This thesis is submitted to Imperial College London in partial fulfilment of the requirements for the degree of Doctor of Philosophy. Except for where indicated, it presents entirely my own work and describes the results of my own research.
Declaration

I herewith certify that all material in this dissertation which is not my own work has been properly acknowledged.

Petros Giataganas

‘The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commerciacl No Derivitives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work.’ (Imperial College London PhD regulations 2012)
Abstract

Advances in surgery have had a significant impact on cancer treatment and management. Recurrence, however, is still a major issue, and is often associated with incomplete tumour removal. Thus far, histopathological examination is still the “gold standard” for assessing tumour resection completeness. However, it is operator-dependent and too slow for intraoperative use. Recently developed endomicroscopy techniques enable the acquisition of high resolution images at a cellular level in situ, in vivo, thus significantly extending the information content available intraoperatively. The miniaturised imaging probes incorporate flexible fibre bundles and allow the ease of integration with surgical instruments. However, manual control of these probes is challenging, particularly in terms of maintaining consistent tissue contact and performing large area surveillance of complex, deformable, 3D structures.

This thesis explores the use of surgical robots and robotically-assisted probe manipulation to provide stable, precise, consistent and dexterous manipulation of endomicroscopy probes for surgical applications. Following a discussion of image enhancement techniques, a first approach towards robotically-assisted probe manipulation using existing surgical robotic platforms is demonstrated in the form of multi-purpose, pick-up probes. They also incorporate novel force adaptive mechanisms for consistent tissue contact.

The development of bespoke, mechatronically-enhanced robotic devices is then presented. Firstly, a handheld robotic scanning device is proposed for breast conserving surgery, allowing accurate, high speed scanning over wide deformable tissue areas. An energy delivery fibre is integrated into the scanning mechanism for image-guided ablation or intraoperative marking of tumour margins. Secondly, a dexterous 5-degree-of-freedom robotic instrument is proposed for use in endoluminal microsurgeries. The instrument offers increased flexibility and by using a master-slave control scheme, we demonstrate how efficient, large area scanning over curved endoluminal surfaces can be performed. Finally, the fusion of ultrasound imaging with endomicroscopy is investigated through the development of a robotically-actuated articulated instrument for multi-modality image fusion.
Acknowledgements

I would like to start by expressing my special thanks to my supervisor Professor Guang-Zhong Yang, without whom this work would not have been possible. His leadership, vision, technical and clinical skills have truly helped me towards the completion of the work presented in this thesis. He has always set the example for everybody in the lab by working very hard and being able to provide support at any time of the day (and night). This PhD period was a life-changing experience for me which I owe mostly to Professor Yang. Thank you for believing in me from the first moment and giving me the opportunity to work in such a great and multi-disciplinary environment with so many talented, dedicated and truly inspiring people.

I would also like to thank my second supervisor Professor Lord Ara Darzi. I am really grateful that I had the opportunity to work alongside and learn from such an inspirational person, someone with fundamental contributions in minimally invasive and robot-assisted surgery. Furthermore, I would also like to acknowledge with much appreciation the crucial role of the Institute of Global Health Innovation, who supported this research work.

Throughout this thesis I have been fortunate to work closely with an exceptional physicist and friend - Mike Hughes. His technical abilities never cease to amaze me, alongside his never ending support to my work. Without his novel research in endomicroscopy and the multiple systems that he has built throughout these years, I would never have achieved the work presented in this thesis. Moreover, I will always be grateful for his invaluable contribution to proof-read this thesis and encourage me every day to keep working hard.

In addition, I have to mention that I have been lucky enough to work alongside and be motivated by Konrad Leibrandt and Piyamate Wisanuvej, two talented engineers with exceptional software and electronic skills. Together we faced many deadlines, late nights and weekends but it was always fun! I also have to mention George Mylonas and Valentina Vitiello for their invaluable advice and friendly support. You were an inspiration to me and thank you for being there whenever I needed. I owe many thanks to Jianzhong Shang who introduced me to design engineering and had a big impact on my work with his advice and guidance all these years.
I would like also to thank all these people that contributed in different ways to the work presented here and supported me throughout these years. For that I would like to thank Chris Payne, Lin Zhang, George Dwyer, Christos Bergeles, Khushi Vyas, Carlo Seneci, Gauthier Gras, Menglong Ye, Hedyeh Rafii, Maura Power, Pierre Berthet-Rayne, Su-Lin Lee, Dan Elson, Philip Pratt, Maria Grammatikopoulou, Alessandro Vandini, Matina Giannarou, Karen Kerr, Raphaele Raupp and Yakov Boglev.

The last years wouldn’t have been the same without Sofia. Thank you for making me feel home here in London and for your own unique support during the last crazy months!

Last but not least, I owe an enormous depth of gratitude to my family, particularly to my parents Eleni and Theocharis, my brother Nikos and my grandparents Petros and Marika. Their support, enthusiasm and encouragement towards what I do has always been incredible and I will be forever grateful. Σας ευχαριστώ για όλα!
'To avoid injuring your patient.'

_Hippocrates, 460 BC_


'One thing I know, that I know nothing.'

_Socrates, 5th century BC_


'Stay Hungry. Stay Foolish.'

_Steve Jobs, 2005_
Contents

List of Figures 25

List of Tables 26

Nomenclature 27

1 Introduction 28
  1.1 Introduction ................................ 28
  1.2 Motivation and Objectives ...................... 28
  1.3 Original Contributions .......................... 32
  1.4 List of Publications of the Work Presented .......... 34

2 Optical Biopsy for Intraoperative Tissue Characterisation 37
  2.1 Introduction ................................ 37
  2.2 Intraoperative Tissue Characterisation and Tumour Margin Delineation 38
  2.3 Optical Imaging and Biopsy ........................ 41
    2.3.1 Technical Requirements of in vivo Optical Imaging ........... 42
    2.3.2 Optical Imaging Modalities ........................ 43
  2.4 Endomicroscopy and its Clinical Translation .............. 47
    2.4.1 Approaches to Scanning Miniaturisation in Confocal Microscopy 47
    2.4.2 Fibre Optic Technology .......................... 54
    2.4.3 Image Processing for Fibre-based Endomicroscopy ........... 56
    2.4.4 Commercial Endomicroscopy Systems .................... 58
  2.5 Surgical Robots and Minimally Invasive Methods ............. 61
    2.5.1 Historical Developments .......................... 63
    2.5.2 Rigid Teleoperated Robotic Systems ................... 63
    2.5.3 Highly Articulated and Flexible Robotic Systems ........... 67
    2.5.4 Hands-on and Handheld Surgical Robotic Systems ............ 72
  2.6 Robotic-Assisted Endomicroscopy ........................ 78
  2.7 Discussions and Conclusions .......................... 84
7.3.1 Mechanical Design Evaluation ......................... 187
7.3.2 Endomicroscopy Scanning of a Large Surface Area .......... 188
7.3.3 Calibration of Endomicroscopy and Ultrasound ............... 189
7.3.4 Ultrasound Imaging ........................................ 190
7.3.5 Ex Vivo Test with Porcine Colon Tissue ..................... 191
7.4 Discussions and Conclusions .................................... 192

8 Conclusions and Future Research Directions .......................... 194
8.1 Contributions of this Thesis ..................................... 194
8.2 Future Research Directions ..................................... 198

A Optical Imaging Systems .......................................... 202
A.1 Virtual Slit-Scanning Endomicroscopy System .................. 202
A.2 Widefield Fluorescence Endomicroscopy ....................... 204

B Autonomous Scanning for Endomicroscopic Mosaicking and 3D Fusion 205
B.1 System Overview ............................................... 205
B.2 Results ......................................................... 208

C 3-DoF Robotic Instrument for Large-Area Transanal Microscopic Imaging 212
C.1 Mechatronic Design for Large Area Scanning .................. 212
C.2 Optimal Scanning Velocity ..................................... 215
C.3 Results and Validation .......................................... 215

D DH Parameters of the 6-DoF Robotic Instrument .................. 218

E DH Parameters of the Articulated Instrument for US and Microscopy Fusion 219

F Mechanical Details of the Presented Instruments .................. 220

G Image Reuse Permissions ........................................... 223

Bibliography ......................................................... 225
List of Figures

2.1 Cancer progression and tissue cells; normal cells, before changing to cancer cells, become abnormal (dysplasia). Dysplasia may or may not become cancer [1]. ........................................... 38

2.2 Comparison between medical imaging modalities in terms of the type of information they provide. ........................................... 39

2.3 Comparison between imaging modalities in terms of their resolution and penetration depth. ........................................... 40

2.4 Exemplar images from different optical imaging systems; (a-b) porcine oesophagus tissue, (c) cleaning lens paper and (d) stomach tissue. ........ 45

2.5 Fibre-based confocal laser endomicroscopy system overview. ............ 47

2.6 Scanning mechanisms for fibre-based endomicroscopic methods; (a) Proximal scanning; (b) Proximal scanning using a spatial light modulator; (c) Proximal line-scanning; (d) Proximal multi-point scanning; (e) Distal single axis scanning using a cantilever (electromagnetic or piezoelectric actuators) mechanism to scan the fibre across the surface of the static objective; (f) Distal single axis scanning using a cantilever (electromagnetic or piezoelectric actuators) mechanism to scan both the fibre and the objective and (g) Distal dual axis scanning using a 2D MEMS mirror. ........ 48
2.7 (a) Most popular types of fibres used in endomicroscopy, (b,c) Exemplar fibres (Schott leached fibre bundle and Fujikura fibre bundle), (d) Image showing the individual fibre cores in a fibre bundle.

2.8 Commercial endomicroscopy systems (a) PENTAX ISC-1000 system (image courtesy of Optiscan Pty Ltd, 2016), (a) Cellvizio Endomicroscopy system @2016 Mauna Kea technologies), (c) Optiscan Five-1 handheld instrument @2016 Optiscan Pty Ltd and (d) NinePoint Medical Nvision-VLE Imaging System (image courtesy of Farm, 2016).

2.9 General schematic of Computer Assisted Robotic Surgery. Adapted from [2]

2.10 Da Vinci® Xi robotic system from Intuitive Surgical Inc.; (left) surgeon’s console, (centre) patient-side cart with 4 robotic arms and (right) vision stack. @2015 Intuitive Surgical Inc.

2.11 Exemplar highly articulated / flexible robotic systems; (a) HARP ©2008 IEEE, (b) iSNAKE ©2011 IEEE, (c) SPRINT robot ©2014 ASME, (d) Titan SPORT system ©Titan Medical Inc., (e) Da Vinci SP ©Intuitive Surgical Inc. and (f) Hansen Sensei X ©Hansen Medical Inc..

