar with both control modes and the simulator. The experiments –and the training period– required users to perform multiple reaching tasks that emulate positions the arms will reach during the dissection stage of ESD. The user experiments were followed by an evaluation of performance, where several metrics were evaluated. Statistical tests were conducted to find the statistical difference between the performance of the manual and the assisted mode. The metrics were separated into three categories: duration, surgical and kinematic metrics. Results show that the surgical and kinematic metrics perform better with the assisted mode at a duration cost. The assisted mode enables correct coordinated control with overall better results –with regards to the number of movements of the body, and smoother movements of all subsystems– while being able to correctly perform the surgical task –by the non-violation of the reaching task. In addition, the participants were asked to fill in the NASA TLX questionnaire to provide a subjective evaluation of users' opinions, which pointed out that users didn't find a large difference between the assisted and manual modes. Given the fact that novice users with and without clinical backgrounds perform the experiments, the lack of added complexity to the control is a successful outcome from the experiments.

Metrics and user opinion also highlight some shortcomings. While the overall performance of the assisted control mode outperforms in the surgical and kinematic metrics it is performed at a penalty cost. The user takes longer to perform the experiments with the assisted mode, which is due to several factors. First, the user sets unreachable positions of the arms at the current body position. This is a limitation of the approach which does not have context about the target positions and if they are reachable, such cases increased the duration of outlier cases that overall affect the duration metrics. Another limitation is regarding the lack of awareness of the surgical target (reaching positions)

which can lead users into performing unwanted movements, which impact the overall duration and some of the kinematic metrics. The lack of awareness in the system can cause undesired body movements (false positives), increasing the user's frustration. Due to the false positives and lack of economy of movement, users are not able to improve the duration metrics. The second factor is regarding the user frustration in the NASA TLX questionnaire. This is because the short period of training and experiments prevents users from reaching the plateau of performance and the lack of expertise frustrates the user-assisted control. The frustration under the NASA TLX has to be carefully considered, as the reaching task emulates the basis of the surgical task and it is not as mentally demanding as ex- or in-vivo experiments in which bleeding and safety play a huge role in the user's cognitive load. The proposed framework performs better in the surgical and kinematic metrics, yet the duration metrics have an area of opportunity.

Another solution to the shortcoming could be path prediction for the tip of the arms (Pasini et al., 2023; Amini Khoiy et al., 2016; Kossowsky and Nisky, 2022). Predicting the approximate path of the arms (Wagner et al., 2021) to better suit arm-body movements can be implemented to anticipate imminent body movements, as tools get close to the limits of the operational workspace or specific regions in the task space. Path prediction can be considered by the awareness or by feature vectors (position, speed, orientation, the position of the targets, distance to the target, among others) (Kossowsky and Nisky, 2022) that allow to predict if the tool will be leaving the desired area or not, reducing the number of false-positive cases. Context-dependent tasks and awareness could link specific behaviours to different steps of the surgical task, e.g. how tension evolves over the entire dissection or how the tool approaches certain lesions based on their location. Another case of context-awareness could be the estimated distance between the objectives and the tip of the arms. If the tip of the arm and the target are within reach, no body movement should be triggered and the user has free telemanipulation of the arms. Future work requires dealing with the optimal level of awareness that enables correct assistance to better suit complex surgical procedures. Close consideration of how the context information impacts false positives without increasing the false negatives and without impacting the surgical, kinematic and duration metrics must be taken. An interesting concept will be to consider the level of training[2] with respect

[2]since multiple users have tested one or more versions of the control approach of chapter 4

to the level of assistance that better suits them, and how they can impact the training performance to better stabilise the learning plateau.

The control framework also requires physical validation on the follower side of FSIR. The proposed control framework chapter 4 is translated into the second surgical scenario, colonoscopy and biopsy and it is a physical validation of the essence of the control framework. The assistance for colonoscopy and biopsy relies on a visual detection module that segments and detects the lumen, the lesion and the OCT probe arm to enable visual servoing inside a colon model. The control framework deals with multiple objectives and performs individual and simultaneous eye-in-hand and eye-to-hand visual servoing while managing the challenges of FSIR platforms. Simultaneous visual servoing is analogous to coordinated control and serves as a validation of the control framework. The high-level control framework commands a low-level controller that governs the follower side. The user is not required to take any action during the experiment, allowing for increased assistance. The approach developed in chapter 5 uses the information of the screen and the current state of the robot for driving the control framework. The visual detection module is the first approach to context awareness in the telemanipulation of FSIR. Consideration of the limitations (Poignonec et al., 2020; Barrientos-Diez et al., 2021; Poignonec et al., 2023; Cursi et al., 2022) for precise control of FSIR should be taken in future works.

Furthermore, special effort has to be considered to develop a solution that includes the context of the surgical task, and visual and kinematic information, e.g. when the positions of the arms are in contradictory directions or when a position for the arms is only reachable at specific body positions. Open research questions in regards to path planning for the arms with respect to the next position of the body is set to be a hard problem due to inherited dependencies of the subsystems. Estimating how the interactions affect the state of the robot could be included to increase the accuracy of the tools. Accurate state estimation through added sensors to the channels of FSIR robots could improve the state estimation and could even allow 3D registration. The information from the OCT tool and an accurate body and arm state could be used to estimate the position of the lesion, which increases the amount of information the system is capable of generating and analysing in situ.

Jointly, the two main contributions of this thesis reduce the gap between manual telemanipulation and autonomous surgery by means of multiple levels of assistance. The developments of chapter 4 explore assistance by autonomously controlling the system on specific circumstances and user commands. While chapter 5 explores autonomous and assisted tasks to enable colonoscopy and biopsy through autonomous tasks with user supervision. The control framework developed handles the challenges of telemanipulating FSIR by modelling complex surgical procedures into reaching tasks for the subsystems. To enhance the duration metrics and reduce the mental demand the control frameworks developed require including context-awareness by visual detection modules or registration information that allows moving the body as required.

Chapter 7

List of publications of this thesis

Internationally review academic journals

- **Gonzalez-Herrera, F.**, Nageotte, F., Zanne, P., Borghesan, G., de Mathelin, M., Vander Poorten, E., & Rosa, B. (2024) A semi-autonomous control mode for flexible steerable intraluminal platforms, IEEE Transactions on Medical Robotics and Bionics. In press, DOI 10.1109/TMRB.2024.3385990

International scientific conferences and symposia, published in full in proceedings

- **Herrera, J. F. G.**, Nageotte, F., Borghesan, G., Vander Poorten, E., & Rosa, B. (2023). Semi-autonomous control for endoluminal robotic platforms. In Computer and Robot-Assisted Surgery (CRAS), Paris, France, September 2023
- **Herrera, J. F. G.**, Pore, A., Sestini, L., Sahu, S. K., Liao, G., Zanne, P., ... & Gora, M. (2022). Autonomous image guided control of endoscopic orientation for OCT scanning. In Computer and Robot-Assisted Surgery (CRAS), Naples, Italy, April 2022
- Liao, G., **Herrera, F. G.**, Zhang, Z., Pore, A., Sestini, L., Sahu, S. K., ... & Gora, M. J. (2022, May). Autonomous OCT volumetric scanning with robotic endoscope. In Clinical Biophotonics II (p. PC1214602), SPIE

Appendix A

NASA Task Load Index

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Appendix B

Extra concepts

B.1 Smoothness by spectral arc length

The spectral arc length ($SPARC$) smoothness is a revised and improved version of the spectral arc length measure (SAL) (Balasubramanian et al., 2011) which estimates smoothness by calculating the arc length of the Fourier magnitude spectrum within the frequency range 0 to 20 Hz of a given speed profile $v(t)$. To compute the smoothness (Balasubramanian et al., 2015) via spectral arc length ($SPARC$) metric for measuring smoothness multiple steps are required.

1. Segment speed profiles, $\{v[n], n \in \{0, \ldots, N-1\}\}$.

2. Compute Fourier magnitude spectrum. Compute the K-point fast Fourier transform (FFT) magnitude spectrum of the speed profile by padding $(K-1)$ zeros to the speed profile. $V[k] = |FFT(V_{zp}[np])|, k \in \{0, \ldots, K-1\}$
   $v_{zp}[n] = \begin{cases} v[n] & 0 \leq n \leq N-1 \\ 0 & N \leq n \leq K-1 \end{cases}$
   $K = 2^{roundup(\log_2 N)+4}$

3. Normalize Fourier magnitude spectrum with respect to its DC value.
   $\hat{V}[k] \triangleq \frac{V[k]}{V[0]}$

4. Choose the cut-off frequency.
   $K_c \triangleq \min\{K_c^{max}, \min_K\{K|\hat{V}[r] < \bar{V} \forall r > K\}\}$

5. Compute smoothness. The formula of smoothness is applied.

$$SPARC \triangleq -\sum_{k=1}^{K_c-1} \sqrt{\left(\frac{1}{K_c-1}\right)^2 + \left(\Delta \hat{V}[k]\right)^2} \tag{B.1}$$

An implementation code is available at https://github.com/siva82kb/smoothness as part of the work described in Balasubramanian et al. (2015).

B.2 Effect size

Effect size (Kelley and Preacher, 2012) is a value that measures the strength of the relationship between two variables in a population, or a sample-based estimate of that quantity. Effect size completes statistical hypothesis testing and is an essential component when evaluating the strength of a statistical claim.

The common language effect size A_w is used for nonparametric distributions, and its equation expresses effect size on the basis of the probability that a random observation of population **p** scores higher than a random observation of population **q**—that is,

$$A_w = \sum_{i=1}^{n_p} \sum_{j=1}^{n_q} \frac{[\#(\mathbf{p}_i > \mathbf{q}_j) + 0.5\#(|\mathbf{p}_i - \mathbf{q}_j| \leq \varepsilon)]}{m\,n} \tag{B.2}$$

where # is the count function, p and q are vectors of scores for the two samples, and n_i is the sample size in the group $i = p, q$. ε is a small threshold which allows handling the case of continuous data (Vargha and Delaney, 2000). Let's consider an example with $p = \{5, 7, 6, 5\}$ and $q = \{3, 4, 5, 3\}$. In this case, it is straightforward to compute $A = (3.5 + 4 + 4 + 3.5)/16 = .9375$, meaning that there is a 93.75% chance that the observation would be higher for a randomly selected member of group **p** than for a randomly selected member of group **q**. Ruscio (2008) found that the nonparametric A_w was generally accurate and suggested that researchers and practitioners should report this measure.

Further details about the correct use for the effect size can be found in Li (2016); Vargha and Delaney (2000); Fritz et al. (2012).

Appendix C

Résumé en français

C.1 Introduction

Le cancer colorectal (CCR) occupe le quatrième rang mondial des cancers en termes d'incidence (19,5 %) et de mortalité (9,0 %). Le cancer colorectal provient généralement d'amas anormaux de cellules appelés polypes. Les polypes peuvent se développer dans le système digestif jusqu'à former des métastases, puis se propager à d'autres organes et systèmes du corps. Pour prévenir les métastases, les polypes colorectaux (PRC) sont détectés aux premiers stades de leur développement grâce à une coloscopie. La coloscopie permet d'inspecter visuellement les parois du côlon à la recherche de polypes. Lorsque les gastro-entérologues soupçonnent la présence d'un tissu malin, ils en prélèvent un échantillon qui sera analysé dans le cadre d'une biopsie. L'analyse détermine le degré de malignité et des procédures d'ablation des polypes sont effectuées le cas échéant, ce qui permet d'enrayer efficacement le développement du cancer colorectal.

Pour éviter les taux élevés d'incidence et de mortalité du RCC, les chirurgiens peuvent utiliser des endoscopes flexibles qui permettent d'inspecter le système digestif. Des prototypes récents d'endoscopes flexibles permettent de monter des outils chirurgicaux sur leur extrémité, ce qui permet aux chirurgiens d'interagir avec les tissus et d'enlever les tissus suspects. Les endoscopes flexibles munis d'outils à leur extrémité nécessitent que plusieurs utilisateurs contrôlent leur corps et leurs bras. Leur contrôle peut être mentalement exigeant pour les chirurgiens. Le stress et les exigences mentales peuvent nuire aux performances des chirurgiens et à la santé des patients.

Par le passé, le standard pour l'ablation des polypes était l'accès via la cavité abdominale. De petites incisions sont réalisées dans la paroi abdominale pour y introduire des instruments chirurgicaux qui sont utilisés pour procéder à l'ablation des parties du colon contenant des polypes. Ces interventions sont très inconfortables pour le patient, elles présentent un risque élevé d'infection en raison des multiples incisions et elles requièrent l'ablation inutile de larges portions du côlon. Les patients doivent également rester longtemps à l'hôpital et faire l'objet d'une surveillance continue. Pour pallier ces inconvénients, il est possible de traiter les polypes en passant par les orifices naturels du patient. Des endoscopes flexibles peuvent être insérés par l'anus du patient pour atteindre le PRC.

Des capacités robotiques ont été développées pour les endoscopes flexibles. Les robots intraluminaux flexibles et orientables (De l'anglais, flexible steerable intraluminal robots, FSIR) sont l'évolution des endoscopes flexibles dotés de capacités robotiques. Les robots intraluminaux flexibles et orientables permettent à un seul utilisateur de contrôler le corps et les bras du robot. L'utilisateur unique doit prendre en compte tous les mouvements de ces plates-formes. La complexité de ces plateformes et les procédures chirurgicales complexes sont les deux principaux défis à relever. Lors de la télémanipulation manuelle de robots intraluminaux flexibles et orientables, les chirurgiens doivent relever les défis afin d'exécuter correctement les procédures chirurgicales complexes.

L'objectif de cette thèse est d'aider la télémanipulation de robots intraluminaux flexibles et orientables à relever les défis tout en permettant une exécution correcte des procédures chirurgicales.

C.2 Robots intraluminaux flexibles et orientables

Ces plateformes sont télémanipulées dans une architecture leader-suiveur. Le côté suiveur est placé à l'intérieur du corps du patient tandis que le côté leader se trouve dans la salle d'opération. Les bras et le corps sont les deux principaux sous-systèmes qui constituent le côté suiveur de la FSRI.

Les robots intraluminaux flexibles et orientables sont composés d'un corps principal avec 4 degrés de liberté ou plus (de l'anglais, degree of freedom, DOF). À l'extrémité du corps, deux canaux sont dotés de bras chirurgicaux avec des instruments à leur

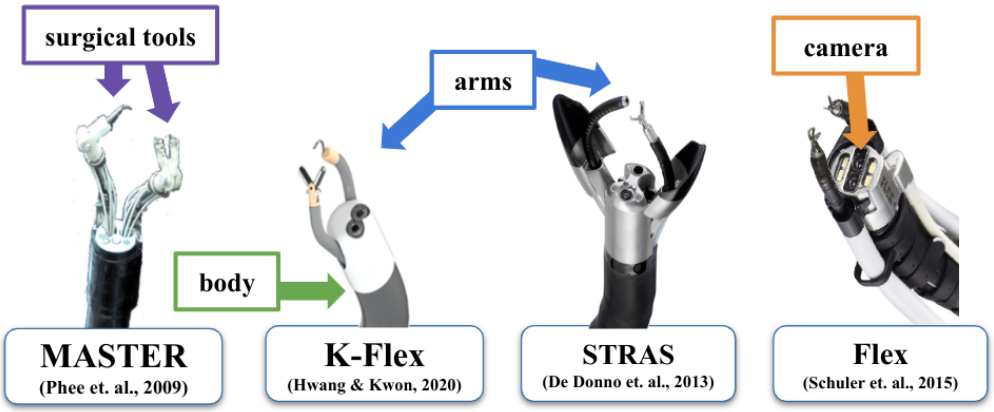


FIGURE C.1: Architecture commune aux plateformes FSIR.

extrémité. Chaque bras a 3 DOF et la pointe qui est montée par des outils chirurgicaux peut être activée. À l'extrémité du corps, ces plateformes sont dotées d'une caméra qui permet de visualiser les bras. Les plateformes représentent au moins 10 DOF. Plusieurs plateformes, telles que STRAS (Nageotte et al., 2020), MASTER Hwang and Kwon (2020b), K-FLEX (Schuler et al., 2015), suivent cette architecture.

Ces plateformes sont contrôlées par les utilisateurs au moyen de l'interface leader. Le côté leader est constitué de poignées, de joysticks et de pédales qui traduisent les commandes du chirurgien au côté suiveur dans le système digestif.

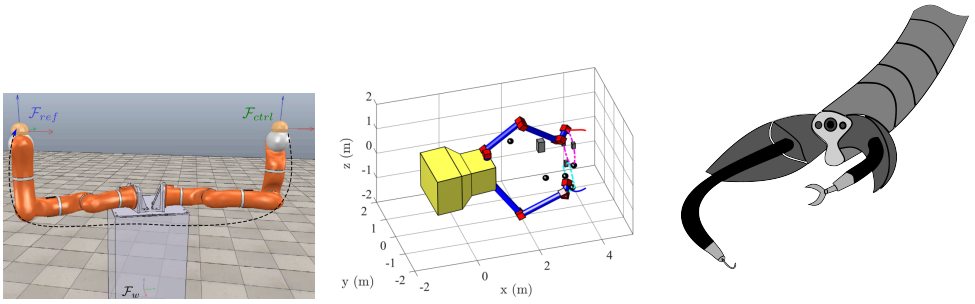
C.3 Problème

L'architecture robotique complexe du FSIR est difficile à contrôler pendant les procédures chirurgicales complexes qui nécessitent de contrôler le corps et les bras (fig. C.4(a)). Le contrôle simultané du bras et du corps est appelé contrôle coordonné. Lors des procédures d'ablation, telles que l'étape de dissection de la sous-muqueuse endoscopique (de anglais, endoscopic submucosal dissection, ESD), le contrôle coordonné est particulièrement nécessaire. Au cours de ces procédures, le bras et le corps doivent être contrôlés simultanément afin d'effectuer correctement les tâches chirurgicales.

Le contrôle coordonné d'un bras et du corps exige des chirurgiens qu'ils contrôlent 7 DOF, les deux bras et le corps nécessitant 10 DOF. Lors de procédures complexes, il est difficile de contrôler tous les DOF et d'effectuer correctement la tâche chirurgicale. Afin de préserver la sécurité du patient, les chirurgiens évitent le contrôle coordonné

et préfèrent le contrôle séquentiel. Le contrôle séquentiel peut nécessiter davantage de mouvements de la part des bras et du corps et interrompt le flux de travail chirurgical. Ces interruptions augmentent la quantité d'interaction entre l'outil et les tissus, ce qui n'est pas souhaitable et accroît l'inconfort du patient.

L'assistance peut permettre de réduire le nombre total de DOF pendant la commande coordonnée, tout en maintenant les mêmes mouvements du côté suiveur. Un mode assisté doit contrôler les multiples DOF tout en tenant compte de la tâche individuelle exécutée par chaque sous-système. Les tâches de tous les sous-systèmes doivent permettre au chirurgien de passer en douceur à l'exécution des tâches chirurgicales. Le mode assisté nécessite de contrôler plusieurs DOF qui sont couplés dans une base mobile.



(a) (b) (c)

FIGURE C.2: Exemples de structures robotiques bras-corps similaires tirés de la littérature. a) Robot humanoïde à deux bras dans le haut du corps (Tarbouriech et al., 2022), b) simulation de robot spatial (Cai et al., 2022) c) STRAS robot

Dans d'autres plateformes robotiques dotées de bras indépendants et d'une base mobile, l'assistance ou l'automatisation a été réalisée de multiples façons. Deux approches principales sont évaluées. Les tâches multiples pour les multiples effecteurs finaux peuvent être gérées par la projection dans l'espace nul. Dans la projection dans l'espace vide, les tâches multiples sont évaluées avec une hiérarchie stricte.

Une autre proposition consiste à utiliser des approches d'optimisation. Avec les approches d'optimisation, il n'y a pas de hiérarchie stricte. Une solution efficace pour traiter les tâches consiste à utiliser l'approche tâche-fonction. Cette approche permet de définir des tâches positionnelles qui sont traitées par des fonctions d'optimisation pouvant être résolues par des solveurs numériques. Pour gérer les comportements conflictuels

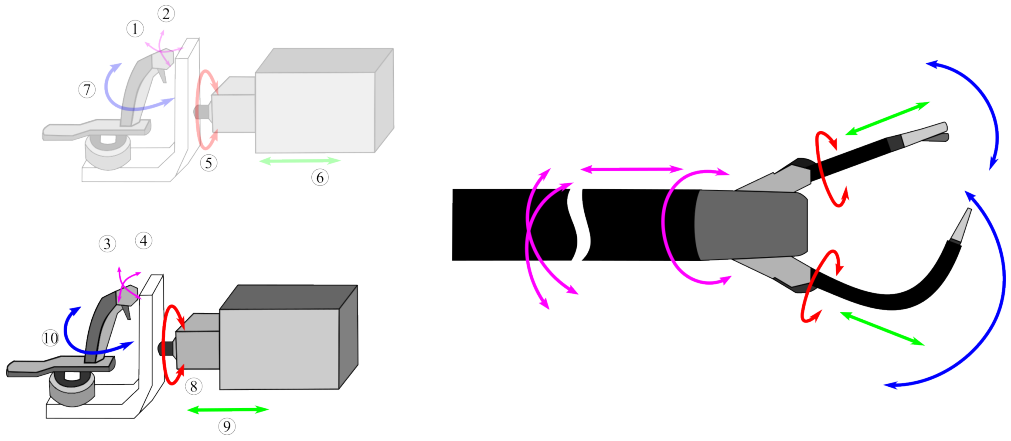


FIGURE C.3: Vue d'ensemble du robot STRAS. Le côté leader est à gauche de l'image, le côté suiveur est à droite.

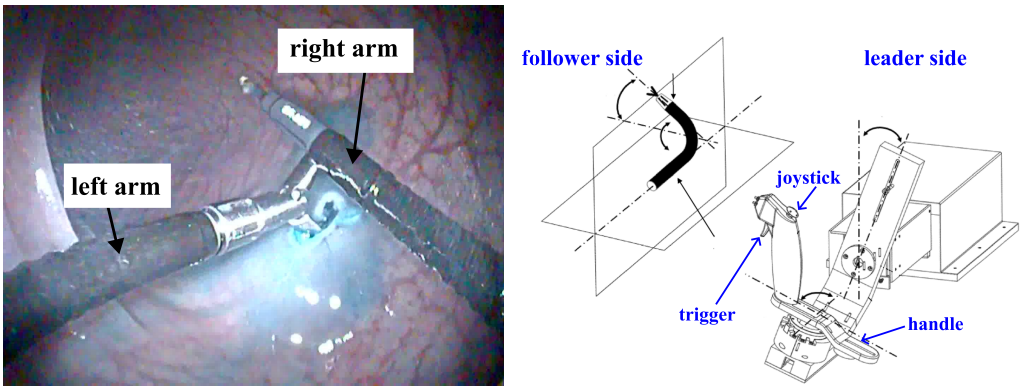
entre les tâches, des pondérations peuvent être introduites. Une approche pondérée permet d'éviter la hiérarchie et d'imposer à la place des mouvements des multiples sous-systèmes qui permettent un comportement global de tous les effecteurs finaux.

Le cadre contrôlé proposé est développé sous le robot STRAS (fig. C.3) qui partage une architecture similaire entre plusieurs plateformes. Le robot STRAS a 10 DOF dont 4 DOF pour le corps, avec déflexion dans deux directions, rotation et flexion. Les bras peuvent se plier, pivoter et être insérés. Les 10 DOF du système sont augmentés par les pédales et les actionneurs des outils chirurgicaux à l'extrémité des bras. Les 10 DOF sont mappés sur la console leader, qui est mappée pour émuler les positions des bras du chirurgien (fig. C.4(b)).

C.4 Proposition

La proposition de cette thèse est de créer un cadre de contrôle qui permet l'assistance à la télémanipulation de FSIR. Le cadre de contrôle commande tous les sous-systèmes (bras et corps) tout en réduisant le nombre total de DOF qui doivent être contrôlés par le chirurgien. L'approche agit sur des cas spécifiques lors de procédures chirurgicales complexes qui nécessitent un contrôle coordonné. Deux cas chirurgicaux principaux sont envisagés. Le premier est celui où un contrôle coordonné est nécessaire pour assister la télémanipulation. Dans ce cas, les mouvements de la poignée latérale du leader sont transmis au corps tandis que le cadre de contrôle commande l'ensemble de la plate-forme, ce qui facilite la transition entre l'assistance et le contrôle par l'utilisateur. L'autre cas

chirurgical est celui d'une automatisation complète. Dans ce cas, un contrôle coordonné est nécessaire pour accomplir des tâches chirurgicales spécifiques. Le cadre de contrôle est développé dans le cadre de l'architecture commune du FSIR. Le cadre de contrôle est évalué pour assister ou effectuer des tâches chirurgicales autonomes.



(a) (b)

FIGURE C.4: Vue d'ensemble des bras à partir de la caméra corporelle. À gauche : vue de la caméra endoscopique. A droite : Vue d'ensemble du bras du côté du suiveur et du leader.

En règle générale, la solution proposée repose sur une formulation basée sur l'optimisation en utilisant l'approche de la fonction de la tâche. Des poids sont introduits pour équilibrer l'approche coordonnée pour les multiples sous-systèmes. La tâche est interprétée comme la définition de positions objectives pour les effecteurs au cours des tâches les plus complexes. Les pédales de la plateforme robotique sont utilisées pour assister le chirurgien lors des multiples étapes de ces procédures complexes. Les pédales permettent au chirurgien de déplacer la plate-forme et d'ancrer la position sur un bras. Plusieurs poids sont définis pour les différents comportements favorisant les multiples objectifs définis.

Figure C.5 présente une vue d'ensemble du cadre de contrôle proposé. Dans le cas chirurgical de l'étape de dissection de l'ESD, le mode de contrôle assisté est régi par un leader virtuel qui commande le corps et le bras, tandis que le cadre de contrôle prend en charge les DOF du suiveur dans le simulateur. Dans le scénario chirurgical de la coloscopie et de la biopsie, le nombre total de DOF est contrôlé par le mode assisté.

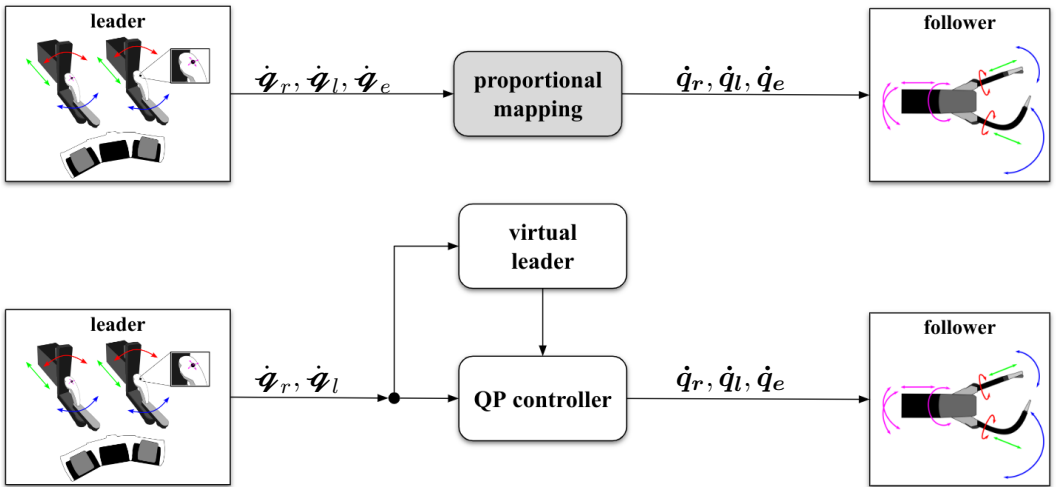


FIGURE C.5: Aperçu du cadre de contrôle. En haut, télémanipulation manuelle (eqs. (3.23) and (3.25)), sur le bas, télémanipulation assistée par le mode de contrôle semi-autonome bras-corps. Lorsqu'il n'y a pas de mouvement du corps dans le mode de contrôle semi-autonome bras-corps, cas 0, les bras $\dot{q}_r$, $\dot{q}_l$ sont régis par la console du chef $\dot{q}_r$, $\dot{q}_l$. Lorsque les cas 1 et 2 sont actifs, l'ensemble de la plate-forme est contrôlé par le mode assisté, qui déplace le corps indirectement par les mouvements des bras, comme indiqué ci-dessous eq. (4.26). Dans le cas 1, le bras non dominant n'est pas ancré $\dot{q}_r, \dot{q}_e$ as the output comme sortie du contrôleur QP. Dans le cas 2, les deux bras seront contrôlés, en ancrant le bras non dominant comme le souhaite la traction pendant la phase de dissection de l'ESD, tout en contrôlant en toute sécurité la position du bras dominant.