2.12 Exemplar cooperatively-controlled and handheld instruments for orthopaedic surgery (a) Acrobot ©2006 IEEE, (b) Stryker Mako Surgical Arm ©Stryker Corp. and (c) Navio™ PFS system ©BlueBelt Technologies, Inc.

2.13 Exemplar motion-compensation instruments for micro-surgery (a) Steady-Hand robot ©2010 IEEE, (b) Initial prototype of Micron @CMU [3], (c) ITrem - Latt et al. @2009 IEEE, (d) SMART instrument ©2012 OSA, with permissions from the author (e) Micron - MacLachlan et al. @2012 IEEE and (f) Micron - Yang et al. @2012 IEEE.
2.14 Exemplar handheld instruments with robotic articulation (a) Dario et al. ©2000 IEEE, (b) He et al. ©2015 IEEE, (b) Payne et al. ©2014 Biomedical Engineering Society and (d) Hackethal et al. ©2011 reproduced from Hackethal et al. with permission from Springer. .......................................................... 76

2.15 Exemplar handheld instruments with force control or/and force feedback capabilities (a) Yuen et al. ©2010 IEEE, (b) Latt et al. ©2010 IEEE, (c) Lee et al. @2013 with permission from Springer, (d) Payne et al. @2012 IEEE, (e) Payne et al. @2013 IEEE and (f) Payne et al. @2014 IEEE. .......................................................... 77


2.17 Exemplar robotic articulated approaches for endomicroscopy (a) Newton et al., (b) Zuo et al. ©2015 IEEE, (c) Zuo et al. @2015 with permission from Springer, (d) Payne et al. ©2015 IEEE. .......................................................... 81

2.18 (a) Exemplar approaches for endomicroscopy with the Da Vinci robot by Brahmbhatt et al. @2015 Arch. Plast. Surg. and (b) by Patsias et al. @2014 with permission from Springer, (c) SPRINT robot with a point-based Raman sensor by Ashok et al. @2014 with permission from Willey and (d) “Endodrone” robot by Avila et al. @2015 IEEE. .......................................................... 84

3.1 Challenges associated with fibre bundle based technology that necessitate the existence of an image reconstruction algorithm; (a) honeycomb-like pattern artefact, (b) non-uniform circular image and (c) damaged or broken fibres. .......................................................... 88

3.2 Monochrome and coloured raw images acquired from a fluorescence confocal endomicroscope and a colour reflectance endomicroscope respectively. .......................................................... 90

3.3 Overall image reconstruction process. .......................................................... 90
3.4 Core pattern removal process for an endocytoscopic image; the same process if followed with any fibre-based imaging approach. .......................... 91

3.5 Triangle definition of barycentric coordinates (left) and interpolated result of a monochrome image for one exemplar triangle (right). .......................... 92

3.6 Intensity correction process; all core pattern-free images are multiplied by a 2D intensity map to reconstruct a uniformly illuminated microscopic image. The top images correspond to the reconstruction of the calibration image that when is multiplied by the intensity map, it renders it uniform. .......................... 93

3.7 Initial acquired image from cartilage structure using colour reflectance microscopy (left) and imaging results from the image reconstruction algorithm (right). .......................... 94

3.8 Exemplar imaging results from reconstructed microscopic images; (a) USAF resolution target, (b) porcine colon mucosa, (c) printed ink on paper, (d) unstained onion skin, (e) porcine small bowel mucosa and (f) porcine colon mucosa. .......................... 95

3.9 Normalised cross correlation algorithm. .......................... 97

3.10 Mosaic images with and without blending processing. .......................... 101

3.11 Exemplar spiral mosaics of lens cleaning paper (left) and porcine fat colonic tissue (right) using the NCC mosaicking approach (false colour is applied). .......................... 102

3.12 Large area lens paper mosaic using a robotic 3-DoF scanner. .......................... 102

3.13 Exemplar colour reflectance endomicroscopy mosaic of colonic tissue. .......................... 103
3.14 (a) Exemplar mosaic using the Radon Transform algorithm to recognise the relative rotation between each individual image frame and (b) (upper) the relative rotations used to generate each individual frame compared to the calculated using the radon transform and (down) their relative errors for each frame (errors $0.196 \pm 0.47$ degrees (normality is tested based on Lilliefors test)).

3.15 Mosaic created from large area spiral scan over lens tissue paper stained with acriflavine. Circled insets are shown at approximately 2.5X magnification. (Zuo et al. [4]©2014 IEEE)

3.16 (a) Checkerboard tracking in the stereo images and (b) augmented visualisation by overlaying the microscopic image on top of the tracked checkboard.

3.17 (a) FoV in the global coordinate frame of the endoscope, (b) overlay of the microscopic images on the reconstructed field-of-view, and (c) endoscopic mosaic rendered in 3D based on tracking.

4.1 (a) CAD design illustration of the single modality pick-up probe; (b-e) Snapshots of phantom and ex vivo experiments presenting the robotic manipulation of the probe with the Da Vinci® instrument and in (c) and (e) a snapshot of the user interface is depicted, where the KeyDot® marker was overlaid with the streaming microscopic images for augmented reality.

4.2 (a) CAD design illustration of the CLE/OCT pick-up probe; (b) demonstration of the correlation between the distance to the tissue surface measured in the OCT image (bottom row) and the quality of pCLE images (top row). As the surface moves closer to the probe, from left to right in the figure, the probe-to-tissue distance measured by the OCT probe becomes smaller while the pCLE image quality improves.
4.3 (Left) Computer-Assisted Drawing (CAD) model of the proposed drop-in probe and (right) fabricated instrument handled by the da Vinci Endowrist® Needle Driver.

4.4 Illustration of the proposed concept; CAD model of the proposed drop-in probe (left) and its section view (right) demonstrating the air flow path.

4.5 (a) Graph of a 60 s sample recording of the scale demonstrating that the adaptive probe proposed can apply consistent low magnitude forces to the surface of a scientific scale. Force measurements were made by recording a video of the scales as the probe was driven across. Then, from this video, each measured weight value was manually recorded every 0.22 s; (b) Frames from the recorded video presenting the adaptive probe (left) and the rigid probe (right) and their respective weight values in grams.

4.6 (a-b) Snapshots during the tissue task where the users had to scan, with the adaptive robotic probe (a) and a rigid probe (b), a large curved 3D surface and acquire consistent microscopy images. (c-d) Indicative microscopy images sampled during the tissue task with the adaptive probe (c) and the rigid probe (d).

4.7 Indicative mosaic results of all the subjects that participated at the tissue study, demonstrating the largest mosaics achieved with the adaptive probe and the rigid probe respectively. The mosaics depicted have been scaled with the same ratio to allow comparison of their lengths. Zoomed versions of microscopy segments are presented to assess the changes in the morphology of the cells, with and without the adaptive force probe.

4.8 Average size of generated mosaics from each user during the 1 minute allowance for the tissue task. The images correspond to the raw acquisition rate of 120 fps.
4.9 Average size of generated mosaics from each user during the 1 minute allowance for the tissue task. The images correspond to the downsampled acquisition rate of 30 fps. .......................... 124

5.1 Overview of the robotic system. ................................. 129

5.2 Robotic system overview along with the fabricated high-speed handheld robotic scanning device. ................................. 130

5.3 Framework overview of the proposed system. ..................... 131

5.4 (a) Exploded view of the handheld robotic scanning device and (b) three different scanning positions presenting the steel level positions with the resulting deflection on the tube. ................................. 133

5.5 Fabricated high-speed handheld robotic scanning device. .............. 134

5.6 Control schematic of the proposed system. .......................... 138

5.7 Shaft length analysis: (left) demonstration of change in axial position at the extremes of the instrument’s workspace and (right) axial change versus shaft length. ................................. 139

5.8 Analysis of the cam-roller mechanism positions: (left) four different position configurations explored during the analysis and (right) simulated workspace results for the four different configurations. ............... 140

5.9 Workspace analysis (a) Custom camera rig, (b) zoomed region of the tracked tip after binary conversion showing the tracked circle and its centre, (c) zoomed image presenting an exemplar tracked spiral trajectory and the instrument tip before binary conversion, and (d) overall workspace region with linear and non-linear parts. .......................... 141

5.10 Repeatability evaluation of two different scanning patterns along with the corresponding trajectory errors relative to the commanded trajectory, (a) three repeats of spiral trajectories, (b) four repeats of raster trajectories. 142
5.11 Visual servoing results on a grid pattern. (a) Grid along with a zoomed picture of the printed pattern, (b) a mosaic of approximately 1.4 mm in diameter without and with visual servoing control. A dashed line is shown for visual assessment of the faithful reconstruction of straight lines from the phantom. (c) The two spiral mosaics (orange tint) are overlaid manually on the printed pattern (grey) for visual comparison.

5.12 Visual servoing mosaic results when motion is introduced using the motorised translation stage, along with the graphs showing the current position (real position from the mosaic) and the desired (commanded) position. In open loop, the real position deviates from the desired trajectory whereas, with the visual servoing the real position follows the desired trajectory. The control trajectory shows the commanded output to the instrument. Pseudo-colour is applied to the mosaics for improved visualisation.

5.13 Mosaic results without and with visual servoing on the grid pattern phantom with random motion disturbance of 1.25 mm/s.

5.14 Mosaic results without and with visual servoing on lens paper and colonic tissue while an external disturbance is applied. (a) Lens paper with disturbance along x-axis of 0.1 mm/s, (b) colonic tissue with random motion disturbance of 0.1 mm/s. Pseudo-colour is applied on the lens paper mosaics.

5.15 Mosaic results for breast tissue scanning without and with visual servoing showing the difference in the scanning area covered.

5.16 Example of ex vivo tissue mosaics. (a) Human normal breast tissue (fibrous connective tissue), (b) human normal breast tissue (adipose tissue admixed with fibrous connective tissue), (c-d) human breast invasive carcinoma, (e) porcine large bowel and (f) porcine large bowel fat.
5.17 Example mosaics of lens paper while manipulating the instrument in a handheld manner or clamped on the robotic arm. Pseudo-colour is applied for improved contrast.

5.18 Example *ex vivo* mosaic results using the in-house endocytoscopy with human normal breast tissue and the dual-wavelength slit scanning confocal system with lens paper.

5.19 (a) Tip of the instrument presenting the CO\(^2\) laser and microscopy fibres, (b) instance from the laser ablation procedure on a paper card when the laser fired, (c) image from the custom tracking camera presenting the mechanical offset between the CO\(^2\) fibre (red) and the microscopy fibre (blue), (d) two instances overlaid, from different timestamps, presenting the movement by the offset value in order the red dot can coincide with the initial position of the blue dot, and (e) real-time mosaic, from the same region of a paper card, before and after the ablation in the centre of the spiral pattern.