C.5 Résultats

Expériences réalisées dans le cadre des deux scénarios chirurgicaux. Le premier scénario chirurgical, l'étape de dissection de l'ESD, est testé dans un simulateur avec le système de chef physique. Dans le simulateur, l'utilisateur doit atteindre des positions qui reproduisent celles que les bras atteindraient pendant la procédure chirurgicale réelle. Trois tâches d'atteinte sont définies. L'utilisateur contrôle le système simulé avec l'approche assistée pour déplacer les sous-systèmes et atteindre le premier objectif avec le bras non dominant. Ensuite, sans enfreindre le premier objectif, l'utilisateur atteint les deux objectifs restants, ce qui nécessite de déplacer les sous-systèmes simultanément. La procédure complexe est assistée par le cadre de contrôle. Pendant la télémanipulation manuelle, l'utilisateur doit tenir compte de l'architecture couplée, des multiples DOF et des multiples tâches pour les effecteurs finaux. Les paramètres relatifs à la performance et au temps sont enregistrés au cours des essais des deux modes de contrôle. Le mode de contrôle assisté donne de meilleurs résultats en ce qui concerne les paramètres de la tâche chirurgicale, avec une trajectoire plus douce et une légère pénalité en termes de temps (fig. C.6). En mode de commande assistée, la tâche du bras non dominant

n'est pas violée comme en mode manuel. Dans le scénario réel, cela nécessiterait des mouvements supplémentaires des bras, ce qui augmenterait le nombre de fois où le tissu est saisi.

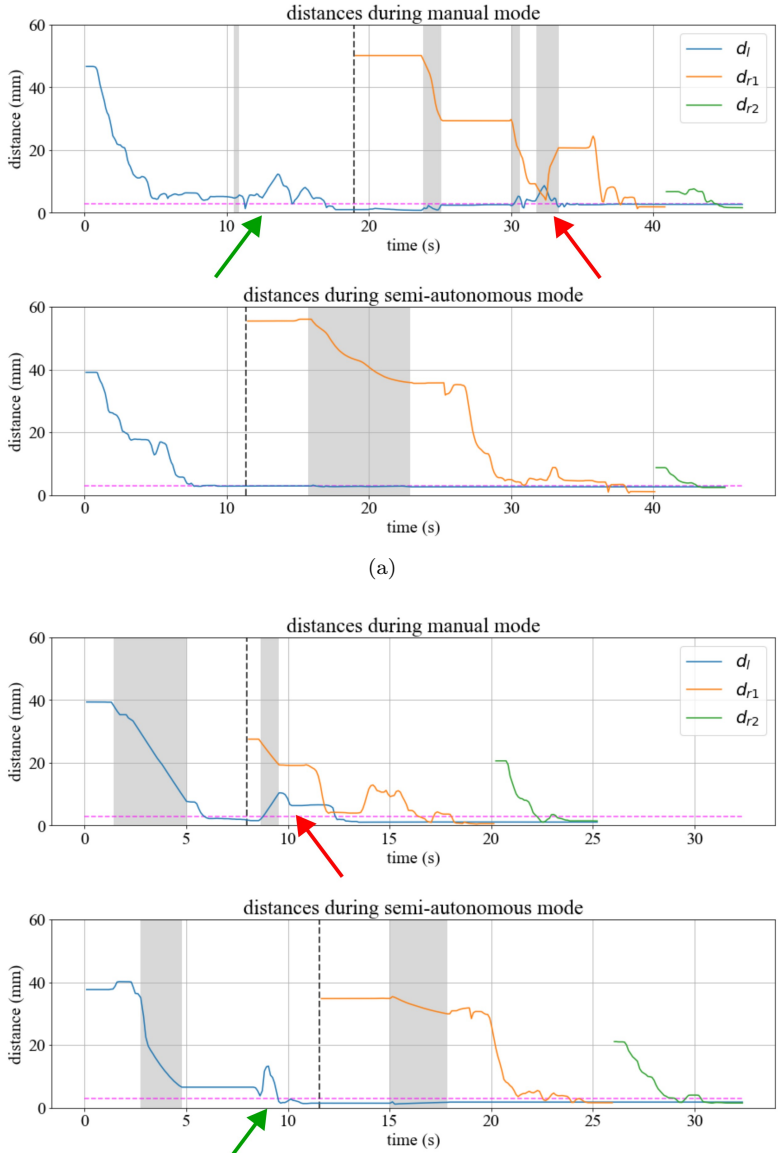
Pour le scénario de la coloscopie, l'automatisation complète de la tâche est mise en œuvre, en envisageant de s'approcher de la zone suspecte et de déployer un outil de balayage pour évaluer le tissu. La procédure nécessite le contrôle coordonné de plusieurs sous-systèmes, le bras et le corps, tout en effectuant un suivi visuel. L'utilisateur doit suivre la lumière, le tissu suspect et les bras pour approcher et déployer l'outil de balayage. Dans les expériences réalisées, l'approche de contrôle segmente et suit visuellement la lumière, le bras et le tissu pour effectuer un asservissement visuel à l'aide de la position de l'image. L'outil de balayage est ensuite déployé à proximité du tissu et l'expérience est terminée. Le contrôleur gère l'architecture couplée et définit des problèmes d'optimisation similaires pour le contrôleur, de sorte que de multiples expériences ont été réalisées de manière autonome par le contrôleur (figs. C.7 and C.8).

Les résultats obtenus dans le cadre des deux scénarios chirurgicaux ont permis d'évaluer le cadre de contrôle proposé dans le cadre d'un suivi physique et d'un simulateur.

C.6 Conclusion

Le contrôleur basé sur le QP permet d'assister et d'effectuer des tâches automatiques par une interprétation géométrique d'une tâche chirurgicale complexe. L'architecture couplée et le grand nombre de DOF sont gérés, ce qui permet aux utilisateurs de se concentrer sur la tâche chirurgicale tandis que le contrôleur relève les défis de la télémanipulation FSIR.

Les travaux futurs devront inclure des tests plus réalistes avec le côté suiveur, au lieu de se contenter d'une évaluation sur le simulateur. Ces tests devraient reformuler l'approche de contrôle pour tenir compte de l'état imprécis du robot et traduire la tâche de positionnement en contrôle de force. Les non-linéarités du robot FSIR rendent la traduction du cadre proposé difficile, en particulier lorsque les bras et les tissus interagissent. Des approches visuelles ont été développées pour estimer l'état du robot. Une intégration correcte devrait incrémenter la précision du modèle et utiliser à la fois l'état cinématique et l'état de l'image pour agir avec précision sur le suiveur physique.



(a)

(b)

FIGURE C.6: Exemple de distances entre les bras et les objectifs en mode manuel (en haut) et en mode semi-autonome de contrôle bras-corps (en bas). a) et b) sont des essais réalisés par deux utilisateurs différents. Les mouvements du corps sont représentés en gris. $\boldsymbol{d}_l$ est la distance entre le bras non dominant et son objectif ${}_w\boldsymbol{p}^l_d$, $\boldsymbol{d}_{r1}$ (resp. $\boldsymbol{d}_{r2}$) est la distance entre le bras dominant et son premier objectif ${}_w\boldsymbol{p}^r_{d1}$ (resp. deuxième objectif ${}_w\boldsymbol{p}^r_{d2}$). Les distances ne sont représentées que pour les étapes pertinentes. Les seuils de réalisation des tâches d'atteinte sont représentés par la ligne horizontale en pointillés roses, de 3 mm. La ligne verticale en pointillés représente la transition entre le stade I et le stade II de la dissection. Les pointes de flèches rouges représentent la violation de la tâche chirurgicale après l'intervention chirurgicale ${}_w\boldsymbol{p}^l_d$ a été atteint (après la ligne verticale en pointillés noirs).La pointe de flèche verte indique une augmentation de la distance de ${}_w\boldsymbol{p}^l_d$ mais avant que l'objectif d'atteinte par le bras non dominant ne soit atteint.

L'interprétation géométrique des tâches chirurgicales permet de définir des objectifs pour le contrôleur, qui sont activés par la commande de l'utilisateur. La prévision

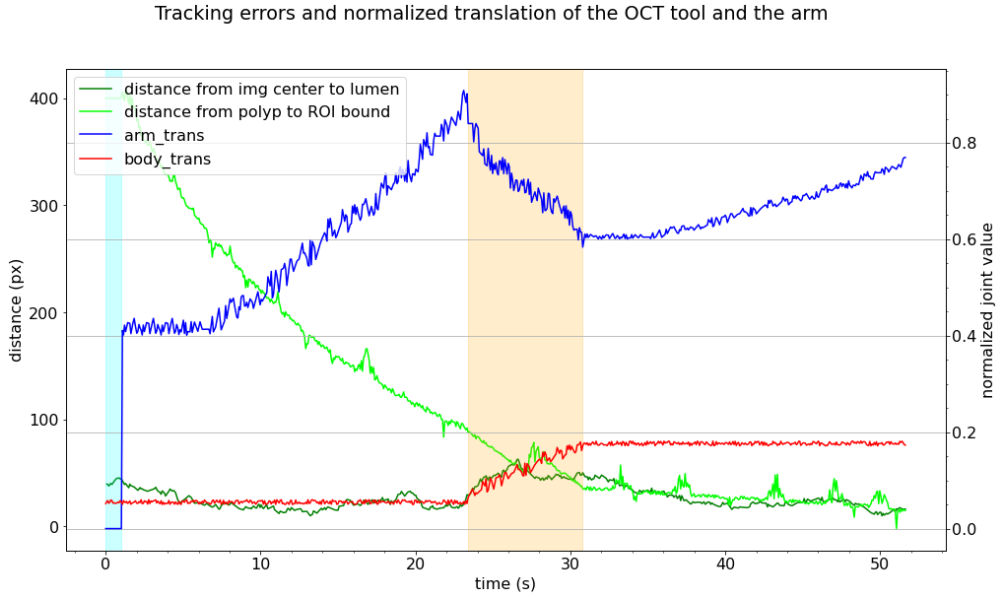


FIGURE C.7: Expérience 1 : Erreurs de suivi pendant l'expérience. En vert : distance de la lumière au centre de l'image. En vert : distance entre l'outil OCT et le centre de la région d'intérêt. La surbrillance orange compense la translation du bras (après avoir atteint la translation maximale). La surbrillance cyan indique que l'OCT est sorti du canal pour atteindre la position d'image souhaitée.

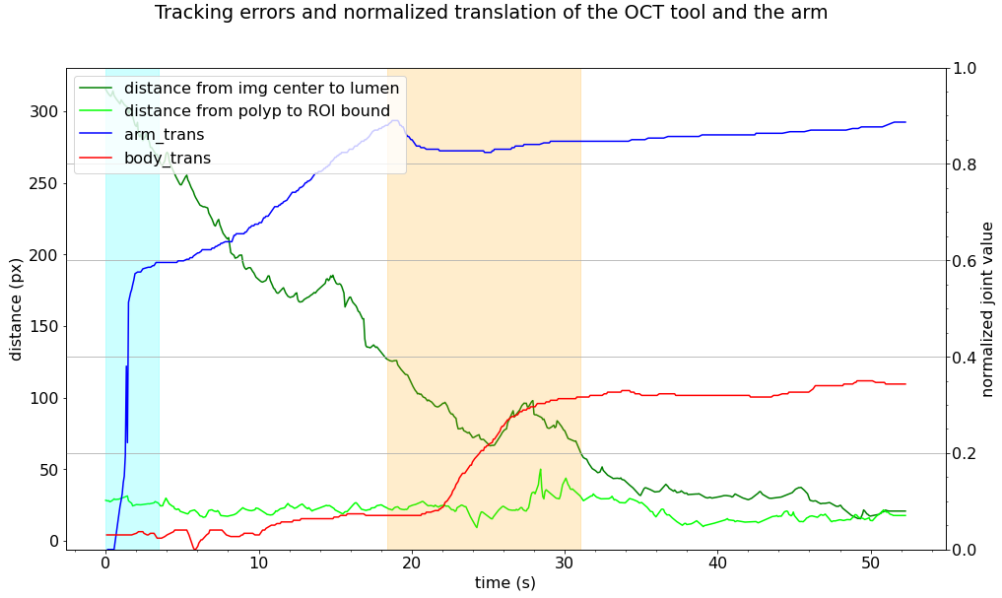


FIGURE C.8: Expérience 2 : Erreurs de suivi pendant l'expérience. En vert : distance de la lumière au centre de l'image. En vert : distance entre l'outil OCT et le centre de la région d'intérêt. La surbrillance orange compense la translation du bras (après avoir atteint la translation maximale). La surbrillance cyan indique que l'OCT est sorti du canal pour atteindre la position d'image souhaitée.

de la position de l'outil et la détection de la phase chirurgicale permettraient d'agir automatiquement. Il est ainsi possible de passer du mode assisté à la téléopération libre.

Bibliography

Abbott, D. J., Becke, C., Rothstein, R. I., and Peine, W. J. (2007). Design of an endoluminal notes robotic system. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 410–416. IEEE.

Abdelaal, A. E., Mathur, P., and Salcudean, S. E. (2020). Robotics in vivo: A perspective on human–robot interaction in surgical robotics. *Annual review of control, robotics, and autonomous systems*, 3(1):221–242.

Abe, Y., Da Silva, M., and Popović, J. (2007). Multiobjective control with frictional contacts. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 249–258.

Aertbeliën, E. and De Schutter, J. (2014). etasl/etc: A constraint-based task specification language and robot controller using expression graphs. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1540–1546. IEEE.

Agustinos, A., Wolf, R., Long, J.-A., Cinquin, P., and Voros, S. (2014). Visual servoing of a robotic endoscope holder based on surgical instrument tracking. In *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 13–18. IEEE.

Ahn, J., Kim, J., Lee, H., Hwang, M., and Kwon, D.-S. (2021). A highly intuitive and ergonomic redundant joint master device for four-degrees of freedom flexible endoscopic surgery robot. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 17(1):1–14.

Aimaq, R., Akopian, G., and Kaufman, H. S. (2011). Surgical site infection rates in laparoscopic versus open colorectal surgery. *The American surgeon*, 77(10):1290–1294.

Al-Bander, B., Mathew, A., Magerand, L., Trucco, E., and Manfredi, L. (2022). Real-time lumen detection for autonomous colonoscopy. In *MICCAI Workshop on Imaging Systems for GI Endoscopy*, pages 35–44. Springer.

Alexander III, A. D. (1972). Impacts of telemation on modern society. *On Theory and Practice of Robots and Manipulators*, pages 121–136.

Allen, P. K., Timcenko, A., Yoshimi, B., and Michelman, P. (1993). Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Transactions on Robotics and Automation*, 9(2):152–165.

Amini Khoiy, K., Mirbagheri, A., and Farahmand, F. (2016). Automatic tracking of laparoscopic instruments for autonomous control of a cameraman robot. *Minimally Invasive Therapy & Allied Technologies*, 25(3):121–128.

Atallah, S., Hodges, A., and Larach, S. (2018). Direct target notes: prospective applications for next generation robotic platforms. *Techniques in Coloproctology*, 22:363–371.

Athiniotis, S., Srivatsan, R., and Choset, H. (2019). Deep q reinforcement learning for autonomous navigation of surgical snake robot in confined spaces. In *Proceedings of the The Hamlyn Symposium on Medical Robotics, London, UK*, pages 23–26.

Attanasio, A., Scaglioni, B., De Momi, E., Fiorini, P., and Valdastri, P. (2021). Autonomy in surgical robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, 4:651–679.

Attanasio, A., Scaglioni, B., Leonetti, M., Frangi, A. F., Cross, W., Biyani, C. S., and Valdastri, P. (2020). Autonomous tissue retraction in robotic assisted minimally invasive surgery–a feasibility study. *IEEE Robotics and Automation Letters*, 5(4):6528–6535.

Avellino, I., Bailly, G., Arico, M., Morel, G., and Canlorbe, G. (2020). Multimodal and mixed control of robotic endoscopes. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Azizian, M., Khoshnam, M., Najmaei, N., and Patel, R. V. (2014). 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*, 10(3):263–274.

Azizian, M., Najmaei, N., Khoshnam, M., and Patel, R. (2015). Visual servoing in medical robotics: a survey. part ii: tomographic imaging modalities–techniques and applications. *The international journal of medical robotics and computer assisted surgery*, 11(1):67–79.

Babenko, B., Yang, M.-H., and Belongie, S. (2009). Visual tracking with online multiple instance learning. In *2009 IEEE Conference on computer vision and Pattern Recognition*, pages 983–990. IEEE.

Balasubramanian, S., Melendez-Calderon, A., and Burdet, E. (2011). A robust and sensitive metric for quantifying movement smoothness. *IEEE transactions on biomedical engineering*, 59(8):2126–2136.

Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A., and Burdet, E. (2015). On the analysis of movement smoothness. *Journal of neuroengineering and rehabilitation*, 12(1):1–11.

Bardou, B., Nageotte, F., Zanne, P., and de Mathelin, M. (2009). Design of a telemanipulated system for transluminal surgery. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5577–5582. IEEE.

Bardou, B., Zanne, P., Nageotte, F., and de Mathelin, M. (2010). Control of a multiple sections flexible endoscopic system. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2345–2350. IEEE.

Barrientos-Diez, J., Dong, X., Axinte, D., and Kell, J. (2021). Real-time kinematics of continuum robots: modelling and validation. *Robotics and Computer-Integrated Manufacturing*, 67:102019.

Bateux, Q., Marchand, E., Leitner, J., Chaumette, F., and Corke, P. (2017). Visual servoing from deep neural networks. *arXiv preprint arXiv:1705.08940*.

Bateux, Q., Marchand, E., Leitner, J., Chaumette, F., and Corke, P. (2018). Training deep neural networks for visual servoing. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 3307–3314. IEEE.

Batmaz, A. U., Falek, M. A., Zorn, L., Nageotte, F., Zanne, P., De Mathelin, M., and Dresp-Langley, B. (2017). Novice and expert haptic behaviours while using a

robot controlled surgery system. In *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*, pages 94–99. IEEE.

Battaglia, E., Boehm, J., Zheng, Y., Jamieson, A. R., Gahan, J., and Fey, A. M. (2021). Rethinking autonomous surgery: focusing on enhancement over autonomy. *European urology focus*, 7(4):696–705.

Bénard, F., Barkun, A. N., Martel, M., and von Renteln, D. (2018). Systematic review of colorectal cancer screening guidelines for average-risk adults: Summarizing the current global recommendations. *World journal of gastroenterology*, 24(1):124.

Berthet-Rayne, P., Leibrandt, K., Kim, K., Seneci, C. A., Shang, J., and Yang, G.-Z. (2018). Rolling-joint design optimization for tendon driven snake-like surgical robots. In *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pages 4964–4971. IEEE.

Bhurwal, A., Bartel, M. J., Heckman, M. G., Diehl, N. N., Raimondo, M., Wallace, M. B., and Woodward, T. A. (2016). Endoscopic mucosal resection: learning curve for large nonpolypoid colorectal neoplasia. *Gastrointestinal endoscopy*, 84(6):959–968.

Bihlmaier, A. (2016). Endoscope robots and automated camera guidance. In *Learning Dynamic Spatial Relations*, pages 23–102. Springer.

Bodenstedt, S., Padoy, N., and Hager, G. (2012). Learned partial automation for shared control in tele-robotic manipulation. In *2012 AAAI Fall Symposium Series*.

Bonjer, H., Hop, W., Nelson, H., Sargent, D. J., Lacy, A. M., Castells, A., Guillou, P. J., Thorpe, H., Brown, J., Delgado, S., et al. (2007). Transatlantic laparoscopically assisted vs open colectomy trials study group. laparoscopically assisted vs open colectomy for colon cancer: a meta-analysis. *Arch Surg*, 142(3):298–303.

Bouyarmane, K., Chappellet, K., Vaillant, J., and Kheddar, A. (2018). Quadratic programming for multirobot and task-space force control. *IEEE Transactions on Robotics*, 35(1):64–77.

Bouyarmane, K. and Kheddar, A. (2011). Using a multi-objective controller to synthesize simulated humanoid robot motion with changing contact configurations. In *2011 IEEE/RSJ international conference on intelligent robots and systems*, pages 4414–4419. IEEE.

Buhl, J. F., Grønhøj, R., Jørgensen, J. K., Mateus, G., Pinto, D., Sørensen, J. K., Bøgh, S., and Chrysostomou, D. (2019). A dual-arm collaborative robot system for the smart factories of the future. *Procedia manufacturing*, 38:333–340.

Cabras, P., Nageotte, F., Zanne, P., and Doignon, C. (2017). An adaptive and fully automatic method for estimating the 3d position of bendable instruments using endoscopic images. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(4):e1812.

Cai, P., Yue, X., Wang, M., and Cui, Y. (2022). Hierarchical motion planning at the acceleration level based on task priority matrix for space robot. *Nonlinear Dynamics*, 107(3):2309–2326.

Cao, C. G., MacKenzie, C. L., and Payandeh, S. (1996). Task and motion analyses in endoscopic surgery. In *ASME International Mechanical Engineering Congress and Exposition*, volume 15281, pages 583–590. American Society of Mechanical Engineers.

Cao, Y., Liao, C., Tan, A., Gao, Y., Mo, Z., and Gao, F. (2009). Meta-analysis of endoscopic submucosal dissection versus endoscopic mucosal resection for tumors of the gastrointestinal tract. *Endoscopy*, 41(09):751–757.

Caravaca-Mora, O., Zanne, P., Liao, G., Zulina, N., Heroin, L., Zorn, L., De Mathelin, M., Rosa, B., Nageotte, F. P., and Gora, M. J. (2023). Automatic intraluminal scanning with a steerable endoscopic optical coherence tomography catheter for gastroenterology applications. *Journal of Optical Microsystems*, 3(1):011005.

Caravaca-Mora, O., Zanne, P., Zorn, L., Nageotte, F., Zulina, N., Gravelyn, S., Montgomery, P., De Mathelin, M., Dallemagne, B., and Gora, M. J. (2020). Steerable oct catheter for real-time assistance during teleoperated endoscopic treatment of colorectal cancer. *Biomedical optics express*, 11(3):1231–1243.

Cataldi, E., Real, F., Suárez, A., Di Lillo, P., Pierri, F., Antonelli, G., Caccavale, F., Heredia, G., and Ollero, A. (2019). Set-based inverse kinematics control of an anthropomorphic dual arm aerial manipulator. In *2019 International Conference on robotics and automation (ICRA)*, pages 2960–2966. IEEE.

Chaumette, F., Hutchinson, S., and Corke, P. (2016). Visual servoing. *Springer handbook of robotics*, pages 841–866.

Chen, R., Liu, Q., Chen, Z., Guo, K., Yu, X., and Guo, L. (2023). Online trajectory generation for aerial manipulator subject to multi-tasks and inequality constraints. In *2023 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 933–939. IEEE.

Chikhaoui, M. T. and Rosa, B. (2022). Modeling and control strategies for flexible devices. In *Endorobotics*, pages 187–213. Elsevier.

Chitrakaran, V. K., Behal, A., Dawson, D. M., and Walker, I. D. (2007). Setpoint regulation of continuum robots using a fixed camera. *Robotica*, 25(5):581–586.

Chiu, H.-M., Wang, H.-P., Wu, M.-S., Lee, Y.-C., and Lin, J.-T. (2007). Endoscopic management of early colorectal cancer. In *Journal of Gastroenterology and Hepatology*, volume 22, pages A121–A121. BLACKWELL PUBLISHING 9600 GARSINGTON RD, OXFORD OX4 2DQ, OXON, ENGLAND.

Chiu, P. W. Y., Ho, K. Y., and Phee, S. J. (2021). Colonic endoscopic submucosal dissection using a novel robotic system (with video). *Gastrointestinal Endoscopy*, 93(5):1172–1177.

Chiu, P. W. Y. W., Phee, S. J., and Ho, K.-Y. (2019). Tu2016 colonic endoscopic submucosal dissecion using ease robotic system-a preclinical study. *Gastrointestinal Endoscopy*, 89(6):AB658–AB659.

Choi, M. S., Yun, S. H., Oh, C. K., Shin, J. K., Park, Y. A., Huh, J. W., Cho, Y. B., Kim, H. C., and Lee, W. Y. (2022). Learning curve for single-port robot-assisted rectal cancer surgery. *Annals of Surgical Treatment and Research*, 102(3):159–166.

Church, J. M. (2002). Warm water irrigation for dealing with spasm during colonoscopy: simple, inexpensive, and effective. *Gastrointestinal endoscopy*, 56(5):672–674.

Colan, J., Davila, A., and Hasegawa, Y. (2023). Constrained motion planning for a robotic endoscope holder based on hierarchical quadratic programming. In *2023 8th International Conference on Control and Robotics Engineering (ICCRE)*, pages 198–203. IEEE.

Corsi, D., Marzari, L., Pore, A., Farinelli, A., Casals, A., Fiorini, P., and Dall'Alba, D. (2023). Constrained reinforcement learning and formal verification for safe colonoscopy

navigation. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10289–10294. IEEE.

Cursi, F., Bai, W., Yeatman, E. M., and Kormushev, P. (2022). Model learning with backlash compensation for a tendon-driven surgical robot. *IEEE Robotics and Automation Letters*, 7(3):7958–7965.

Da Col, T., Mariani, A., Deguet, A., Menciassi, A., Kazanzides, P., and De Momi, E. (2020). Scan: System for camera autonomous navigation in robotic-assisted surgery. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2996–3002. IEEE.

Daneshgar Rahbar, M., Ying, H., and Pandya, A. (2021). Visual intelligence: Prediction of unintentional surgical-tool-induced bleeding during robotic and laparoscopic surgery. *Robotics*, 10(1):37.

De Donno, A., Nageotte, F., Zanne, P., Zorn, L., and de Mathelin, M. (2013a). Master/slave control of flexible instruments for minimally invasive surgery. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 483–489. IEEE.

De Donno, A., Zorn, L., Zanne, P., Nageotte, F., and de Mathelin, M. (2013b). Introducing stras: A new flexible robotic system for minimally invasive surgery. In *2013 IEEE International Conference on Robotics and Automation*, pages 1213–1220. IEEE.

De Lasa, M. and Hertzmann, A. (2009). Prioritized optimization for task-space control. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5755–5762. IEEE.

De Leon, M. P. and Di Gregorio, C. (2001). Pathology of colorectal cancer. *Digestive and Liver Disease*, 33(4):372–388.

De Mathelin, M., Le Bastard, F., Nageotte, F., Zanne, P., and Zorn, L. (2020). Master interface device for a motorised endoscopic system and installation comprising such a device. US Patent 10,660,719.

de Mathelin, M., Nageotte, F., Zanne, P., and Dresp-Langley, B. (2019). Sensors for expert grip force profiling: towards benchmarking manual control of a robotic device for surgical tool movements. *Sensors*, 19(20):4575.

de Moura, D. T. H., Aihara, H., Jirapinyo, P., Farias, G., Hathorn, K. E., Bazarbashi, A., Sachdev, A., and Thompson, C. C. (2019). Robot-assisted endoscopic submucosal dissection versus conventional esd for colorectal lesions: outcomes of a randomized pilot study in endoscopists without prior esd experience (with video). *Gastrointestinal endoscopy*, 90(2):290–298.

De Rossi, G., Minelli, M., Roin, S., Falezza, F., Sozzi, A., Ferraguti, F., Setti, F., Bonfè, M., Secchi, C., and Muradore, R. (2021). A first evaluation of a multi-modal learning system to control surgical assistant robots via action segmentation. *IEEE Transactions on Medical Robotics and Bionics*, 3(3):714–724.

De Schutter, J., De Laet, T., Rutgeerts, J., Decré, W., Smits, R., Aertbeliën, E., Claes, K., and Bruyninckx, H. (2007). Constraint-based task specification and estimation for sensor-based robot systems in the presence of geometric uncertainty. *The International Journal of Robotics Research*, 26(5):433–455.

Decré, W., Smits, R., Bruyninckx, H., and De Schutter, J. (2009). Extending itasc to support inequality constraints and non-instantaneous task specification. In *2009 IEEE International Conference on Robotics and Automation*, pages 964–971. IEEE.

Di Lillo, P., Pierri, F., Caccavale, F., and Antonelli, G. (2020). Experiments on whole-body control of a dual-arm mobile robot with the set-based task-priority inverse kinematics algorithm. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9096–9101. IEEE.

Diamond, S. and Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5.

Diana, M., Chung, H., Liu, K.-H., Dallemagne, B., Demartines, N., Mutter, D., and Marescaux, J. (2013). Endoluminal surgical triangulation: overcoming challenges of colonic endoscopic submucosal dissections using a novel flexible endoscopic surgical platform: feasibility study in a porcine model. *Surgical endoscopy*, 27:4130–4135.

Dietrich, A., Albu-Schäffer, A., and Hirzinger, G. (2012). On continuous null space projections for torque-based, hierarchical, multi-objective manipulation. In *2012 IEEE international conference on robotics and automation*, pages 2978–2985. IEEE.

Dietrich, A., Ott, C., and Albu-Schäffer, A. (2015). An overview of null space projections for redundant, torque-controlled robots. *The International Journal of Robotics Research*, 34(11):1385–1400.

Dik, V. K., Moons, L. M., and Siersema, P. D. (2014). Endoscopic innovations to increase the adenoma detection rate during colonoscopy. *World journal of gastroenterology: WJG*, 20(9):2200.

Djeha, M., Gergondet, P., and Kheddar, A. (2023). Robust task-space quadratic programming for kinematic-controlled robots. *IEEE Transactions on Robotics*.

Doignon, C., Nageotte, F., and De Mathelin, M. (2005). Segmentation and guidance of multiple rigid objects for intra-operative endoscopic vision. In *International Worksop on Dynamical Vision*, pages 314–327. Springer.

Drexler, W. and Fujimoto, J. G. (2015). *Optical Coherence Tomography: Technology and Applications*. SpringerReference.