5.20 Example of a real-time mosaic image from the same region of the grid pattern before and after the laser ablation. The grid pattern was then imaged through a bench-top microscope, presenting an ablation of 104 µm in diameter and thermal damage of approximately 210 µm.

6.1 (left) Transanal Endoscopic Microsurgery procedure, (centre) endoscopic images provide no pathological information that can assist in tumour margin identification and (right) the Micro-IGES robotic system.

6.2 Key characteristics of the Micro-IGES robotic assisted microscopy instrument.
6.3 (a) Interior design and components of the instrument’s actuation pack, (b) distal joint components, (c) rack and pinion mechanism between the rails of the Micro-IGES system and the instrument; translation stability is enhanced by rollers integrated on the case, (d) two channels for motor wires and microscopy optical fibres respectively and (e) an exploded view of the distal joint components showing the tendon routings and the fabricated pulley mechanism.

6.4 Tendon-routing of the articulated end-effector. Dashed lines depict that the tendons are in different planes and circles in the end of the tendons are the fixation points.

6.5 Motor controllers and interfaces.

6.6 Overall system framework.

6.7 Graphical User Interface.

6.8 Illustrations of the (a) instrument’s joints, (b) coupled mechanism in the distal part to constrain the 2nd and 4th link to move parallel to each other.

6.9 Illustrations of the instrument’s workspace in comparison to the available endoluminal workspace for a TEM operation. The latter is depicted in the left figure, from a side view, as a cylinder of 50 mm diameter and in the right figure, from the front view, as two cylinders of 40 mm and 50 mm diameter. In both figures the instrument is displaced in an offset position equal to its insertion point from the surgical port.

6.10 Workspace dexterity evaluation; red: high dexterity, blue: low dexterity.

6.11 (a) Backlash hysteresis experimental setup presenting the EM transmitter and the sensors with the extensions and (b) backlash hysteresis result for the translation joint showing the backlash area.

6.12 Backlash hysteresis results for each individual joint of the instrument.
6.13 Large area mosaics using tissue paper stained with fluorescence contrast agent (acrilavine). .......................... 173

6.14 Mosaicking results of the grid pattern over an approximately 26 mm$^2$ area (6.1 mm height/ 4.3 mm width). .......................... 174

6.15 (a) Experimental setup for ex vivo experiments using the TEM simulator and porcine colon tissue; the instrument inside the Micro-IGES rail is depicted and in the background the clinical graphical interface shows the real-time microscopic images and mosaics, (b) images from the experiments demonstrating insertion of the instrument using the master-slave control, assuming perpendicular orientation to the tissue and then scanning automatically once the tip is in place, and (c) images from an external camera showing the real-time mosaics and images from the clinical interface. .......................... 175

6.16 (a) Experimental setup using a curved artificial phantom and the master-slave control and (b) zoomed view of the curved phantom with the colonic tissue placed on top. .......................... 175

6.17 (a) Automatic raster scanning mosaicking results (first mode) and (b) real-time mosaicking results under completely manual control using the master device. .......................... 176

7.1 Computer rendering of the presented instrument, (a) entire instrument, (b) random configuration of the lateral DoF, (c) passive force adaptive mechanism mounted on scanning mechanism and (d) exploded view of the two part segment design. .......................... 180

7.2 Overall system framework of the presented robotic instrument. .......................... 182

7.3 Actuation pack assembly (left) CAD rendering along with individual mechanical components and (right) fabricated actuation pack. .......................... 184
7.4 (a) Bottom view of the instrument’s end-effector, showing the ultrasound image plane and the microscopy image and scanning plane. The mechanical offset between the two imaging planes is presented as well as the spatial alignment offsets. (b-c) Calibration phantom, real (b) and computer rendering (c) containing two spatially known features, easily identifiable by each modality. ............................... 187

7.5 Performance evaluation by endoscopic tracking; (a) Tracking information for repeated rotational trajectories (blue points), with fitted circle (red) for correlating the angle of the shaft to the encoder angle values; (b) the correlation between the position encoder sensing (blue) and the measured rotation angle (green) by the endoscope. ................................. 188

7.6 Comparison between mosaics using a fixed probe clamp (a) and passive force adaptive mechanism (b). Areas covered are $140\, mm^2$ and $126\, mm^2$ respectively. ........................................... 189

7.7 Calibration image showing the alignment between the mosaic and US image, and the spatial offset of $4.6\, mm$ between the centre of the images. .............................................................................. 190

7.8 3D reconstruction of kidney phantom, (a) experimental setup along with the 3D reconstructed overlay of the artificial tumour, (b) basic 3D reconstruction of the US dataset and (c) 3D reconstruction using advanced transfer function for visualisation. ........................................ 191

7.9 Ex vivo imaging test with porcine colon tissue experimental setup, (a) overview of the instrument during scanning, (b) details of the endoscopic tracking using a stereo endoscope, and (c) artificial phantom to simulate a potential tumorous structure to assess the US 3D reconstruction. .............................. 191
7.10 *Ex vivo* imaging of porcine colon tissue imaging study, (a) instance from the endoscopic view (operator’s view) presenting in real-time the microscopic images on the right and an overlay of the real-time US image on the US probe, (b) 3D reconstruction result of the US images presenting the details of the tissue but also the artificial tumour structure below (c) $4\text{mm} \times 12\text{mm}$ area mosaic showing surface detail of the colon including the crypts.

A.1 (a) Photograph and (b) schematic of the endomicroscopy imaging system, showing only the key components. Intermediate relay optics, and optomechanics are not shown for clarity. (c) Enclosed endomicroscopy system with custom fibre probe connected.

A.2 Enclosed widefield endomicroscopy system with a Fujikura fibre bundle attached.

B.1 A description of coordinate systems (frames) and related transformations. B: robot base frame; E: end-effector frame; M: marker frame; D: probe tip frame; C: camera frame.

B.2 3D fusion. The surface is densely interpolated on the area (colored as gray) for pathology analysis, and 2D-3D mapping is created so that visualisation of mosaics on 3D can be updated efficiently.

B.3 Exemplar 2D mosaicing results. (a) Mosaics of a 3x3 mm region on a kidney phantom; (b) Mosaics of a 2x2 mm region on an *ex vivo* porcine colon fat tissue sample.

B.4 Exemplar 3D fusion results. (a) Online fusion results and (b) multi-region fusion results of the phantom experiments, (c) and (d) online fusion results of the *ex vivo* colon experiment.

B.5 Example of 3D fusion results using microscopic and macroscopic visual servoing, (a) lens paper and (b) *ex vivo* porcine stomach tissue.
C.1 3D CAD renderings and photographs of the developed scanning instrument depicting: (1) the joint mechanism, (2) tip joint at various angles, (3) the rotational actuator for the shaft’s roll, (4) the microscopic probe’s channel, (5) the rotational actuator connected with a screw-based mechanism for driving the distal joint, (6) the tensioning mechanism for the tendons driving the joint, (7) the voice coil actuator for the shaft’s translation, (8) the workspace of the instrument while inside the patient, (9) the fabricated driver unit and (10) the fabricated instrument with casing.

C.2 Relative positions of endoscope, checkerboard pattern, and endomicroscopy probe.

C.3 Performance evaluation by probe tracking. (a) Repeated measured linear trajectories (black, blue red) compared with expected position (green line), (b) tracking information for repeated rotation-al trajectories (blue points), with fitted circle (red) for correlating the angle of the shaft to the encoder positions, (c) the correlation between the position encoder sensing (blue) and the measured rotation angle (green).

C.4 (a) Mosaic image of a USAF resolution test chart using 70 reconstructed images with 1 mm/s linear velocity and plot showing calculated position of each frame (blue points) with a best fit line (red), (b) 2D lens paper mosaic showing large area coverage, (c) \textit{ex vivo} colorectal tissue 2D mosaic and (d) \textit{ex vivo} colorectal tissue linear mosaic.

F.1 Pick-up mounts presented in Chapter 4; (a) Single modality pick-up probe dimensions and bottom view; (b) Multi-modality pick-up probe dimensions and bottom view; and (c) passive adaptive pick-up probe dimensions.
F.2 Handheld robotic scanning device presented in Chapter 5 overview; (a) full and cross-sectional view of the instrument presenting its main design features; (b) visualisation of the approximate workspace of the cantilevered tube; and (c) the scanning mechanism in different instances demonstrating how the two cam rollers push the triangular cam and as a result the cantilevered tube, that is rigidly attached to the cam, moves.  

F.3 Micro-IGES robotic assisted microscopy instrument presented in Chapter 6 overview; (a) dimensions of the articulated instrument’s end-effector; (b) main components of the proximal actuation and motor pack; and (c) overall dimensions of the presented instrument.  

F.4 Bottom view of the multi-modality robotic instrument presented in Chapter 7, showing the linear scanning range (13.6 mm) and the angular range ($\pm 66^\circ$); (left) different instances of the linear scanning mechanism presenting the microscopy probe scanned in parallel to the ultrasound image plane and (right) different instances of the US probe sweeping motion. The combined sweeping motion of the US with the linear scanning of the microscopy probe allow for multi-modality 3D image fusion.
List of Tables

2.1 Technical requirements for in vivo optical imaging .......................... 43

2.2 List of commercially available surgical robotic systems (as of September 2016) .......................................................... 68

4.1 Demonstration of each user’s expertise (+ = Limited experience, ++ = Intermediate experience, +++ = Expert) and total mosaic length in mm achieved by each user over a large 3D tissue surface in the period of 1 minute. .............................................................. 121

5.1 Comparison between the bench-top microscope’s measurements and the generated mosaics with and without visual servoing. .......................... 145

6.1 Individual joint ranges ............................................................... 165

6.2 Master-slave control, presenting how the joints and positions of the master device are mapped to drive each of the robotic instrument’s joints. .......... 167

D.1 Denavit–Hartenberg parameters of the 5-Dof robotic instrument. ........ 218

E.1 Denavit–Hartenberg parameters of the robotic instrument for US and pCLE fusion. ................................................................. 219
# Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCS</td>
<td>Breast Conserving Surgery</td>
</tr>
<tr>
<td>CLE</td>
<td>Confocal Laser Endomicroscopy</td>
</tr>
<tr>
<td>CT</td>
<td>X-ray Computed Tomography</td>
</tr>
<tr>
<td>DH</td>
<td>Denavit–Hartenberg</td>
</tr>
<tr>
<td>DoF</td>
<td>Degrees-of-Freedom</td>
</tr>
<tr>
<td>EM</td>
<td>Electromagnetic</td>
</tr>
<tr>
<td>FLIM</td>
<td>Fluorescence Lifetime Imaging</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional magnetic resonance imaging</td>
</tr>
<tr>
<td>FoV</td>
<td>Field-of-View</td>
</tr>
<tr>
<td>GI</td>
<td>Gastrointestinal Tract</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HRME</td>
<td>High-Resolution Microendoscopy</td>
</tr>
<tr>
<td>MEMS</td>
<td>Microelectromechanical systems</td>
</tr>
<tr>
<td>MIS</td>
<td>Minimally Invasive Surgery</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>NBI</td>
<td>Narrow Band Imaging</td>
</tr>
<tr>
<td>OCT</td>
<td>Optical Coherence Tomography</td>
</tr>
<tr>
<td>PA</td>
<td>Photo-acoustic</td>
</tr>
<tr>
<td>pCLE</td>
<td>Probe-based Confocal Laser Endomicroscopy</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PID</td>
<td>Proportional–Integral–Derivative</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SPECT</td>
<td>Single-photon emission computed tomography</td>
</tr>
<tr>
<td>TEM</td>
<td>Transanal Endoscopic Microsurgery</td>
</tr>
<tr>
<td>TPFM</td>
<td>Two-photon fluorescence microscopy</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Introduction

This chapter provides a brief overview of the thesis, highlighting the aims, motivation and original contributions of the PhD research.

1.2 Motivation and Objectives

Optical microscopy, a technology that dates back to the sixteenth century to the development of the first “compound” microscope by Zacharias Jansen, is a well-established topic in the life sciences community. A very common application of optical microscopy in modern medicine is the histological examination of discrete biological samples for diagnosis. The extraction of tissue samples can be performed either with a biopsy or as part of a surgical procedure. The excised specimens are then processed into thin slices (sections), stained, and examined under a light microscope or an electron microscope. This process, however, is discrete, and most suitable for offline, bench-top analysis [5]. The biopsy process is invasive and error prone and there is a significant clinical demand to reduce the invasiveness whilst permitting real-time diagnosis. A situation that these requirements are particularly relevant to is in oncological surgery where tumour margins
need to be accurately identified in order to intraoperatively guide tissue excision.

Only recently has there been an increased trend towards conducting *in vivo* optical microscopy. An example of apparently successful combination of advanced computer technology and miniaturised optics in a clinical environment is the development of endomicroscopy and its applications to tissues *in vivo*. These advances allow the acquisition of *in vivo* and *in situ* microscopic images through miniature optical probes at a cellular level resolution for real-time assessment of tissue pathology. Using optical fibres to relay light to and from the tissue allows probes to be compact and passive, while still delivering high resolution microscopy images at near video rates. In particular, endomicroscopy is an emerging technique for real-time, highly magnified, continuous visualisation of histological details for *in vivo* and *in situ* assessment of tissue pathology, a process known as “optical biopsy” [6]. Current optical biopsy technologies include a plethora of biophysical methods, including autofluorescence endoscopy, narrow band imaging, optical coherence tomography and confocal laser endomicroscopy [7].

A significant hurdle that these technologies face is the need for miniaturisation. The use of microscopic imaging for skin examination is not particularly challenging because there are no space restrictions. For internal organs, microscopic imaging probes need to pass through small working channels of endoscopes to image targets such as the gastrointestinal tract (GI). This presents tremendous challenges for the clinicians in terms of stability, ergonomics and image quality. These are factors that affect the diagnostic yield and accuracy of these techniques, thus influencing the clinical uptake.

Further motivation towards miniaturisation of existing optical microscopy technology is due to the radical changes in the way that surgical operations are performed. Increasingly, open surgical procedures are being replaced by minimally invasive surgery (MIS) and in the last decade further by robotic-assisted MIS for specific procedures. MIS, also known as keyhole surgery, is a pioneering surgical technique that allows operations to be performed through small incisions. The technique leads to a number of patient benefits including smaller incisions, which reduce trauma and pain, reduced blood loss, better cosmetic results, shorter recovery time, earlier return to work and reduced cost for the hospital. On the other hand, there are significant disadvantages such as: reduced haptic
feedback, loss of direct hand–eye coordination and reduced field-of-view (FoV) that can lead to dis-orientation problems. Furthermore, the procedure must be performed with long instruments that have to enter into the patient body and be manoeuvred around the entry point, called the fulcrum effect, restricting the motion further.

Advances in robotics and mechatronics in medicine and especially in surgical theatres are becoming more and more evident as technology evolves, allowing systems to either compensate for the weaknesses of the surgeons or fully assist them during the operation. For example, the use of a master-slave control paradigm allows the exploitation of an additional processing level between the surgeon and the robot to overcome many of the issues related to traditional MIS. The fulcrum effect is eliminated by the use of ergonomic user interfaces and computerised kinematic mapping. At the same time, precise motion control of the instrument tip is achieved by tremor removal and motion scaling. Instruments with wristed articulation are introduced, adding flexibility at the slave level. Recent advances in imaging and computerised vision technologies, such as high definition stereoscopic displays with augmented reality, have further enhanced the capabilities of intraoperative navigation [8, 9]. However, the high cost and large footprint of the current surgical robotic systems limit their uptake in many hospitals and healthcare facilities.

The aim of this thesis is to demonstrate the potential of combining surgical robotic instruments with endomicroscopy approaches, both from the software perspective through advanced robotic control and image fusion algorithms, and from the hardware perspective through stable miniaturised and sophisticated mechanisms that enhance the capabilities of the surgeons by providing intraoperative cellular level imaging with increased consistency.

The thesis begins with a brief overview of my PhD work in Chapter 1 including the motivation and original contributions of the work.

In Chapter 2, a detailed review of in vivo optical imaging technologies, including laser scanning methods and image processing techniques for cellular level image augmentation, will be presented. The chapter has a particular focus on endomicroscopy approaches. The chapter will also analyse the impact of robotic assisted MIS and aim to provide a
deeper understanding of how optical imaging, can be integrated seamlessly in the operating theatre.

In Chapter 3, endomicroscopic image enhancement techniques will be introduced and further image processing algorithms will be demonstrated with the aim of optimally producing histology-like images intraoperatively in the form of large area in vivo mosaics for tumour margin identification using real-time microscopic images. All these imaging approaches can be integrated with the developed robotic and non-robotic instruments and will assist the clinical translation of these instruments.

A first approach towards robotically-assisted probe manipulation will be described in Chapter 4. The design and implementation of three different pick-up endomicroscopy probes for existing surgical MIS and robotic systems are presented. Initially, a simple pick-up instrument that integrates an endomicroscope probe to facilitate scanning of large areas will be briefly described along with experimental results. This is then followed by a short description of a pick-up probe that includes an OCT and an endomicroscopy probe for multi-modality optical imaging. Finally, a simple passively adaptive pick-up probe that can be integrated with the da Vinci instruments for intraoperative endomicroscopy imaging will be presented. The device uses a novel low-friction air bearing with open loop adaptive axial force control to maintain constant contact between the tissue and the imaging probe, facilitating microscopy scans over complex and curved surfaces such as those encountered during in vivo tumour margin delineation. Detailed ex vivo experiments and results from user trials are presented to demonstrate the effectiveness of the technique.

In Chapter 5, a novel high-speed, high-precision robotic instrument for breast conserving surgery will be described. The instrument can be used in a handheld manner but can also be mounted to a robotic arm to improve its stability. The high-speed characteristics of the device allow it to scan areas as large as 16 mm² in less than 10 seconds, while being able to deliver intervention in situ by integrating a CO₂ laser ablation fibre. The optical fibres are interchangeable and therefore multiple endomicroscopy modalities can be integrated. Finally, to further improve the accuracy of mosaics and the mechanism itself, a microscopic visual servoing algorithm will be described and ex vivo tissue studies
will demonstrate the clinical applicability of the instrument.

In order to facilitate image mosaicking for endoluminal endoscopic microsurgeries, a new robotically controlled articulated micro-endoscope is proposed in Chapter 6. Compared to previously reported devices, this design allows larger areas of tissue to be scanned in a much shorter period of time. Results from \textit{ex vivo} colorectal tissues demonstrate efficient large area scanning with ease of user control. The proposed probe can be used in conjunction with standard laparoscopic instruments and is particularly suited for procedures where margin assessment is critical for the surgical outcome. The key features of the probe are its simple mechatronics and the ability to scan large curved surfaces. It has a small diameter shaft which means that it can be inserted through a standard surgical port, facilitating a wide range of applications in cancer surgery. A potential clinical application will be demonstrated in Transanal Endoscopic Microsurgery (TEM).

In Chapter 7, the integration of endomicroscopy and ultrasound imaging through a robotically actuated instrument will be presented. Endomicroscopy can provide high resolution images at a surface level while ultrasound provides depth resolved information at a macroscopic level. The instrument is manipulated across the surgical workspace, resulting in large area mosaics coupled with ultrasound scanning. In addition, the use of endoscopic tracking is demonstrated, allowing three-dimensional reconstructions of the ultrasound data to be overlaid onto the endoscopic view. \textit{Ex vivo} studies demonstrate the clinical applicability of the instrument.

Finally, Chapter 8 concludes the thesis by discussing the main technical contributions of the thesis, potential limitations of the methods developed, and future work that needs to carried out.

### 1.3 Original Contributions

The original contributions of this work can be summarised as follows:

1. An endomicroscopy imaging framework with real-time large area image mosa-
icking, enhancing the clinical usability of fibre-based endomicroscopic imaging (monochrome and coloured). The developed framework is used in various robotic devices and endomicroscopy approaches, offering high speed, real-time performance;

2. The development of a novel, simple yet versatile adaptive pick-up probe holder for any endomicroscopy fibre bundle, which can be used for robotic and non-robotic applications, allowing large mosaics to be generated over complex 3D surfaces;

3. The development of multi-functionality pick-up probes with integrated tracking, offering the potential to fuse multiple imaging modalities; this enables the combination of microscopic images with stereo-endoscopic images to create a multi-modality functional map of the operating site that facilitates intraoperative tumour margin identification and biopsy area re-targeting;

4. The development of a high speed robotically actuated instrument, that can be used either in a handheld manner or with a supporting robotic arm, with enhanced accuracy and control capabilities for breast conserving surgery. The instrument presents, for the first time, high speed scanning over wide areas together with interventional capabilities for real-time precise treatment;

5. The development of an articulated robotic scanning device offering wide area mosaics even for tubular structures as encountered in endoluminal surgeries. This can be either used in a handheld manner or connected to a robotic system that is operated remotely through a haptic device. This type of robotic manipulation allows smooth and consistent movements inside complex anatomies and the increased degrees-of-freedom (DoF) provide the appropriate dexterity to the surgeon to reach difficult areas of the anatomy. The instrument is developed for intraoperative image guidance for TEM, a very demanding surgical operation with proven benefits over current approaches in terms of invasiveness and patient’s surgical trauma;

6. The development of a robotically actuated instrument that combines ultrasound and endomicroscopy imaging, allowing simultaneous image acquisition, providing the surgeon with the 2D depth information of the ultrasound images and the cellular
level images of the endomicroscopy probe simultaneously, along with an enhanced visualisation environment created by the 3D reconstruction of the ultrasound data.