Eickhoff, A., Jakobs, R., Kamal, A., Mermash, S., Riemann, J., and Van Dam, J. (2006). In vitro evaluation of forces exerted by a new computer-assisted colonoscope (the neoguide endoscopy system). *Endoscopy*, 38(12):1224–1229.

Ellis, R. D., Munaco, A. J., Reisner, L. A., Klein, M. D., Composto, A. M., Pandya, A. K., and King, B. W. (2016). Task analysis of laparoscopic camera control schemes. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 12(4):576–584.

Escande, A., Mansard, N., and Wieber, P.-B. (2010). Fast resolution of hierarchized inverse kinematics with inequality constraints. In *2010 IEEE International Conference on Robotics and Automation*, pages 3733–3738. IEEE.

Escande, A., Mansard, N., and Wieber, P.-B. (2014). Hierarchical quadratic programming: Fast online humanoid-robot motion generation. *The International Journal of Robotics Research*, 33(7):1006–1028.

Espiau, B., Merlet, J.-P., and Samson, C. (1990). Force-feedback control and non-contact sensing: a unified approach. In *Proceedings of the 8th CISM-IFTOMM Symposium on Theory and Practice of Robots Manipulators*.

Estrada, S., Duran, C., Schulz, D., Bismuth, J., Byrne, M. D., and O'Malley, M. K. (2016). Smoothness of surgical tool tip motion correlates to skill in endovascular tasks. *IEEE Transactions on Human-Machine Systems*, 46(5):647–659.

Ewing, D. R., Pigazzi, A., Wang, Y., and Ballantyne, G. H. (2004). Robots in the operating room—the history. In *Seminars in laparoscopic Surgery*, volume 11, pages 63–71. Sage Publications Sage CA: Thousand Oaks, CA.

Fanelli, R. D. (2018). Fundamentals of flexible endoscopy for general surgeons. *Fundamentals of General Surgery*, pages 163–173.

Faverjon, B. and Tournassoud, P. (1987). A local based approach for path planning of manipulators with a high number of degrees of freedom. In *Proceedings. 1987 IEEE international conference on robotics and automation*, volume 4, pages 1152–1159. IEEE.

Feldchtein, F. I., Gelikonov, G., Gelikonov, V., Iksanov, R., Kuranov, R., Sergeev, A. M., Gladkova, N., Ourutina, M., Warren, J., and Reitze, D. (1998). In vivo oct imaging of hard and soft tissue of the oral cavity. *Optics express*, 3(6):239–250.

Fennerty, M. B., Earnest, D. L., Hudson, P. B., and Sampliner, R. E. (1990). Physiologic changes during colonoscopy. *Gastrointestinal endoscopy*, 36(1):22–25.

Ferguson, C. J. (2016). An effect size primer: A guide for clinicians and researchers. *Methodological issues and strategies in clinical research.*

Fernández-Esparrach, G., Bernal, J., López-Cerón, M., Córdova, H., Sánchez-Montes, C., De Miguel, C. R., and Sánchez, F. J. (2016). Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy*, 48(09):837–842.

Ferreau, H. (2007). qpoases–an open-source implementation of the online active set strategy for fast model predictive control. In *Proceedings of the Workshop on Nonlinear Model Based Control–Software and Applications, Loughborough*, pages 29–30.

Fisher, B. L. (2004). Comparison of recovery time after open and laparoscopic gastric bypass and laparoscopic adjustable banding. *Obesity surgery*, 14(1):67–72.

Fisher, D. A., Maple, J. T., Ben-Menachem, T., Cash, B. D., Decker, G. A., Early, D. S., Evans, J. A., Fanelli, R. D., Fukami, N., Hwang, J. H., et al. (2011). Complications of colonoscopy. *Gastrointestinal endoscopy*, 74(4):745–752.

Flacco, F., De Luca, A., and Khatib, O. (2012). Prioritized multi-task motion control of redundant robots under hard joint constraints. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3970–3977. IEEE.

Flores-Abad, A., Ma, O., Pham, K., and Ulrich, S. (2014). A review of space robotics technologies for on-orbit servicing. *Progress in aerospace sciences*, 68:1–26.

Frankel, A., Gardner, R., Maynard, L., and Kelly, A. (2007). Using the communication and teamwork skills (cats) assessment to measure health care team performance. *The Joint Commission Journal on Quality and Patient Safety*, 33(9):549–558.

Freddi, A., Longhi, S., Monteriù, A., and Ortenzi, D. (2016). Redundancy analysis of cooperative dual-arm manipulators. *International Journal of Advanced Robotic Systems*, 13(5):1729881416657754.

Fritscher-Ravens, A., Ghanbari, A., Holland, C., Olagbeye, F., Hardeler, K., Seehusen, F., Jacobsen, B., and Mannur, K. (2009). Beyond notes: randomized controlled study of different methods of flexible endoscopic hemostasis of artificially induced hemorrhage, via notes access to the peritoneal cavity. *Endoscopy*, 41(01):29–35.

Fritz, C. O., Morris, P. E., and Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General*, 141(1):2.

Fuchs, K. (2002). Minimally invasive surgery. *Endoscopy*, 34(02):154–159.

Fuchs, K.-H. and Breithaupt, W. (2012). Transgastric small bowel resection with the new multitasking platform endosamurai™ for natural orifice transluminal endoscopic surgery. *Surgical endoscopy*, 26:2281–2287.

Fujii, K., Gras, G., Salerno, A., and Yang, G.-Z. (2018). Gaze gesture based human robot interaction for laparoscopic surgery. *Medical image analysis*, 44:196–214.

Gaab, M. R. (2013). Instrumentation: endoscopes and equipment. *World neurosurgery*, 79(2):S14–e11.

Garcìa-Garcìa, H. M., Gogas, B. D., Serruys, P. W., and Bruining, N. (2011). Ivus-based imaging modalities for tissue characterization: similarities and differences. *The international journal of cardiovascular imaging*, 27:215–224.

Gatta, C., Pujol, O., Leor, O. R., Ferre, J. M., and Radeva, P. (2009). Fast rigid registration of vascular structures in ivus sequences. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):1006–1011.

Geisert, M., Del Prete, A., Mansard, N., Romano, F., and Nori, F. (2017). Regularized hierarchical differential dynamic programming. *IEEE Transactions on Robotics*, 33(4):819–833.

Goetz, M., Watson, A., and Kiesslich, R. (2011). Confocal laser endomicroscopy in gastrointestinal diseases. *Journal of biophotonics*, 4(7-8):498–508.

Gonzalez Herrera, J. F., Pore, A., Sestini, L., Sahu, S. K., Liao, G., Zanne, P., Dall'Alba, D., Hernansanz, A., Rosa, B., Nageotte, F., et al. (2022). Autonomous image guided control of endoscopic orientation for oct scanning. In *Proceeding of the 11th Joint Workshop on New Technologies for Computer/Robot Assisted Surgery*, pages 68–69. CRAS-eu.

Gora, M. J., Sauk, J. S., Carruth, R. W., Gallagher, K. A., Suter, M. J., Nishioka, N. S., Kava, L. E., Rosenberg, M., Bouma, B. E., and Tearney, G. J. (2013). Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure. *Nature medicine*, 19(2):238–240.

Gora, M. J., Suter, M. J., Tearney, G. J., and Li, X. (2017). Endoscopic optical coherence tomography: technologies and clinical applications. *Biomedical optics express*, 8(5):2405–2444.

Grassmann, R. M., Shentu, C., Hamoda, T., Dewi, P. T., and Burgner-Kahrs, J. (2024). Open continuum robotics–one actuation module to create them all. *Frontiers in Robotics and AI*, 11:1272403.

Gruijthuijsen, C., Garcia-Peraza-Herrera, L. C., Borghesan, G., Reynaerts, D., Deprest, J., Ourselin, S., Vercauteren, T., and Vander Poorten, E. (2022). Robotic endoscope control via autonomous instrument tracking. *Frontiers in Robotics and AI*, 9:832208.

Ha, X. T., Wu, D., Ourak, M., Borghesan, G., Dankelman, J., Menciassi, A., and Vander Poorten, E. (2022). Shape sensing of flexible robots based on deep learning. *IEEE Transactions on Robotics*, 39(2):1580–1593.

Haidegger, T. (2019). Autonomy for surgical robots: Concepts and paradigms. *IEEE Transactions on Medical Robotics and Bionics*, 1(2):65–76.

Harji, D., Houston, F., Burke, J., Griffiths, B., Tilney, H., Miskovic, D., Evans, C., Khan, J., Soomro, N., and Bach, S. P. (2023). The current status of robotic colorectal surgery training programmes. *Journal of Robotic Surgery*, 17(2):251–263.

Harlow, C., Sivananthan, A., Ayaru, L., Patel, K., Darzi, A., and Patel, N. (2020). Endoscopic submucosal dissection: an update on tools and accessories. *Therapeutic Advances in Gastrointestinal Endoscopy*, 13:2631774520957220.

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA: Los Angeles, CA.

Haycock, A., Koch, A. D., Familiari, P., van Delft, F., Dekker, E., Petruzziello, L., Haringsma, J., and Thomas-Gibson, S. (2010). Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointestinal endoscopy*, 71(2):298–307.

Hendrick, R. J., Herrell, S. D., and Webster, R. J. (2014). A multi-arm hand-held robotic system for transurethral laser prostate surgery. In *2014 IEEE international conference on robotics and automation (ICRA)*, pages 2850–2855. Ieee.

Hoffman, E. M., Laurenzi, A., Muratore, L., Tsagarakis, N. G., and Caldwell, D. G. (2018). Multi-priority cartesian impedance control based on quadratic programming optimization. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 309–315. IEEE.

Hong, M., Zhou, H., Liu, L., and Guo, Y. (2022). Coordinated control of free-floating dual-arm space robots based on hybrid task-priority approach. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 236(11):2206–2217.

Hoteya, S., Iizuka, T., Kikuchi, D., and Yahagi, N. (2009). Benefits of endoscopic submucosal dissection according to size and location of gastric neoplasm, compared with conventional mucosal resection. *Journal of Gastroenterology and Hepatology*, 24(6):1102–1106.

Hsu, P., Mauser, J., and Sastry, S. (1989). Dynamic control of redundant manipulators. *Journal of Robotic Systems*, 6(2):133–148.

Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., Hee, M. R., Flotte, T., Gregory, K., Puliafito, C. A., et al. (1991). Optical coherence tomography. *science*, 254(5035):1178–1181.

Huang, H.-C., Chen, C.-P., and Wang, P.-R. (2012). Particle swarm optimization for solving the inverse kinematics of 7-dof robotic manipulators. In *2012 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 3105–3110. IEEE.

Huang, Y., Lai, W., Cao, L., Burdet, E., and Phee, S. J. (2021). Design and evaluation of a foot-controlled robotic system for endoscopic surgery. *IEEE Robotics and Automation Letters*, 6(2):2469–2476.

Hutchinson, S., Hager, G. D., and Corke, P. I. (1996). A tutorial on visual servo control. *IEEE transactions on robotics and automation*, 12(5):651–670.

Hwang, M. and Kwon, D.-S. (2020a). K-flex: a flexible robotic platform for scar-free endoscopic surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 16(2):e2078.

Hwang, M. and Kwon, D.-S. (2020b). K-flex: a flexible robotic platform for scar-free endoscopic surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 16(2):e2078.

Iddan, G., Meron, G., Glukhovsky, A., and Swain, P. (2000). Wireless capsule endoscopy. *Nature*, 405(6785):417–417.

Imanberdiyev, N. and Kayacan, E. (2020). Redundancy resolution based trajectory generation for dual-arm aerial manipulators via online model predictive control. In *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, pages 674–681. IEEE.

Imperiale, T. F., Ransohoff, D. F., Itzkowitz, S. H., Levin, T. R., Lavin, P., Lidgard, G. P., Ahlquist, D. A., and Berger, B. M. (2014). Multitarget stool dna testing for colorectal-cancer screening. *New England Journal of Medicine*, 370(14):1287–1297.

Jarc, A. M. and Curet, M. J. (2017). Viewpoint matters: objective performance metrics for surgeon endoscope control during robot-assisted surgery. *Surgical endoscopy*, 31:1192–1202.

Jha, D., Ali, S., Tomar, N. K., Johansen, H. D., Johansen, D., Rittscher, J., Riegler, M. A., and Halvorsen, P. (2021). Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. *Ieee Access*, 9:40496–40510.

Jiménez-Rodríguez, R. M., Rubio-Dorado-Manzanares, M., Díaz-Pavón, J. M., Reyes-Díaz, M. L., Vazquez-Monchul, J. M., Garcia-Cabrera, A. M., Padillo, J., and De la Portilla, F. (2016). Learning curve in robotic rectal cancer surgery: current state of affairs. *International journal of colorectal disease*, 31:1807–1815.

Johns Hopkins Medicine, J. (2024). Colonoscopy. https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/colonoscopy. Accessed: 2024-04-10.

Kang, J. and Gwak, J. (2019). Ensemble of instance segmentation models for polyp segmentation in colonoscopy images. *IEEE Access*, 7:26440–26447.

Kelley, K. and Preacher, K. J. (2012). On effect size. *Psychological methods*, 17(2):137.

Kenngott, H. G., Fischer, L., Nickel, F., Rom, J., Rassweiler, J., and Müller-Stich, B. (2012). Status of robotic assistance—a less traumatic and more accurate minimally invasive surgery? *Langenbeck's archives of surgery*, 397:333–341.

Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *The international journal of robotics research*, 5(1):90–98.

Khatib, O. (1987). A unified approach for motion and force control of robot manipulators: The operational space formulation. *IEEE Journal on Robotics and Automation*, 3(1):43–53.

Kim, J., de Mathelin, M., Ikuta, K., and Kwon, D.-S. (2022). Advancement of flexible robot technologies for endoluminal surgeries. *Proceedings of the IEEE*, 110(7):909–931.

Kim, J., Hwang, M., Lee, D., Kim, H., Ahn, J., You, J., Baek, D., and Kwon, D.-S. (2019). Effects of flexible surgery robot on endoscopic procedure: Preliminary benchtop user test. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–6. IEEE.

Kim, Y. J., Kim, E. S., Cho, K. B., Park, K. S., Jang, B. K., Chung, W. J., and Hwang, J. S. (2013). Comparison of clinical outcomes among different endoscopic resection methods for treating colorectal neoplasia. *Digestive diseases and sciences*, 58:1727–1736.

Kominami, Y., Yoshida, S., Tanaka, S., Sanomura, Y., Hirakawa, T., Raytchev, B., Tamaki, T., Koide, T., Kaneda, K., and Chayama, K. (2016). Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointestinal endoscopy*, 83(3):643–649.