1.4 List of Publications of the Work Presented

The work presented in this thesis has resulted in the following publications (ordered in inverse chronologic order):


Conference on Robotics and Automation (ICRA)


Chapter 2

Optical Biopsy for Intraoperative Tissue Characterisation

2.1 Introduction

There are over 8 million cancer deaths and more than 14 million new cases each year worldwide, with the most common being lung, female breast, bowel and prostate cancer [10]. The survival rates are strongly related to the type and the stage at the time of presentation in the clinic that are determined through histopathological examination. Even though cancer treatment has been improved over the last 20 years, recurrence rates remain high for many types of cancer. One of the major reasons for cancer recurrence is that, during surgical tumour removal, not all cancer cells are removed and new tumours can be regenerated. This leads to the argument that early cancer detection and timely, direct and accurate treatment of cancer can significantly improve survival trends and recurrence rates, thus motivating the need for new diagnostic and treatment technologies to be developed and evaluated. In this chapter, current technologies for tissue identification and characterisation will be reviewed with a specific focus on optical imaging and technologies used in the operating theatre. This will be accompanied by the introduction of surgical robotics, its contribution to MIS and the benefits of combining robotics with optical imaging technologies for more effective diagnostics and cancer treatment.
Normal Cells

Dysplastic Cells

Cancer Cells

Figure 2.1: Cancer progression and tissue cells; normal cells, before changing to cancer cells, become abnormal (dysplasia). Dysplasia may or may not become cancer [1].

2.2 Intraoperative Tissue Characterisation and Tumour Margin Delineation

Intraoperative tissue characterisation has been a long-term endeavour in surgery. Surgical procedures for which intraoperative tissue characterisation is important include cancer and exploratory surgery. In cancer surgery, tissue characterisation is needed to determine resection margins and clearance of cancerous tissue, while in exploratory surgery in vivo tissue characterisation is desired to improve the accuracy of diagnostics and potentially minimise the operating time. Cancer diagnoses are focused on abnormal epithelial cells which are located in the deepest layer of the epithelium on the organs’ surfaces, creating cancers called carcinomas which are responsible for about 85% of cancers [10].

In the last 50 years, imaging modalities such as X-Ray computed tomography (CT) and magnetic resonance imaging (MRI) have offered a new perspective on the inside of the living human body, changing the way that clinicians treat their patients. As can be seen in Figure 2.2, conventional imaging modalities such as X-Ray, CT, MRI, and ultrasound (US) are mainly used to acquire anatomical information, whereas newer imaging modalities such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET) and single-photon emission computed tomography (SPECT) acquire functional
Medical Imaging Visualisation

<table>
<thead>
<tr>
<th>Anatomical</th>
<th>Functional / Metabolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical Imaging</td>
<td>Optical Fluorescence</td>
</tr>
<tr>
<td></td>
<td>fNIRS</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>PET</td>
</tr>
<tr>
<td>MRI / fMRI</td>
<td>SPECT</td>
</tr>
<tr>
<td>X-Ray / CT</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.2: Comparison between medical imaging modalities in terms of the type of information they provide.

and molecular information. Although these modalities are extensively used in different clinical scenarios, their resolution (lateral / depth) is limited and some require ionising radiation. Recent advances in optics, mechatronics and computer science have demonstrated that optical imaging could complement these modalities and provide cellular level information \textit{in vivo} and \textit{in situ}.

White light endoscopy is currently the primary optical imaging method for wide-area imaging. Hopkins’ rod lens endoscope was the first white light endoscope developed, facilitating the adoption of minimally invasive surgery. Advances in camera technology have since permitted high definition and magnifying endoscopy to be developed, providing improved resolution and increased pixel density. However, white light endoscopy still has certain intrinsic limitations as it visualises only reflected light from the superficial tissue surface. In the gastro-intestinal (GI) tract, for example, it is incapable of observing micro-vessels due to poor contrast with the background mucosa [11].

Currently, there are various enhancements made to white light endoscopy, which extend its usage and facilitate tissue identification [12], particularly in the GI tract. Conventional chromoendoscopy uses different local stains applied to the surface of the intestinal mucosa which improve tissue characterisation by enhancing image contrast [13]. Digital chromoendoscopy is an extension to conventional endoscopy, using different light filters
to mimic the behaviour of stains. For example, narrow band imaging (NBI) uses narrow (30 nm FWHM) sub-bands of light, instead of the full white light spectrum. Blue light exhibits high absorption in haemoglobin, unlike green and red light which penetrate further into tissue. This phenomenon highlights the surface vessel architecture and facilitates the identification of early neoplastic signs. Current commercial systems that provide chromoendoscopic features are NBI (Olympus, Japan), FICE (Fuji, Japan), i-Scan (Pentax, Japan) and EndoFlag (Endopix, Israel) [14]. Another advanced endoscopic technique that improves white light endoscopy is fluorescence or autofluorescence endoscopy, which aims to unveil lesions with neoplastic potential [15]. These techniques use intrinsic or exogenous fluorescent markers that, when excited with light of the correct wavelength, emit fluorescence at longer wavelengths. Differences in wavelength and intensity between normal and diseased mucosa allow demarcation of tumours and hence image-guided treatments. Popular systems that incorporate fluorescence endoscopy are the Lucera system (Olympus, Japan) and the LIFE system (Xillix, Canada) [16].

Figure 2.3: Comparison between imaging modalities in terms of their resolution and penetration depth.
pite the high sensitivity (> 90%) of these devices to diseased mucosa layers, it is easy to introduce false positive areas and therefore they have low specificity (> 70%). In more general terms, all the aforementioned imaging modalities based on macroscopic imaging endoscopy are limited by definition to providing macroscopic, i.e. relatively large, views of organs. Accurate tissue diagnosis, therefore, still requires biopsy and histopathological examination under microscopic imaging.

2.3 Optical Imaging and Biopsy

Optical imaging technologies - in terms of optical microscopy – have existed for more than 400 years. One of the main contributions of optical microscopy to medicine is in histopathology, a procedure that allows examination of cellular and nuclear morphology, cellular orientation and density, as well as the overall tissue architecture and cellular composition. Histopathology is still considered the “gold standard” for the diagnosis of many diseases. However, histopathological examination has some limitations, and in particular can cause significant delays in the analysis of the results. Liu et al. [17] summarised the issues of histopathology as the following:

- Since it is a discrete sampling method, the tissue samples removed from the area under investigation strongly depend on the operator, which can be a particularly problematic issue as it may not be a representative sample of the area.

- It is invasive and destructive.

- The results are dependent on the performance of the human observer (usually the pathologist) and there is a possibility of variations over time and between different observers.

- Results are generated typically after 1 – 2 weeks, making the process certainly not “real-time”.

- Special tissue preparation is needed that can introduce possible artefacts, affecting the morphological and/or molecular interpretation.
Alternatively, the cryosection technique (“frozen section”) has been introduced as the fastest method to get histopathological information. However, this method is still not real-time as the results take a significant amount of intraoperative time, more than 20 minutes, to be viewed. Additionally, the cryosection approach can create, occasionally, artefacts due to freezing, providing less reliability to its results [18].

To alleviate these problems, recent work has focused on bringing optical microscopy into the operating theatre to provide \textit{in vivo} tissue examination. This process is known as “optical biopsy”, referring to methods that are non-destructive and can be performed \textit{in situ}. This type of biopsy is a recent development, as the technical challenges related to the miniaturisation of microscopes were a significant hurdle [19]. Until recently, only systems for scanning easily accessible organs or the skin had been developed. However, technological achievements have since allowed further miniaturisation of these optical biopsy systems and imaging GI cancer and other internal organs is now feasible.

\subsection*{2.3.1 Technical Requirements of \textit{in vivo} Optical Imaging}

Before describing the technical advances of \textit{in vivo} optical imaging, the clinical needs and the technical requirements that engineers face when developing these technologies will be presented. Initially, the clinical applications can be generally grouped into three main categories: easily accessible surfaces / organs (e.g. eye, skin, neck), hollow structures / organs (e.g. oesophagus, colon, lungs) or surgically-accessed structures / organs (e.g. breast, kidney, brain). These clinical categories set different levels of technical requirements which are described in the Table 2.1.

From all the points listed above, it is apparent that the most essential technical requirements to propel optical imaging into the clinical realm include the optical head dimensions, frame rate, FoV and imaging depth along with the tether flexibility and multiple wavelength/imaging capabilities.
### Table 2.1: Technical requirements for in vivo optical imaging

<table>
<thead>
<tr>
<th>Easily accessible surfaces / organs</th>
<th>Hollow structures / organs</th>
<th>Surgically-accessed structures / organs</th>
</tr>
</thead>
<tbody>
<tr>
<td>No need of miniaturisation</td>
<td>Need of miniaturisation</td>
<td></td>
</tr>
<tr>
<td>Sufficient portability and manoeuvrability</td>
<td>Flexibility</td>
<td></td>
</tr>
<tr>
<td>Ergonomic and ease of use designs</td>
<td>Endoscope-compatible design</td>
<td>Laparoscope-compatible design</td>
</tr>
<tr>
<td>Large penetration depth</td>
<td>Integration with other surgical tools (e.g. diathermy)</td>
<td></td>
</tr>
<tr>
<td>Large FoV</td>
<td>High frame rate for motion compensation</td>
<td></td>
</tr>
<tr>
<td>Multi-colour or multi-imaging capabilities for improved diagnostics and targeted therapy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 2.3.2 Optical Imaging Modalities

Optical imaging technologies use the properties of light and its interaction with tissue to obtain morphological and structural details about the tissue. In general, systems can be categorised into image-based approaches or techniques which produce a spectroscopic or time-varying intensity signal from a single point on the tissue. In this section, a more detailed description will be provided of both categories, but with more emphasis on image-based modalities as these will be the main focus of this work.

**Optical Coherence Tomography** (OCT) is a high resolution, non-invasive optical imaging modality, providing real-time anatomical and histological details. Similarly to ultrasound, it can perform cross-sectional, three-dimensional and en face visualisation, offering axial and lateral resolution of approximately 10 $\mu m$ and 25 $\mu m$ respectively and penetration depths exceeding 2 $mm$ [20]. In contrast to ultrasound, though, it uses infrared light rather than acoustic waves and does not require tissue contact or a transducing medium. OCT is most widely used in ophthalmology, for imaging both the posterior and anterior segments of the eye, but there has been progress in translating OCT to areas such as the gastrointestinal tract and urology [21, 22, 23] using fibre-based approaches. An inter-
esting approach to screen Barrett’s oesophagus disease has been developed recently to map volumetric information over a large circumferential area of the oesophagus using a balloon-based OCT catheter [24]. With this volumetric OCT approach, the endoscopists are able to measure the Barrett’s epithelium thickness and detect the buried glands that could be a marker of disease recurrence [25].

**Photo-acoustic (PA) imaging and microscopy** methods use short laser pulses to generate heat due to the optical absorption in the tissue, leading to its thermal expansion and generation of acoustic signals. The excitation wavelength is tuned to specific absorbers (e.g. haemoglobin) to maximise contrast. The ultrasound signal is detected by a transducer and used to form an image. The technique can be used for a number of clinical applications including monitoring blood vessel growth in tumours, detecting skin melanomas, and tracking blood oxygenation in tissues [26].

**Optical spectroscopy** is based on wavelength-dependent interactions of light with the tissue. Techniques includes measuring elastic scattering and absorption as a function of wavelength, as well as measuring the spectral output of non-elastic scattering, such as fluorescence and Raman scattering (where light can be absorbed and re-emitted or scattered at longer wavelengths). These spectra can be correlated to changes in histological features or the presence of exogenous or endogenous fluorescent compounds (fluorophores). Spectroscopic methods have been explored as in vivo diagnostic tools for characterising cancer cells and early malignant changes at a molecular level [27, 28, 29] but the full clinical value has not yet been determined.

**Fluorescence Lifetime Imaging (FLIM)** is a time-resolved spectroscopic technique that, instead of using the intensity of fluorescence coming from the imaged sample, uses the lifetime of the fluorophore signal (by measuring its exponential decay rate). This requires the light source to be pulsed or rapidly modulated and the resulting measurement is less prone to wavelength-ratiometric artefacts such as fluorophore concentration, non-uniform illumination, tissue deformation and scattering. FLIM can be combined with other optical imaging approaches such as confocal microscopy to add molecular contrast to microscopic images by displaying the fluorescence decay profiles [30, 31].
Figure 2.4: Exemplar images from different optical imaging systems; (a-b) porcine oesophagus tissue, (c) cleaning lens paper and (d) stomach tissue.

*Epifluorescence microscopy, or widefield fluorescence microscopy* is the simplest approach for fluorescence imaging as the sample is flood-illuminated excited by a specific excitation light wavelength and the emitted fluorescent light is detected at an eye-piece or a camera. It is one of the cheapest microscopy modalities and it allows viewing either directly by eye or with high frame-rate video, even allowing visualisation of red blood cell flow without the need of laser sources or fast scanners. The main disadvantage is that there is no optical sectioning; meaning that all out of focus light is detected, the effective imaging depth is low and the overall signal to noise ratio is poor [32] compared to other fluorescence optical imaging techniques. Epifluorescence microscopy can be found also in the literature under the name of *High-Resolution Microendoscopy (HRME)*.

**Structured Illumination** Microscopy, which involves illumination of the sample with a set of three shifted line patterns, allows removal of out-of-focus light by demodulation, adding optical sectioning to epifluorescence microscopy. Recently, the approach was used to accomplish optical sectioning in endomicroscopy [33, 34]. The drawback of structured illumination is the low frame rate (three images are acquired to obtain one final image), the introduction of artefacts due to inaccurate pattern illumination and motions between acquisitions, and noisier images due to digital subtraction.

**Two-photon fluorescence microscopy (TPFM)** is a novel imaging approach with high penetration ability and three dimensional optical sectioning capability. TPFM typically uses near-infrared femtosecond laser pulses to excite fluorescence from the tissue. The
concept is based on the fact that two photons of comparably lower energy can also excite a fluorophore normally excited by a photon with higher energy (i.e. its absorption peak is at a shorter wavelength) if a near-simultaneous absorption occurs. This double absorption is likely only near the focal plane, providing optical sectioning. Despite the important fact that TPFM requires an expensive high power, ultra short-pulse laser, the key advantage of TPFM is the robustness to photo-bleaching, scattering and increased penetration compared to confocal microscopy [35, 36, 37].

A recent emerging optical biopsy technique is **Endocytoscopy** that provides microscopic images of live cells with the assistance of magnifying optics [38]. It provides similar images to fluorescence and widefield imaging but the contrast of its images is derived by the scattering effects of white light within the region of interest. This imaging approach was initially introduced by Olympus with a flexible, catheter-style endoscope [39]. Recently, fibre-based endocytoscopy approaches have been developed that can be combined with other types of endomicroscopy simultaneously [40].

The optical imaging method that will be presented in this thesis is based on **Confocal Microscopy**. Marvin Minsky invented the confocal microscope in 1955 and it was termed as “confocal” because the objective lens and the condenser had the same focal point [41]. Confocal microscopy is a form of scanning microscopy, where the sample is imaged point-by-point, and uses a detector pinhole aperture to reject out of focus light, i.e. light coming from either above or below the focal plane. In particular, a laser scanning confocal microscope uses a laser beam focused into a small focal volume within the sample by the objective lens. Reflected and fluorescence light is emitted back by the sample, collected by the objective lens, and imaged onto a detector through a pinhole which obstructs light coming from out of focus points. This light spot corresponds to one image pixel in the resulting image. In order to reconstruct a whole microscopic image, a 2D scanning mechanism is used to scan this point in the lateral plane for 2D imaging and also in the axial plane for 3D imaging. The main drawback of confocal microscopy is that penetration depth is limited to 50 – 100 µm and light scattering reduces image quality.
Figure 2.5: Fibre-based confocal laser endomicroscopy system overview.

Apart from the camera-based or point-based approaches described in the previous section, most of the optical imaging methods require a scanning method to form either two dimensional (2D) or three dimensional images (3D) through fibres, especially the ones that require point-by-point scanning. Therefore, in the following section, an overview of different scanning approaches proposed or developed for fibre-based optical imaging methods will be presented; with a specific focus on confocal laser endomicroscopy since it is the main topic of this work.

2.4 Endomicroscopy and its Clinical Translation

2.4.1 Approaches to Scanning Miniaturisation in Confocal Microscopy
Figure 2.6: Scanning mechanisms for fibre-based endomicroscopic methods; (a) Proximal scanning - cascaded galvanometer mirrors that scan the excitation laser across the fibre bundle’s proximal surface; (b) Proximal scanning using a spatial light modulator to illuminate separate pixels across the 2D proximal face of the fibre bundle; (c) Proximal line-scanning as the illumination is performed across a line in the fibre bundle rather than a point; (d) Proximal multi-point scanning using Nipkow disks (lens and pinhole); (e) Distal single axis scanning using a cantilever (electromagnetic or piezoelectric actuators) mechanism to scan the fibre across the surface of the static objective; (f) Distal single axis scanning using a cantilever (electromagnetic or piezoelectric actuators) mechanism to scan both the fibre and the objective and (g) Distal dual axis scanning using a 2D MEMS mirror which pivots in two directions.
Nowadays, optical biopsy is feasible in vivo using confocal microscopy principles, by inserting a fibre (or fibre bundle) between the laser source/microscope optics and the objective lens. This facilitates the development of flexible, fibre-based, confocal microscopes, which can be inserted into the limited working channels of flexible endoscopes (microendoscopes) and perform real-time in vivo biopsies. Confocal Laser Endomicroscopy (CLE), as this technology is mainly called, aside from its sophisticated miniaturised optical components, requires precise and high-speed scanning mechanisms to scan the laser spot over the tissue. The scanning can either be proximal or distal to the light guide. A proximal scanning mechanism is integrated in the system side, while distal mechanisms are placed at the endoscopic tip. Proximal scanning systems use a coherent image guide (typically a fibre bundle) to transmit the scanned pattern from the scanning system to the tissue, and the image from the tissue back to the system [42]. The advantage of having a proximal scanning mechanism is that the size of the scanning system is not limited and does not affect the design of the imaging probe. However, since it relies on a fibre bundle, the resolution is limited by the number of fibres in the bundle, typically ranging from 10,000 to 100,000 fibres. By comparison, a distal end scanning device uses a single, scanned fibre to transmit the image to the system, providing better image resolution and quality.

The most common scanning mechanism used for proximal scanning is the combination of two cascaded galvanometer mirrors, one for horizontal scanning (resonant scanner - high speed up to \(4kHz\)) and one for vertical direction (lower speed), resulting in a raster scan pattern across the surface of the fibre bundle (see Figure 2.6(a)). The fact that illumination has to be swept across the bundle to illuminate each separate fibre can result in simultaneous coupling into adjacent cores, reducing the image resolution and contrast. This point-by-point scanning also limits the frame rate of systems (typically < \(12Hz\)) which is problematic when motion is introduced; motion artefacts are prominent in point scanning systems especially for in vivo or handheld clinical operations [43]. The resonant mirror’s speed, as a result, is the limiting factor for this scanning method. Some exemplar research integrations of point-by-point scanning devices can be found in the work of Knittel et al. [44] with acquisition rate of 0.5 fps and Sung et al. [45] with 15 fps.

49
To alleviate the speed issue of the point-by-point scanning methods, methods that parallelised the illumination and detection paths have been developed. A first approach is to use multi-point scanning methods such as the Nipkow disks [46], two rotatable disks with a series of small apertures arranged in a pattern (see Figure 2.6(d)). Here, the manufacturing complexity is a limiting factor, and the use of discrete holes results in image quality issues due to a trade-off between light-throughput and optical sectioning. Another scanning approach is slit-scanning (see Figure 2.6(c)), this incorporates a line illumination system and a slit aperture detector to scan larger areas and increase the image acquisition speed with slight degradation, however, in axial resolution. In the work of Sabharwal et al. [47], a slit-based fluorescence confocal microendoscope has been fabricated that incorporates a scanning mirror that scans an illumination line across the bundle, scans the returning fluorescent light through a fixed detection slit aperture and also sweeps the collected, spatially filtered light onto a 2D CCD camera to recreate real-time microscopic images. Rouse et al. [48] added a dispersing prism to this setup to collect both spectral and spatial information and increase the diagnostic yield. This work was further improved and used in vivo by Rouse et al. [49, 50]. Compared to the previous work’s limited frame rate of 30 fps, Hughes et al. [43] accomplished a higher frame rate of 120 fps using a linear CCD detector (1D line-scan camera). In a further variation, Dwyer et al. [51] used a theta line scanning approach that uses a second objective lens to collect the fluorescence light at an angle to the illumination axis. An alternative solution for the line scan mechanism, provided by Lin et al. [52], used an incoherent fibre bundle to multiplex and de-multiplex light towards and from the tissue. The system therefore combines the simplicity of a slit scan mechanism with the increased contrast of the point-scan technique. To summarise, the main disadvantage for line-scanning systems is that along the focal line the confocality is lost, resulting in reduced optical sectioning and limited imaging depth. However, comparing line-by-line with point-by-point scanning approaches, line scanning slightly reduces the complexity of the mechanism and can potentially increase the signal-to-noise ratio (SNR) for a given frame rate and FoV compared to point-by-point systems by increasing the pixel dwell time [53]. In the study of Wang et al. [54], the line scanning approach provided comparable imaging results to point-by-point approaches at shallow depths while being superior under low signal conditions, e.g. weak fluorescence. Finally,
as will be presented in the next chapters, due to the improved frame rate of line scanning systems, the mosaicking of microscopic images is improved significantly, providing enlarged FoV images for real time assessment.