Kossowsky, H. and Nisky, I. (2022). Predicting the timing of camera movements from the kinematics of instruments in robotic-assisted surgery using artificial neural networks. *IEEE Transactions on Medical Robotics and Bionics*, 4(2):391–402.

Kragic, D. and Christensen, H. I. (2001). Cue integration for visual servoing. *IEEE transactions on robotics and automation*, 17(1):18–27.

Krüger, J., Schreck, G., and Surdilovic, D. (2011). Dual arm robot for flexible and cooperative assembly. *CIRP annals*, 60(1):5–8.

Krupa, A., Gangloff, J., Doignon, C., De Mathelin, M. F., Morel, G., Leroy, J., Soler, L., and Marescaux, J. (2003). Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE transactions on robotics and automation*, 19(5):842–853.

Kume, K. (2009). Endoscopic mucosal resection and endoscopic submucosal dissection for early gastric cancer: current and original devices. *World Journal of Gastrointestinal Endoscopy*, 1(1):21.

Kwoh, Y. S., Hou, J., Jonckheere, E. A., and Hayati, S. (1988). A robot with improved absolute positioning accuracy for ct guided stereotactic brain surgery. *IEEE transactions on biomedical engineering*, 35(2):153–160.

Latif, H. O., Sherkat, N., and Lotfi, A. (2008). Telegaze: Teleoperation through eye gaze. In *2008 7th IEEE International Conference on Cybernetic Intelligent Systems*, pages 1–6. IEEE.

Lazo, J. F., Lai, C.-F., Moccia, S., Rosa, B., Catellani, M., de Mathelin, M., Ferrigno, G., Breedveld, P., Dankelman, J., and De Momi, E. (2022). Autonomous intraluminal navigation of a soft robot using deep-learning-based visual servoing. *arXiv preprint arXiv:2207.00401*.

Lee, C., Wang, Y.-F., Uecker, D. R., and Wang, Y. (1994). Image analysis for automated tracking in robot-assisted endoscopic surgery. In *Proceedings of 12th International Conference on Pattern Recognition*, volume 2, pages 88–89. IEEE Computer Society.

Lee, D.-H., Hwang, M., and Kwon, D.-S. (2019). Robotic endoscopy system (easyendo) with a robotic arm mountable on a conventional endoscope. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 367–372. IEEE.

Lee, W. S., Cho, J. W., Kim, Y. D., Kim, K. J., and Jang, B. I. (2011). Technical issues and new devices of esd of early gastric cancer. *World journal of gastroenterology: WJG*, 17(31):3585.

Leijte, E., de Blaauw, I., Van Workum, F., Rosman, C., and Botden, S. (2020). Robot assisted versus laparoscopic suturing learning curve in a simulated setting. *Surgical endoscopy*, 34:3679–3689.

Levy, L. C., Adrales, G., and Rothstein, R. I. (2008). Training for notes. *Gastrointestinal Endoscopy Clinics of North America*, 18(2):343–360.

Li, J. C.-H. (2016). Effect size measures in a two-independent-samples case with non-normal and nonhomogeneous data. *Behavior research methods*, 48:1560–1574.

Li, P., Chaumette, F., and Tahri, O. (2005). A shape tracking algorithm for visual servoing. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 2847–2852. IEEE.

Li, Z. and Chiu, P. W.-Y. (2018). Robotic endoscopy. *Visceral medicine*, 34(1):45–51.

Liberman, A., Shrier, I., and Gordon, P. (2005). Injuries sustained by colorectal surgeons performing colonoscopy. *Surgical Endoscopy and Other Interventional Techniques*, 19:1606–1609.

Lieberman, D. A., Rex, D. K., Winawer, S. J., Giardiello, F. M., Johnson, D. A., and Levin, T. R. (2012). Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the us multi-society task force on colorectal cancer. *Gastroenterology*, 143(3):844–857.

Liegeois, A. et al. (1977). Automatic supervisory control of the configuration and behavior of multibody mechanisms. *IEEE transactions on systems, man, and cybernetics*, 7(12):868–871.

Liu, M., Micaelli, A., Evrard, P., Escande, A., and Andriot, C. (2011). Interactive dynamics and balance of a virtual character during manipulation tasks. In *2011 IEEE International Conference on Robotics and Automation*, pages 1676–1682. IEEE.

Lohsiriwat, V. (2010). Colonoscopic perforation: incidence, risk factors, management and outcome. *World journal of gastroenterology: WJG*, 16(4):425.

Légner, A., Diana, M., Halvax, P., Liu, Y.-Y., Zorn, L., Zanne, P., Nageotte, F., De Mathelin, M., Dallemagne, B., and Marescaux, J. (2017). Endoluminal surgical triangulation 2.0: A new flexible surgical robot. preliminary pre-clinical results with colonic submucosal dissection. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(3):e1819. e1819 RCS-16-0150.R2.

Mahapatra, S. K., Ghosal, A., et al. (2020). 3d printed cable-driven continuum robots with generally routed cables: modeling and experiments. *arXiv preprint arXiv:2003.04593*.

Mandapathil, M., Duvvuri, U., Güldner, C., Teymoortash, A., Lawson, G., and Werner, J. A. (2015). Transoral surgery for oropharyngeal tumors using the medrobotics® flex® system–a case report. *International journal of surgery case reports*, 10:173–175.

Mansard, N., Khatib, O., and Kheddar, A. (2009). A unified approach to integrate unilateral constraints in the stack of tasks. *IEEE Transactions on Robotics*, 25(3):670–685.

Marchand, E., Chaumette, F., Spindler, F., and Perrier, M. (2002). Controlling an uninstrumented manipulator by visual servoing. *The International Journal of Robotics Research*, 21(7):635–647.

Marescaux, J., Leroy, J., Gagner, M., Rubino, F., Mutter, D., Vix, M., Butner, S. E., and Smith, M. K. (2001). Transatlantic robot-assisted telesurgery. *Nature*, 413(6854):379–380.

Marshall, J. B. (1995). Technical proficiency of trainees performing colonoscopy: a learning curve. *Gastrointestinal endoscopy*, 42(4):287–291.

Martin, J. W., Scaglioni, B., Norton, J. C., Subramanian, V., Arezzo, A., Obstein, K. L., and Valdastri, P. (2020). Enabling the future of colonoscopy with intelligent and autonomous magnetic manipulation. *Nature machine intelligence*, 2(10):595–606.

Mascagni, P., Lim, S. G., Fiorillo, C., Zanne, P., Nageotte, F., Zorn, L., Perretta, S., de Mathelin, M., Marescaux, J., and Dallemagne, B. (2019). Democratizing endoscopic submucosal dissection: single-operator fully robotic colorectal endoscopic submucosal dissection in a pig model. *Gastroenterology*, 156(6):1569–1571.

McCarty, T. R., Bazarbashi, A. N., Thompson, C. C., and Aihara, H. (2020). Hybrid endoscopic submucosal dissection (esd) compared with conventional esd for colorectal lesions: a systematic review and meta-analysis. *Endoscopy*, 53(10):1048–1058.

McCashland, T., Brand, R., Lyden, E., De Garmo, P., et al. (2000). The time and financial impact of training fellows in endoscopy. *Official journal of the American College of Gastroenterology— ACG*, 95(11):3129–3132.

Mehta, P. P., Sanaka, M. R., Parsi, M. A., Albeldawi, M. J., Dumot, J. A., Lopez, R., Zuccaro, G., and Vargo, J. J. (2014). Association of procedure length on outcomes and adverse events of endoscopic retrograde cholangiopancreatography. *Gastroenterology Report*, 2(2):140–144.

Meng, Z., Huang, Z., Deng, B., Ling, L., Ning, Y., and Rafiq, S. M. (2022). Robotic-assisted vs non-robotic traction techniques in endoscopic submucosal dissection for malignant gastrointestinal lesions. *Frontiers in Oncology*, 12:1062357.

Menon, S. and Trudgill, N. (2014). How commonly is upper gastrointestinal cancer missed at endoscopy? a meta-analysis. *Endoscopy international open*, 2(02):E46–E50.

Misra, G. and Bai, X. (2017). Task-constrained trajectory planning of free-floating space-robotic systems using convex optimization. *Journal of Guidance, Control, and Dynamics*, 40(11):2857–2870.

Mistry, M., Nakanishi, J., Cheng, G., and Schaal, S. (2008). Inverse kinematics with floating base and constraints for full body humanoid robot control. In *Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots*, pages 22–27. IEEE.

Mistry, M., Nakanishi, J., and Schaal, S. (2007). Task space control with prioritization for balance and locomotion. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 331–338. IEEE.

Mitsuishi, M., Morita, A., Sugita, N., Sora, S., Mochizuki, R., Tanimoto, K., Baek, Y. M., Takahashi, H., and Harada, K. (2013). Master–slave robotic platform and its feasibility study for micro-neurosurgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 9(2):180–189.

Moghul, M. R., Sodergren, M. H., Clark, J., Teare, J., Yang, G.-Z., and Darzi, A. (2013). Education and training in notes: a systematic review. *Surgical innovation*, 20(3):282–291.

Moore, L. J., Wilson, M. R., McGrath, J. S., Waine, E., Masters, R. S., and Vine, S. J. (2015). Surgeons' display reduced mental effort and workload while performing robotically assisted surgical tasks, when compared to conventional laparoscopy. *Surgical endoscopy*, 29:2553–2560.

Morgan, E., Arnold, M., Gini, A., Lorenzoni, V., Cabasag, C., Laversanne, M., Vignat, J., Ferlay, J., Murphy, N., and Bray, F. (2023). Global burden of colorectal cancer in 2020 and 2040: Incidence and mortality estimates from globocan. *Gut*, 72(2):338–344.

Mudunuri, A. V. (2010). *Autonomous camera control system for surgical robots*. Wayne State University.

Mulhall, B. P., Veerappan, G. R., and Jackson, J. L. (2005). Meta-analysis: computed tomographic colonography. *Annals of Internal Medicine*, 142(8):635–650.

Nadeau, C. and Krupa, A. (2013). Intensity-based ultrasound visual servoing: Modeling and validation with 2-d and 3-d probes. *IEEE Transactions on Robotics*, 29(4):1003–1015.

Nadeau, C., Ren, H., Krupa, A., and Dupont, P. (2014). Intensity-based visual servoing for instrument and tissue tracking in 3d ultrasound volumes. *IEEE Transactions on Automation Science and Engineering*, 12(1):367–371.

Nageotte, F., Zorn, L., Zanne, P., and De Mathelin, M. (2020). Stras: A modular and flexible telemanipulated robotic device for intraluminal surgery. In *Handbook of Robotic and Image-Guided Surgery*, pages 123–146. Elsevier.

Nasseri, Y., Stettler, I., Shen, W., Zhu, R., Alizadeh, A., Lee, A., Cohen, J., and Barnajian, M. (2021). Learning curve in robotic colorectal surgery. *Journal of Robotic Surgery*, 15:489–495.

Nazari, A. A., Zareinia, K., and Janabi-Sharifi, F. (2022). Visual servoing of continuum robots: Methods, challenges, and prospects. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 18(3):e2384.

NCI, N. C. I. (2022). Colorectal-cancer-screening. https://www.cancer.gov/types/colorectal/patient/colorectal-screening-pdq [Accessed: 2024-02-28].

Oka, S., Tanaka, S., Kaneko, I., Mouri, R., Hirata, M., Kanao, H., Kawamura, T., Yoshida, S., Yoshihara, M., and Chayama, K. (2006). Endoscopic submucosal dissection for residual/local recurrence of early gastric cancer after endoscopic mucosal resection. *Endoscopy*, 38(10):996–1000.

Ott, L., Nageotte, F., Zanne, P., and De Mathelin, M. (2011). Robotic assistance to flexible endoscopy by physiological-motion tracking. *IEEE Transactions on Robotics*, 27(2):346–359.

Ott, L., Zanne, P., Nageotte, F., de Mathelin, M., and Gangloff, J. (2008). Physiological motion rejection in flexible endoscopy using visual servoing. In *2008 IEEE International Conference on Robotics and Automation*, pages 2928–2933. IEEE.

Pandya, A., Reisner, L. A., King, B., Lucas, N., Composto, A., Klein, M., and Ellis, R. D. (2014). A review of camera viewpoint automation in robotic and laparoscopic surgery. *Robotics*, 3(3):310–329.

Parker, J. K., Khoogar, A. R., and Goldberg, D. E. (1989). Inverse kinematics of redundant robots using genetic algorithms. In *1989 IEEE International Conference on Robotics and Automation*, pages 271–272. IEEE Computer Society.

Pasini, N., Mariani, A., Deguet, A., Kazanzides, P., and De Momi, E. (2023). Grace: Online gesture recognition for autonomous camera-motion enhancement in robot-assisted surgery. *IEEE Robotics and Automation Letters*.

Patel, N., Seneci, C. A., Shang, J., Leibrandt, K., Yang, G.-Z., Darzi, A., and Teare, J. (2015). Evaluation of a novel flexible snake robot for endoluminal surgery. *Surgical endoscopy*, 29:3349–3355.

Patel, P. and Thakkar, S. J. (2011). Simulation and training in transluminal endoscopy. *Techniques in Gastrointestinal Endoscopy*, 13(2):140–145.

Phee, S., Low, S., Sun, Z., Ho, K., Huang, W., and Thant, Z. (2008). Robotic system for no-scar gastrointestinal surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 4(1):15–22.

Philippsen, R., Sentis, L., and Khatib, O. (2011). An open source extensible software package to create whole-body compliant skills in personal mobile manipulators. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1036–1041. IEEE.

Poignonec, T., Nageotte, F., and Bayle, B. (2023). A fading memory discontinuous ekf for the online model identification of cable-driven robots with backlash. In *2023 American Control Conference (ACC)*, pages 949–954. IEEE.

Poignonec, T., Zanne, P., Rosa, B., and Nageotte, F. (2020). Towards in situ backlash estimation of continuum robots using an endoscopic camera. *IEEE Robotics and Automation Letters*, 5(3):4788–4795.

Pore, A., Finocchiaro, M., Dall'Alba, D., Hernansanz, A., Ciuti, G., Arezzo, A., Menciassi, A., Casals, A., and Fiorini, P. (2022). Colonoscopy navigation using end-to-end deep visuomotor control: A user study. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9582–9588. IEEE.

Pore, A., Tagliabue, E., Piccinelli, M., Dall'Alba, D., Casals, A., and Fiorini, P. (2021). Learning from demonstrations for autonomous soft-tissue retraction. In *2021 International Symposium on Medical Robotics (ISMR)*, pages 1–7. IEEE.

Porto, R. A., Nageotte, F., Zanne, P., and de Mathelin, M. (2019). Position control of medical cable-driven flexible instruments by combining machine learning and kinematic analysis. In *2019 international conference on robotics and automation (ICRA)*, pages 7913–7919. IEEE.

Portolés, S., Vanbiervliet, P., Rosa, B., Tomassetti, C., Meuleman, C., Vander Poorten, E. B., and Reynaerts, D. (2015). Force control for tissue tensioning in precise robotic laser surgery. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 579–585. IEEE.

Price, K., Peine, J., Mencattelli, M., Chitalia, Y., Pu, D., Looi, T., Stone, S., Drake, J., and Dupont, P. E. (2023). Using robotics to move a neurosurgeon's hands to the tip of their endoscope. *Science Robotics*, 8(82):eadg6042.

Pugin, F., Bucher, P., and Morel, P. (2011). History of robotic surgery: from aesop and zeus to da vinci. *Journal of visceral surgery*, 148(5):e3–e8.

Qu, J., Zhang, F., Wang, Y., and Fu, Y. (2019). Human-like coordination motion learning for a redundant dual-arm robot. *Robotics and Computer-Integrated Manufacturing*, 57:379–390.

Que, D., Xinyu, L., and Shaohua, M. (2022). Design of a space dual-arm robot and its ground microgravity simulation platform. In *2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, pages 405–409. IEEE.

Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., Ng, A. Y., et al. (2009). Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, Japan.

Rabeneck, L., Paszat, L. F., Hilsden, R. J., Saskin, R., Leddin, D., Grunfeld, E., Wai, E., Goldwasser, M., Sutradhar, R., and Stukel, T. A. (2008). Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. *Gastroenterology*, 135(6):1899–1906.

Rader, S., Kaul, L., Fischbach, H., Vahrenkamp, N., and Asfour, T. (2016). Design of a high-performance humanoid dual arm system with inner shoulder joints. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 523–529. IEEE.

Ravindran, S., Haycock, A., Woolf, K., and Thomas-Gibson, S. (2021). Development and impact of an endoscopic non-technical skills (ents) behavioural marker system. *BMJ Simulation & Technology Enhanced Learning*, 7(1):17.

Regnard, J.-F., Grunenwald, D., Spaggiari, L., Girard, P., Elias, D., Ducreux, M., Baldeyrou, P., and Levasseur, P. (1998). Surgical treatment of hepatic and pulmonary metastases from colorectal cancers. *The Annals of thoracic surgery*, 66(1):214–218.

Reilink, R. (2013). *Image-based robotic steering of advanced flexible endoscopes and instruments*. PhD thesis, Thesis, Zutphen.

Reilink, R., Stramigioli, S., and Misra, S. (2010). Image-based flexible endoscope steering. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2339–2344. IEEE.

Ribeiro, E., Uhl, A., Wimmer, G., and Häfner, M. (2016). Exploring deep learning and transfer learning for colonic polyp classification. *Computational and mathematical methods in medicine*, 2016.

Rivas-Blanco, I., Estebanez, B., Cuevas-Rodriguez, M., Bauzano, E., and Muñoz, V. F. (2014). Towards a cognitive camera robotic assistant. In *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 739–744. IEEE.

Rivera-Serrano, C. M., Johnson, P., Zubiate, B., Kuenzler, R., Choset, H., Zenati, M., Tully, S., and Duvvuri, U. (2012). A transoral highly flexible robot: Novel technology and application. *The Laryngoscope*, 122(5):1067–1071.

Rosa, B., Bordoux, V., and Nageotte, F. (2019). Combining differential kinematics and optical flow for automatic labeling of continuum robots in minimally invasive surgery. *Frontiers in Robotics and AI*, 6:86.

Rosenberg, L. B. (1993). Virtual fixtures: Perceptual tools for telerobotic manipulation. In *Proceedings of IEEE virtual reality annual international symposium*, pages 76–82. Ieee.

Ruiter, J., Bonnema, G. M., van der Voort, M. C., and Broeders, I. A. M. J. (2013). Robotic control of a traditional flexible endoscope for therapy. *Journal of robotic surgery*, 7:227–234.

Ruscio, J. (2008). A probability-based measure of effect size: robustness to base rates and other factors. *Psychological methods*, 13(1):19.

Saito, I., Tsuji, Y., Sakaguchi, Y., Niimi, K., Ono, S., Kodashima, S., Yamamichi, N., Fujishiro, M., and Koike, K. (2014). Complications related to gastric endoscopic submucosal dissection and their managements. *Clinical endoscopy*, 47(5):398–403.

Sakr, M., Li, Z. J., Van der Loos, H. M., Kulić, D., and Croft, E. A. (2022). Quantifying demonstration quality for robot learning and generalization. *IEEE Robotics and Automation Letters*, 7(4):9659–9666.

Salini, J., Padois, V., and Bidaud, P. (2011). Synthesis of complex humanoid whole-body behavior: A focus on sequencing and tasks transitions. In *2011 IEEE international conference on robotics and automation*, pages 1283–1290. IEEE.

Saltzman, J. R., Cash, B. D., Pasha, S. F., Early, D. S., Muthusamy, V. R., Khashab, M. A., Chathadi, K. V., Fanelli, R. D., Chandrasekhara, V., Lightdale, J. R., et al. (2015). Bowel preparation before colonoscopy. *Gastrointestinal endoscopy*, 81(4):781–794.

Samson, C. (1991). Velocity and torque feedback control of a nonholonomic cart. In *Advanced Robot Control: Proceedings of the International Workshop on Nonlinear and Adaptive Control: Issues in Robotics, Grenoble, France, Nov. 21–23, 1990*, pages 125–151. Springer.

Samson, C. and Espiau, B. (1990). Application of the task-function approach to sensor-based control of robot manipulators. *IFAC Proceedings Volumes*, 23(8):269–274.

Samson, C., Espiau, B., and Borgne, M. L. (1991). *Robot control: the task function approach*. Oxford University Press, Inc.

Sathyanarayana, S. (2006). Nonuniform rotational distortion (nurd) reduction. US Patent 7,024,025.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of modern applied statistical methods*, 8:597–599.

Saxena, A., Pandya, H., Kumar, G., Gaud, A., and Krishna, K. M. (2017). Exploring convolutional networks for end-to-end visual servoing. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3817–3823. IEEE.

Schäfer, M., Krähenbühl, L., and Büchler, M. (1998). Comparison of adhesion formation in open and laparoscopic surgery. *Digestive surgery*, 15(2):148–152.

Schroeder, W., Martin, K., and Lorensen, B. (2006). The visualization toolkit, 4th edn. kitware. *New York*.

Schuler, P. J., Duvvuri, U., Friedrich, D. T., Rotter, N., Scheithauer, M. O., and Hoffmann, T. K. (2015). First use of a computer-assisted operator-controlled flexible endoscope for transoral surgery. *The Laryngoscope*, 125(3):645–648.

Schwaitzberg, S. D., Roberts, K., Romanelli, J. R., Desilets, D. J., Earle, D., Horgan, S., Swanstrom, L., Hungness, E., Soper, N., Kochman, M. L., et al. (2018). The novel trial: natural orifice versus laparoscopic cholecystectomy—a prospective, randomized evaluation. *Surgical Endoscopy*, 32:2505–2516.

Schwaner, K. L., Dall'Alba, D., Jensen, P. T., Fiorini, P., and Savarimuthu, T. R. (2021). Autonomous needle manipulation for robotic surgical suturing based on skills learned from demonstration. In *2021 IEEE 17th international conference on automation science and engineering (CASE)*, pages 235–241. IEEE.

Seeliger, B. and Swanström, L. L. (2020). Robotics in flexible endoscopy: current status and future prospects. *Current opinion in gastroenterology*, 36(5):370–378.

Selvaggio, M., Abi-Farraj, F., Pacchierotti, C., Giordano, P. R., and Siciliano, B. (2018). Haptic-based shared-control methods for a dual-arm system. *IEEE Robotics and Automation Letters*, 3(4):4249–4256.

Sentis, L. and Khatib, O. (2004a). Prioritized multi-objective dynamics and control of robots in human environments. In *4th IEEE/RAS International Conference on Humanoid Robots, 2004.*, volume 2, pages 764–780. IEEE.

Sentis, L. and Khatib, O. (2004b). Task-oriented control of humanoid robots through prioritization. In *IEEE International Conference on Humanoid Robots*, pages 1–16.

Sentis, L., Park, J., and Khatib, O. (2010). Compliant control of multicontact and center-of-mass behaviors in humanoid robots. *IEEE Transactions on robotics*, 26(3):483–501.

Sestini, L., Rosa, B., De Momi, E., Ferrigno, G., and Padoy, N. (2021). A kinematic bottleneck approach for pose regression of flexible surgical instruments directly from images. *IEEE Robotics and Automation Letters*, 6(2):2938–2945.

Shabanzadeh, D. M. and Sørensen, L. T. (2012). Laparoscopic surgery compared with open surgery decreases surgical site infection in obese patients: a systematic review and meta-analysis. *Annals of surgery*, 256(6):934–945.

Shahkoo, A. A. and Abin, A. A. (2023). Autonomous tissue manipulation via surgical robot using deep reinforcement learning and evolutionary algorithm. *IEEE Transactions on Medical Robotics and Bionics*, 5(1):30–41.

Shi, H., Sun, G., Wang, Y., and Hwang, K.-S. (2018). Adaptive image-based visual servoing with temporary loss of the visual signal. *IEEE Transactions on Industrial Informatics*, 15(4):1956–1965.

Siau, K., Hawkes, N. D., and Dunckley, P. (2018). Training in endoscopy. *Current Treatment Options in Gastroenterology*, 16:345–361.

Siegel, R. L., Wagle, N. S., Cercek, A., Smith, R. A., and Jemal, A. (2023). Colorectal cancer statistics, 2023. *CA: a cancer journal for clinicians*, 73(3):233–254.

Slotine, S. B. and Siciliano, B. (1991). A general framework for managing multiple tasks in highly redundant robotic systems. In *proceeding of 5th International Conference on Advanced Robotics*, volume 2, pages 1211–1216.

Soetikno, R. M., Gotoda, T., Nakanishi, Y., and Soehendra, N. (2003). Endoscopic mucosal resection. *Gastrointestinal endoscopy*, 57(4):567–579.

Song, S., Li, Z., Yu, H., and Ren, H. (2015). Electromagnetic positioning for tip tracking and shape sensing of flexible robots. *IEEE Sensors Journal*, 15(8):4565–4575.

Spada, C., De Vincentis, F., Cesaro, P., Hassan, C., Riccioni, M. E., Minelli Grazioli, L., Bolivar, S., Zurita, A., and Costamagna, G. (2012). Accuracy and safety of second-generation pillcam colon capsule for colorectal polyp detection. *Therapeutic advances in gastroenterology*, 5(3):173–178.

Spaun, G. O., Zheng, B., Martinec, D. V., Cassera, M. A., Dunst, C. M., and Swanström, L. L. (2009). Bimanual coordination in natural orifice transluminal endoscopic surgery: comparing the conventional dual-channel endoscope, the r-scope, and a novel direct-drive system. *Gastrointestinal endoscopy*, 69(6):e39–e45.

Stiff, G., Rhodes, M., Kelly, A., Telford, K., Armstrong, C., and Rees, B. (1994). Long-term pain: Less common after laparoscopic than open cholecystectomy. *British journal of surgery*, 81(9):1368–1370.

Stolzenburg, J.-U., Franz, T., Kallidonis, P., Minh, D., Dietel, A., Hicks, J., Nicolaus, M., Al-Aown, A., and Liatsikos, E. (2011). Comparison of the freehand® robotic camera holder with human assistants during endoscopic extraperitoneal radical prostatectomy. *BJU international*, 107(6):970–974.

Su, H., Yang, C., Ferrigno, G., and De Momi, E. (2019). Improved human–robot collaborative control of redundant robot for teleoperated minimally invasive surgery. *IEEE Robotics and Automation Letters*, 4(2):1447–1453.

Sullivan, G. M. and Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of graduate medical education*, 4(3):279–282.

Swanstrom, L. L., Volckmann, E., Hungness, E., and Soper, N. J. (2009). Patient attitudes and expectations regarding natural orifice translumenal endoscopic surgery. *Surgical endoscopy*, 23:1519–1525.

Tajbakhsh, N., Gurudu, S. R., and Liang, J. (2015). Automated polyp detection in colonoscopy videos using shape and context information. *IEEE transactions on medical imaging*, 35(2):630–644.

Tarbouriech, S., Navarro, B., Fraisse, P., Crosnier, A., Cherubini, A., and Sallé, D. (2022). An admittance based hierarchical control framework for dual-arm cobots. *Mechatronics*, 86:102814.

Than, M., Witherspoon, J., Shami, J., Patil, P., and Saklani, A. (2015). Diagnostic miss rate for colorectal cancer: an audit. *Annals of gastroenterology: quarterly publication of the Hellenic Society of Gastroenterology*, 28(1):94.

The International Agency for Research on Cancer (IARC) (2020). Global cancer observatory. https://gco.iarc.fr/today. Accessed: 2024-02-28.

Thele, F., Zygmunt, M., Glitsch, A., Heidecke, C.-D., and Schreiber, A. (2008). How do gynecologists feel about transvaginal notes surgery? *Endoscopy*, 40(07):576–580.

Thompson, C. C., Ryou, M., Soper, N. J., Hungess, E. S., Rothstein, R. I., and Swanstrom, L. L. (2009). Evaluation of a manually driven, multitasking platform

for complex endoluminal and natural orifice transluminal endoscopic surgery applications (with video). *Gastrointestinal endoscopy*, 70(1):121–125.

Tokuda, F., Arai, S., and Kosuge, K. (2021). Convolutional neural network-based visual servoing for eye-to-hand manipulator. *IEEE Access*, 9:91820–91835.

Tomczak, M. and Tomczak, E. (2014). The need to report effect size estimates revisited. an overview of some recommended measures of effect size. *Trends in Sport Sciences*.

Trauzettel, F., Vander Poorten, E., Ourak, M., Dankelman, J., and Breedveld, P. (2022). 3flex–3d printable parametric tendon-driven manipulator.

Trivedi, M., Klapheke, R., Youssef, F., Wolfe, S., Jih, L., Chang, M. A., Fehmi, S. A., Krinsky, M. L., Kwong, W., Savides, T., et al. (2022). Comparison of cold snare and hot snare polypectomy for the resection of sporadic nonampullary duodenal adenomas. *Gastrointestinal Endoscopy*, 96(4):657–664.

Van Gossum, A., Munoz-Navas, M., Fernandez-Urien, I., Carretero, C., Gay, G., Delvaux, M., Lapalus, M. G., Ponchon, T., Neuhaus, H., Philipper, M., et al. (2009). Capsule endoscopy versus colonoscopy for the detection of polyps and cancer. *New England Journal of Medicine*, 361(3):264–270.