Spatial light modulator (SLM) techniques, on the other side, use advanced optics or mechanical components such as pixilated micro-mirror arrays with fast responses and high accuracy to illuminate individual fibres across the bundle without illuminating the non-imaging space between the fibres (see Figure 2.6(b)). Significant work on spatial light modulator technique has been conducted by Lane et al. [55] that showcased the significant improvements in optical sectioning and contrast compared to conventional point scanning systems. The major drawback of this technology is the low acquisition speed with limited FoV [41]. Thomson et al. [56] used an SLM for optical phase conjugation to correct phase aberrations and implement scanning and focusing at the distal end of the bundle without any optical or mechanical components, but this requires a prior calibration that is lost when the fibre is bent.

Distal scanning mechanisms allow the use of single or two optical fibres for microscopic image acquisition. Light is transmitted for illumination through a single optical fibre which serves as the pinhole in conventional confocal microscopes, a distal scanning mechanism scans the laser spot over the examined tissue and the light is collected again by the same or another optical fibre. Depending on the coincidence or not of the illumination and collection paths, confocal fluorescence imaging can be either single axis (SAC) or dual axis (DAC). DAC systems require lower numerical aperture objective lenses compared to SAC systems, resulting in deeper working distance and reduced off-axis aberrations. Initial work on single fibre confocal microscopy was performed by Giniunas et al. [57] and Juskaitis et al. [58, 59], showing that the confocal pinhole could be replaced with an optical fibre. Distal scanning mechanisms, initially, included electromagnetic or piezoelectric actuators in the form of a cantilever mechanism to scan the fibre and objective lens at mechanical resonant or non-resonant manner (see Figure 2.6(e-f)). Several implementations used bimorph piezoelectric actuators [60, 61, 62, 63] where scanning patterns, such as Lissajous, are applied by displacing a single mode fibre at either resonance or non-resonance relative to the objective or both together. With the second approach, LeHarzic
et al. [62] eliminated the off axis aberrations and increased the FoV, managing to cover a 420x420 \( \mu m \) area, at the expense of reduced frame rate (0.1 \( fps \)). Other cantilever based implementations utilised piezoelectric tubes (piezotubes) that could scan more complex and controlled trajectories in similar dimensions as biomorphs [64, 32, 65, 66]. Spiral pattern mechanisms at resonant speeds provided miniaturised dimensions with increased video rates (e.g. up to 1 \( mm \) diameter with approximately 200\( \mu m \) diameter FoV and 15 \( fps \) [67]) but suffered from short pixel dwell times as well as non-uniform spatial coverage in comparison to non-resonant scanners [68].

Later, these mechanisms were replaced by microelectromechanical systems (MEMS). Dickensheets et al. [69, 70, 71], initially, pioneered and developed micro-machined silicon scanning mirrors that can fit into a miniaturised confocal probe, providing two dimensional images with aberration correction and focus control (see Figure 2.6(g)). An alternative method to micro-machined silicon scanning mirrors was demonstrated in the works of Fu et al. and Seo et al. [72, 73] that used electrothermal MEMS silicon micro-actuators with an optical fibre directly mounted on them, accomplishing miniature footprints of 5\( mm \) and 1.28\( mm \) respectively, but with reduced frame rates of 5 and 7 \( fps \). Electrostatically actuated MEMS, then, presented improved performances with lower power consumptions, smaller sizes and higher driving forces, especially in vertical comb actuators [32]. Murakami et al. [74] fabricated a 3.3 \( mm \) diameter confocal endomicroscope that incorporates a single MEMS electrostatic gimbal scanning mirror and can be passed through a working channel of an endoscope. Different approaches followed, such as by Hoffmann et al. [75] who developed a two-axis mechanism with electrostatically driven torsional micromirrors and electroplated nickel hinges and frames, whereas authors in [76, 77, 78] followed a more established technique of gimballed scanner using a special process of MEMS devices (SOIMUMPS).

A significant work in the field of DAC MEMS systems has been presented by Ra, Piyawattanametha and Liu. Piyawattanametha et al. [79, 80] developed a handheld dual axes near infrared microscope of 10\( mm \) diameter and with 3D imaging capability based on the MEMS work of Ra et al. [81], while, in 2012, Piyawattanametha et al. [82] fabricated a micro-endoscope that can be inserted through a 6\( mm \) channel of a conventional endo-
scope and provide 3D high-resolution cross-sectional images at the micron scale due to an integrated micro-motor at the distal end. Piyawattanametha et al. [83, 84, 85] demonstrated the use of vertical electrostatic MEMS comb-drives in two photon imaging in combination with femtosecond laser microsurgery in a compact 9.6 mm imaging probe. Similar handheld approaches were demonstrated by Liu et al. [86, 87], who developed a multi-colour, 2 mm diameter, DAC microscope using a gradient-index (GRIN) relay lens, while Yin et al. [53] developed a line-scanning approach in a 12 mm, handheld, DAC microscope, achieving 16 fps with a 300x300 μm FoV. The highest frame rate and resolution achieved to date with electromagnetic MEMS scanning mirrors has been achieved in the work of Arrasmith et al. [88] where in vivo imaging was possible with 56 fps and a resolution of 1.14 μm axially and 0.39 μm laterally.

MEMS approaches are used in most scanning optical imaging modalities such as confocal [76, 89, 90], OCT [91, 92] and multi-photon microscopies [83, 84]. In general, all previous DAC implementations of microscopy provided improved axial, lateral resolution and optical sectioning capabilities, leading to increased contrast and imaging depth [54] compared to most of the SAC and proximal scanning approaches, at the cost of, however, complex scanning mechanisms and miniaturisation. A important drawback of MEMS scanners is that, although they are ideal for miniaturization, much expertise is required for their fabrication.

As mentioned before, confocal endomicroscopy and other optical imaging methods provide optical sectioning abilities, meaning that only light from the focal plane is collected and everything else is blocked. This is desirable for improved quality images, but it does not solve the problem of axial motion of the target; rapid re-focusing techniques are then required. Additionally to this, volumetric imaging in vivo could extend the capabilities of endomicroscopic approaches and provide significant advantages over histological examinations. For these reasons, axial scanning methods developed in parallel with the lateral scanning approaches. Initially, Sung et al. [45] moved the objective lens relative to the tissue by moving a piston with a hydraulic, oil-based approach; however, air bubbles and motion artefacts due to suction hindered its function. Rouse et al. [49] improved the focus mechanism by using a pneumatic mechanism and moving the fibre with respect to
the objective lens; this method allowed a scan of 200 $\mu m$ but introduced hysteresis effects and slow actuation. With the miniaturisation of mechatronics and development of micro-motors, Flusberg et al. [93] used a small DC micro-motor for two-photon fluorescence microendoscopy for animal, brain imaging, while MEMS confocal approaches integrated miniature piezoelectric motors, not only for re-focusing [88], but, also, volumetric imaging up to 90 $\mu m$ depth [89]. With miniature motors the hysteresis effects and position accuracy were greatly improved, leading to embodiments where positioning accuracy was better than the axial resolution [50] in a 5 $mm$ housing. Recently, there has been increasing interest in eliminating the mechatronic parts distally and integrating mechanisms that can optically change the focal plane, generally called tunable focus lenses. Suggested embodiments used either custom-made vari-focal objectives, with water chambers that can control the curvature of an elastic membrane, thus altering the focus [94, 95], or MEMS tunable microlenses [96] using MEMS bimorphs or electrically tunable lenses [97, 98, 99, 100]. The latter provided promising results in terms of both scanning accuracy and range ($> 200 \mu m$) as well as very fast response times ($< 2 ms$) in very compact footprints. The drawback, however, of this approach is that the increased curvature of the lenses, especially in the limits of their range, introduced aberrations in the microscopic images and reduced lateral resolution.

### 2.4.2 Fibre Optic Technology

For optical imaging, the development of fibre optics played an essential role. Fibres can either remotely deliver and collect light in constrained places, while maintaining the advantages of optical sectioning and fluorescence detection, or transmit whole images (using fibre bundles). These features allowed fibre technology to be used for basic biological research as imaging of freely moving animals was possible either by handheld imaging or implanted devices, as well as for minimally invasive clinical diagnostics and surgical operations, as optical fibres can commonly range from 80 $\mu m$ to several millimetres in diameter [32]. In confocal microscopy, apart from the remote delivery of light, fibres in distally scanned approaches could be used as pinholes, automatically aligning the two confocal pinholes by using optical fibres as both the light source and the detector, resulting in a less complex system [101, 41].
Proximal scanning methods were greatly assisted by the development of fibre bundle technology as it allowed mechanical flexibility and significant distal miniaturisation. Fibre bundles consist of up to approximately 100,000 individual step-index fibres in a closely packed arrangement, hundreds of micrometers to a few millimetres in total diameter. Step index fibres provide light transmission by total internal reflection and can guide either a single or multiple spatial modes of light. The fibres are coherent and they maintain the relative arrangement / position in the bundle, allowing the reconstruction of the remote image in a pixelated form; each fibre corresponds to one light intensity and therefore one pixel value. This is an advantage for proximal scanning methods but the drawbacks of fibre bundles are that:

1. this inherent “pixelation” reduces the quality of the images and additional image processing algorithms should be used to eliminate this artefact, and

2. the close adjacency of fibres means that the cladding is very thin and therefore optical cross-talk is possible in which light from neighbouring fibres can leak into one another, resulting in reduced contrast.

A special category of multi-mode fibres is the gradient refractive index (GRIN) fibre
which has the characteristic that the refractive index declines quadratically as the distance increases from the central axis. This allows the electromagnetic spatial modes of light to be transmitted at nearly the same velocity along the fibre. This makes GRIN lenses a good option for objective lenses as they can transmit an image and the focal plane can be adjusted based on the length of the fibre [32]. However, GRIN lenses which can transmit images are rigid, and so cannot be used as flexible probes.