Vandebroek, T., Ourak, M., Gruijthuijsen, C., Javaux, A., Legrand, J., Vercauteren, T., Ourselin, S., Deprest, J., and Poorten, E. (2019). Macro-micro multi-arm robot for single-port access surgery. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 425–432.

Varadarajulu, S., Tamhane, A., and Drelichman, E. R. (2008). Patient perception of natural orifice transluminal endoscopic surgery as a technique for cholecystectomy. *Gastrointestinal endoscopy*, 67(6):854–860.

Vargha, A. and Delaney, H. D. (2000). A critique and improvement of the cl common language effect size statistics of mcgraw and wong. *Journal of Educational and Behavioral Statistics*, 25(2):101–132.

Varier, V. M., Rajamani, D. K., Goldfarb, N., Tavakkolmoghaddam, F., Munawar, A., and Fischer, G. S. (2020). Collaborative suturing: A reinforcement learning approach

to automate hand-off task in suturing for surgical robots. In *2020 29th IEEE international conference on robot and human interactive communication (RO-MAN)*, pages 1380–1386. IEEE.

Vassiliou, M. C., Kaneva, P. A., Poulose, B. K., Dunkin, B. J., Marks, J. M., Sadik, R., Sroka, G., Anvari, M., Thaler, K., Adrales, G. L., et al. (2010). Global assessment of gastrointestinal endoscopic skills (gages): a valid measurement tool for technical skills in flexible endoscopy. *Surgical endoscopy*, 24:1834–1841.

Voros, S., Long, J.-A., and Cinquin, P. (2007). Automatic detection of instruments in laparoscopic images: A first step towards high-level command of robotic endoscopic holders. *The International Journal of Robotics Research*, 26(11-12):1173–1190.

Wagner, M., Bihlmaier, A., Kenngott, H. G., Mietkowski, P., Scheikl, P. M., Bodenstedt, S., Schiepe-Tiska, A., Vetter, J., Nickel, F., Speidel, S., et al. (2021). A learning robot for cognitive camera control in minimally invasive surgery. *Surgical Endoscopy*, 35(9):5365–5374.

Wang, H., Yang, B., Wang, J., Liang, X., Chen, W., and Liu, Y.-H. (2018a). Adaptive visual servoing of contour features. *IEEE/ASME Transactions on Mechatronics*, 23(2):811–822.

Wang, P., Berzin, T. M., Brown, J. R. G., Bharadwaj, S., Becq, A., Xiao, X., Liu, P., Li, L., Song, Y., Zhang, D., et al. (2019). Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut*, 68(10):1813–1819.

Wang, T. D. and Van Dam, J. (2004). Optical biopsy: a new frontier in endoscopic detection and diagnosis. *Clinical gastroenterology and hepatology*, 2(9):744–753.

Wang, Y., Liu, H., Ji, W., and Wang, L. (2018b). Realtime collaborating with an industrial manipulator using a constraint-based programming approach. *Procedia CIRP*, 72:105–110.

Wang, Y., Vina, F., Karayiannidis, Y., Smith, C., and Ogren, P. (2014). Dual arm manipulation using constraint based programming. *IFAC Proceedings Volumes*, 47(3):311–319.

Wang, Y. and Wang, L. (2020). Whole-body collision avoidance control design using quadratic programming with strict and soft task priorities. *Robotics and Computer-Integrated Manufacturing*, 62:101882.

Wang, Y.-F., Uecker, D. R., and Wang, Y. (1998). A new framework for vision-enabled and robotically assisted minimally invasive surgery. *Computerized Medical Imaging and Graphics*, 22(6):429–437.

Waye, J. D., Kahn, O., and Auerbach, M. E. (1996). Complications of colonoscopy and flexible sigmoidoscopy. *Gastrointestinal endoscopy clinics of North America*, 6(2):343–377.

WCH, W. C. H. (2020). General surgery practice.

Webster III, R. J. and Jones, B. A. (2010). Design and kinematic modeling of constant curvature continuum robots: A review. *The International Journal of Robotics Research*, 29(13):1661–1683.

Weede, O., Bihlmaier, A., Hutzl, J., Müller-Stich, B. P., and Wörn, H. (2013). Towards cognitive medical robotics in minimal invasive surgery. In *Proceedings of Conference on Advances In Robotics*, pages 1–8.

Wei, G.-Q., Arbter, K., and Hirzinger, G. (1997). Real-time visual servoing for laparoscopic surgery. controlling robot motion with color image segmentation. *IEEE Engineering in Medicine and Biology Magazine*, 16(1):40–45.

Wenfu, X., Lei, Y., Zhonghua, H., and Liang, B. (2019). Area-oriented coordinated trajectory planning of dual-arm space robot for capturing a tumbling target. *Chinese Journal of Aeronautics*, 32(9):2151–2163.

Wexner, S., Garbus, J., Singh, J., and Group, S. C. O. S. (2001). A prospective analysis of 13,580 colonoscopies: reevaluation of credentialing guidelines. *Surgical endoscopy*, 15:251–261.

Wilcox, B., Tso, K., Litwin, T., Hayati, S., and Bon, B. (1989). Autonomous sensor-based dual-arm satellite grappling. In *Proceedings of the NASA Conference on Space Telerobotics, Volume 3*.

Wilschut, J. A., Habbema, J. D. F., van Leerdam, M. E., Hol, L., Lansdorp-Vogelaar, I., Kuipers, E. J., and van Ballegooijen, M. (2011). Fecal occult blood testing

when colonoscopy capacity is limited. *Journal of the National Cancer Institute*, 103(23):1741–1751.

Winslow, E., Fleshman, J., Birnbaum, E., and Brunt, L. (2002). Wound complications of laparoscopic vs open colectomy. *Surgical Endoscopy And Other Interventional Techniques*, 16:1420–1425.

Wisanuvej, P., Gras, G., Leibrandt, K., Giataganas, P., Seneci, C. A., Liu, J., and Yang, G.-Z. (2017). Master manipulator designed for highly articulated robotic instruments in single access surgery. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 209–214. IEEE.

Wong, S. W. and Crowe, P. (2022). Factors affecting the learning curve in robotic colorectal surgery. *Journal of Robotic Surgery*, 16(6):1249–1256.

Wu, K., Wu, L., and Ren, H. (2014). An image based targeting method to guide a tentacle-like curvilinear concentric tube robot. In *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*, pages 386–391. IEEE.

Yan, L., Xu, W., Hu, Z., and Liang, B. (2020). Multi-objective configuration optimization for coordinated capture of dual-arm space robot. *Acta Astronautica*, 167:189–200.

Yan, L., Yuan, H., Xu, W., Hu, Z., and Liang, B. (2018). Generalized relative jacobian matrix of space robot for dual-arm coordinated capture. *Journal of Guidance, Control, and Dynamics*, 41(5):1202–1208.

Yang, X.-X., Fu, S.-C., Ji, R., Li, L.-X., Li, Y.-Q., and Zuo, X.-L. (2022). A novel flexible auxiliary single-arm transluminal endoscopic robot facilitates endoscopic submucosal dissection of gastric lesions (with video). *Surgical Endoscopy*, 36(7):5510–5517.

Yasunaga, T., Hashizume, M., Kobayashi, E., Tanoue, K., Akahoshi, T., Konishi, K., Yamaguchi, S., Kinjo, N., Tomikawa, M., Muragaki, Y., et al. (2003). Remote-controlled laparoscope manipulator system, naviot™, for endoscopic surgery. In *International Congress Series*, volume 1256, pages 678–683. Elsevier.

Yeung, B. P. M. and Gourlay, T. (2012). A technical review of flexible endoscopic multitasking platforms. *International journal of surgery*, 10(7):345–354.

Yeung, M., Sala, E., Schönlieb, C.-B., and Rundo, L. (2021). Focus u-net: A novel dual attention-gated cnn for polyp segmentation during colonoscopy. *Computers in biology and medicine*, 137:104815.

Yi, J.-s., Ahn, M. S., Chae, H., Nam, H., Noh, D., Hong, D., and Moon, H. (2020). Task planning with mixed-integer programming for multiple cooking task using dual-arm robot. In *2020 17th International Conference on Ubiquitous Robots (UR)*, pages 29–35. IEEE.

Yoshida, K., Kurazume, R., and Umetani, Y. (1991). Dual arm coordination in space free-flying robot. In *IEEE International Conference on Robotics and Automation, 1991. Proceedings*, number 1991 in 3, pages 2516–2521. Tohoku University.

Yoshida, N., Naito, Y., Hirose, R., Ogiso, K., Inada, Y., Fernandopulle, N., Kamada, K., Katada, K., Uchiyama, K., Handa, O., et al. (2014). Prevention of colonic spasm using l-menthol in colonoscopic examination. *International journal of colorectal disease*, 29:579–583.

Yoshizumi, F., Yasuda, K., Kawaguchi, K., Suzuki, K., Shiraishi, N., and Kitano, S. (2009). Submucosal tunneling using endoscopic submucosal dissection for peritoneal access and closure in natural orifice transluminal endoscopic surgery: a porcine survival study. *Endoscopy*, 41(08):707–711.

Zhang, Q., Prendergast, J. M., Formosa, G. A., Fulton, M. J., and Rentschler, M. E. (2020). Enabling autonomous colonoscopy intervention using a robotic endoscope platform. *IEEE Transactions on Biomedical Engineering*, 68(6):1957–1968.

Zhang, R., Zheng, Y., Mak, T. W. C., Yu, R., Wong, S. H., Lau, J. Y., and Poon, C. C. (2016). Automatic detection and classification of colorectal polyps by transferring low-level cnn features from nonmedical domain. *IEEE journal of biomedical and health informatics*, 21(1):41–47.

Zhang, Z., Rosa, B., Caravaca-Mora, O., Zanne, P., Gora, M. J., and Nageotte, F. (2021). Image-guided control of an endoscopic robot for oct path scanning. *IEEE Robotics and Automation Letters*, 6(3):5881–5888.

Zheng, B., Rieder, E., Cassera, M. A., Martinec, D. V., Lee, G., Panton, O. N. M., Park, A., and Swanström, L. L. (2012). Quantifying mental workloads of surgeons

performing natural orifice transluminal endoscopic surgery (notes) procedures. *Surgical endoscopy*, 26:1352–1358.

Zhu, D., Gedeon, T., and Taylor, K. (2011). "moving to the centre": A gaze-driven remote camera control for teleoperation. *Interacting with Computers*, 23(1):85–95.

Zhu, T. S., Godse, N., Clayburgh, D. R., and Duvvuri, U. (2022). Assessing the learning curve associated with a novel flexible robot in the pre-clinical and clinical setting. *Surgical endoscopy*, pages 1–10.

Zhu, Y., Xiong, H., Chen, Y., Liu, Z., Jiang, Z., Huang, R., Gao, F., Zhang, Q., Wang, M., Jin, Y., et al. (2021). Comparison of natural orifice specimen extraction surgery and conventional laparoscopic-assisted resection in the treatment effects of low rectal cancer. *Scientific Reports*, 11(1):9338.

Zinchenko, K., Wu, C.-Y., and Song, K.-T. (2016). A study on motion control of a robotic endoscope holder using speech recognition. In *2016 IEEE International Conference on Industrial Technology (ICIT)*, pages 1472–1475. IEEE.

Zorn, L., Nageotte, F., Zanne, P., Legner, A., Dallemagne, B., Marescaux, J., and de Mathelin, M. (2017). A novel telemanipulated robotic assistant for surgical endoscopy: Preclinical application to esd. *IEEE Transactions on Biomedical Engineering*, 65(4):797–808.

Zorn, L., Zanne, P., Nageotte, F., and De Mathelin, M. (2019). Motorised and modular instrumentation device and endoscopy system comprising such a device. US Patent 10,463,239.

Zulina, N., Caravaca-Mora, O., Liao, G., Gravelyn, S., Schmitt, M., Badu, K., Heroin, L., and Gora, M. J. (2021). Colon phantoms with cancer lesions for endoscopic characterization with optical coherence tomography. *Biomedical optics express*, 12(2):955–968.

Zuo, S., Chen, T., Chen, X., and Chen, B. (2022). A wearable hands-free human-robot interface for robotized flexible endoscope. *IEEE Robotics and Automation Letters*, 7(2):3953–3960.

GONZALEZ HERRERA, Jose Fernando

Flexible steerable intraluminal robots during complex procedures: assistance through high-level control

Résumé

Des robots intraluminaux flexibles et orientables ont été mis au point pour naviguer, inspecter et effectuer des procédures chirurgicales complexes à l'intérieur du système digestif du patient, réduisant ainsi l'incidence élevée et le taux de mortalité du cancer colorectal. Ces robots sont généralement composés de deux bras à l'extrémité du corps, formant une architecture cinématique couplée avec jusqu'à 10 degrés de liberté. Cette thèse se concentre sur le développement d'une assistance robotique pour relever les défis (l'architecture couplée et le nombre élevé de degrés de liberté) des robots intraluminaux flexibles et orientables dans des environnements chirurgicaux complexes, en réduisant les risque associés à des procédures chirurgicales complexes, telles que l'étape de dissection de la dissection sous-muqueuse endoscopique et le cas de la coloscopie et de la biopsie. Une modélisation des tâches chirurgicales est proposée afin de les formuler sous forme d'objectifs et de contraintes dans une commande basée optimisation. Une étude de validation avec 10 sujets pour le cas de la dissection montre une meilleure performance de la tâche chirurgicale avec une légère pénalité temporelle, en comparaison avec le mode manuel dans un environnement simulé. Dans le cas de la coloscopie et de la biopsie, les résultats montrent que pour un modèle de côlon réaliste, la tâche chirurgicale peut être exécutée automatiquement grâce au contrôle coordonné proposé. Les travaux futurs devront réaliser une validation approfondie des approches de commande proposées en vue d'un transfert préclinique et clinique.

Résumé en anglais

Flexible steerable intraluminal robots have been developed to navigate, inspect and perform complex surgical procedures inside the patient's digestive system, reducing the high incidence and mortality rate of colorectal cancer. Flexible steerable intraluminal robots are typically composed of two arms at the tip of the body, forming a coupled kinematic architecture with up to 10 degrees of freedom. This thesis focuses on the development of robotic assistance to handle the challenges (the coupled architecture and the high number of degrees of freedom) of Flexible steerable intraluminal robots in complex surgical environments, reducing the risk of correctly performing complex surgical procedures, such as the dissection stage of endoscopic submucosal dissection and the colonoscopy and biopsy case. We model the surgical tasks and formulate them as objectives and constraints in an optimization-based control framework based on the task-function approach. A validation study with 10 subjects for the dissection case shows a better performance of the surgical task at a time penalty, in comparison with the manual mode in a simulated environment. In the colonoscopy and biopsy case, results show that for a realistic colon model, the surgical task can be performed automatically by coordinated control in a benchtop experiment. Future work should perform in-depth validation of the proposed control approaches for preclinical and clinical transfer.