Finally, applications for needle surgery and lumen organs have been greatly facilitated by the incorporation of prisms in the end of fibre-based imaging probes with the ability to perform side-viewing image acquisition [102, 103].

### 2.4.3 Image Processing for Fibre-based Endomicroscopy

In optical imaging and specifically in fibre-based approaches, image processing is a critical component in the development of these systems as it transforms the raw image data to useful imaging information while eliminating potential artefacts arising from the acquisition process. Due to the dynamic nature of *in vivo* microscopy, real-time computation is an important requirement.

As mentioned before, the image resolution is reduced in proximal-scanning, fibre-based, microscopy approaches by the non-uniform honeycomb pattern introduced by the fibre spacing. Image reconstruction algorithms are applied to the raw microscopic images to restore the true physical signal from the raw data by removing the honeycomb artefact, either by linear triangular interpolation [104, 105] or simple low-pass filtering [106, 107]. A great amount of work in this area has been done in fibre-optic images (fibrescopes), with algorithms using spectral filtering in the frequency domain [108], adaptive spectral masking [109] or spatial interpolation [110] but the speed of these approaches is not adequate for real-time performance. Also, in fluorescence microscopy, an additional level of processing was introduced by Le Goualher *et al.* [104], before the intensity interpolation, where each raw data value was processed through a fluorescence imaging model to recover the true signal coming from the tissue and take into account the intrinsic characteristics of the fibre.
Although high resolution endomicroscopy provides cellular level resolution through the use of non-unity magnification optics at the distal tip of the probe, the interpretation of single microscopic images has been shown in studies to be inadequate for histology-like assessment and tissue interpretation [111]. High-resolution fibre bundle endomicroscopes (e.g. Cellvizio UHD probes) are limited in FoV to the order of 240 µm, making surveillance of large areas of tissue difficult. This lead to the development of image registration techniques, additional to the reconstruction algorithms, in order to stitch together individual microscopic images in a planar fashion and create a larger image that resembles histological images. The majority of endomicroscopy image mosaicking algorithms are developed by Vercauteren et al. [112, 105, 113] as part of the commercial system Cellvizio. From these, only the last [113] could provide real-time mosaicking for the 12 fps Cellvizio system as it was based on a normalised cross correlation approach and the visualisation on a simple dead leaves model (the frame is overlaid on top of the previous frames in the position dictated by the registration algorithm). The other two approaches [112, 105] developed were based on Riemman statistics and non-rigid transformations and can be used as post-processing methods to improve the real-time mosaicking results as they can account for tissue stretching. An alternative approach by Bedard et al. [107] used the cross correlation in the Fourier domain to determine the image shifts.

The challenges with image mosaicking are in dealing with cumulative image registration errors and loss of image information. The first challenge is faced in the work of Loewke et al. [114], where, after a real-time pairwise registration step, a second step is implemented using a global image-alignment algorithm from probabilistic robotics. In the work of Mahe et al. [115] cumulative errors from pairwise registration were eliminated by considering a priori information and in particular the noisy trajectory of a robotic scanner. More recently, Chang et al. [116] improved global alignment by introducing an hierarchical approach which considers sub-mosaics from short video sequences and a group-wise estimation is performed using SIFT matching. A step further from a simple topology mapping towards a more dynamic mosaicking was presented in the work of Mahe et al. [117] where time information was taken into consideration and motion appearance is maintained, an important fact when imaging \textit{in vivo} blood vessels. Finally, the loss of imaging information was the main focus of the work of Vyas et al. [118] that
coupled an EM tracker and a fibre-based endomicroscope and fused both inputs (normalised cross correlation image shifts and EM data) to compensate for mosaicking errors during handheld or endoscopic operation.

2.4.4 Commercial Endomicroscopy Systems

Proximal and distal scanning approaches are both represented also by two clinical, commercially available, endomicroscopy systems. The first approach is the PENTAX ISC-1000 system (Pentax/Hoya, Tokyo, Japan and Optiscan Pty Ltd, Notting Hill, Victoria, Australia collaborated in this product) which refers to a distally-scanning single optical fibre incorporated into a conventional endoscope (see Figure 2.8). The main characteristics of the PENTAX confocal endomicroscope system can be found in the work of Delaney et al. [119] whereas the combined system’s features are summarised by Polglase et al. [120]. In more detail, an electromagnetically-driven scanning mechanism is placed in the distal end of a single optical fibre and it performs a raster scan across the optical window of the probe. Additional to the 2D optical sectioning scan, the system incorporates a piezoelectric z-actuator that provides depth scanning from 0 to 250 $\mu$m beyond the window surface as the probe is in contact with the tissue. The confocal system is integrated in a conventional PENTAX endoscope of 12.8 mm diameter and it covers 5 mm of the overall diameter. The rest of the volume is used for two light guides, a biopsy channel, an imaging
CCD and air and water-jet channels for injecting the contrast agents. The system’s imaging capabilities provide 1024x1024 pixel images at 0.7 seconds per frame or 1024x512 images at 1.2 seconds per frame. The system has been used for imaging mainly the lower gastrointestinal tract, but also for imaging the distal oesophagus, duodenum, stomach and the cervix [121, 122, 123, 124].

The second design approach is the Cellvizio system (Mauna Kea Technologies, Paris, France) (see Figure 2.8). This is a standalone probe-based confocal endomicroscopy (pCLE) system which can be inserted into any conventional endoscope’s working channel [125]. The system integrates a proximal scanning mechanism of high frequency mirror systems that scan across a very high number of optical fibres (up to 30,000) to deliver 488 or 660 nm wavelength laser light to the tissue area under examination. The system also provides dedicated imaging software that enhances the obtained images and provides additional capabilities such as mosaicking of multiple microscopic images. The image acquisition frame rate is 12 frames per second, providing close to video-rate results for enhanced diagnostics. The frame rate is achieved by two mirror systems, one oscillating at 4kHz for horizontal line scanning, and the second by a galvanometric mirror for frame scanning [126]. The company has developed multiple confocal mini-probes to be used for different applications according to their requirements. These probes have different bundle characteristics, such as number of fibres, flexibility, size and resolution, and different distal optics. The selection of different mini-probes provides different working distances, depending on the application, while the focal plane always remains fixed. Recently (March 2016), Mauna Kea Technologies were granted FDA 510(k) clearance for the use of endomicroscopy in surgical procedures.

Additional to the aforementioned products, a handheld fluorescence endomicroscopy system is available from Optiscan (Optiscan FIVE1, Optiscan Pty Ltd, Notting Hill, Victoria, Australia) which was initially developed for pre-clinical animal research but has been used for human tissue as well [127, 128](see Figure 2.8). The handheld instrument has a 6.3 mm diameter rigid shaft and is able to image a 475x475 μm FoV with resolution of 7 μm axially and 0.7 μm laterally. The scanning module is based, similar to the PENTAX endoscope, on a electrically operated tuning fork mechanism [129] that scans a single fibre.
to provide high resolution images at 0.8 or 1.7 fps depending on the image size, while there is a Z-axis actuation micro-motor to change the imaging depth in 4 µm steps up to 250 µm, providing the ability to image different depths without changing probe. The XY motion of the fibre is accomplished when the first drive inside the handheld instrument vibrates at its resonance frequency and the second drive, orthogonal to the first direction, vibrates non-resonantly.

Aside from fluorescence endomicroscopy, optical coherence tomography has been for more than 20 years in the clinical market and used regularly, especially for ophthalmology and cardiology. However in terms of fibre-based, endomicroscopic systems, only one system exists for GI applications (Volumetric Laser Endomicroscopy (VLE) [130]) and it is the product of the company NinePoint Medical, called NVisionVLE (see Figure 2.8). The system consists of a side-firing, fibre-based OCT probe, integrated with a low-pressure balloon catheter, which can be inserted through a channel of an endoscope; when the balloon is stabilised in the suspicious areas of the oesophagus, the flexible probe is stabilised by inflating the balloon and then it starts scanning helically, using a proximally positioned rotating engine, for the whole extent of the balloon (around 6 mm) and provides multiple transverse (up to 3 mm depth) and longitudinal OCT images of the oesophagus. All images are then collected to generate a 3-dimensional OCT dataset that improves the yield of diagnosis [131].

To conclude, all these optical imaging approaches and devices have to follow specific technical requirements in order to be compatible with surgery (exploratory and interventional); these requirements were described in details by Osdoit et al. [132]. In particular:

1. *In vivo* imaging deals with tissue motion that requires increased frame rates (larger than 10 fps) to eliminate motion artefacts; proximal techniques provide higher frame rates than distal scanning mechanisms due to reduced constraints in terms of dimensions. In distal scanning mechanisms the frame rate depends highly on the FoV and accuracy of scanning, leading to frame rates limited to less than 10 fps. This disadvantage is more evident when mosaicking of microscopic images is needed as lower frame rates necessitate slower speeds to achieve the same image
frame overlaps, requiring a degree of control which may not be feasible in flexible endomicroscopy.

2. For most of the endoscopes and colonoscopes, the imaging probes should be less than 3 mm in order to be inserted in their channels, leading to miniaturisation issues; not all the aforementioned technologies can be miniaturised and as a result a limited amount of clinical applications can be investigated.

3. It is desirable not to have any active / electrical components and wires at the tip of the imaging probe as, particularly in the case of high voltages, this is a potential risk for the patients and operators.

4. Complicated distal designs create mechanical robustness issues as any shock or vibration could affect the function of the probe, both short-term and long-term.

5. Disposable mechanisms are desired in clinical scenarios, resulting in issues with complicated distal scanning devices such as MEMS. If sterilisation is required, then the imaging probe should not be susceptible to temperature, chemical agents etc., resulting in advantages if it is a passive component.

2.5 Surgical Robots and Minimally Invasive Methods

All the aforementioned technology is mainly based on bench-top implementations or used in conjunction with conventional endoscopic instruments. However, the way that surgical operations are currently performed presents a shift towards more miniaturised and sophisticated instruments based on robotic control that can potentially assist the application of optical technology and improve its diagnostic yield. Therefore, the latter part of this chapter discusses the developments in surgical robotics, presents the evolution of these systems from passive assistants to active cognitive systems and, simultaneously, describes the robotic features that can assist the incorporation of optical technologies in future surgical robotics.
From the early implementations of medical robotics until the emergence of novel hand-held devices that perform cellular level motions, these instruments have the ultimate aim to augment the capabilities of the surgeons while exhibiting clear patient benefits and increased safety in the operating site. The introduction of the information and robotics technology has enhanced MIS procedures, extending the surgeon’s capabilities by including an intermediate level of processing between the surgeon and the patient, as shown in Figure 2.9. Computerised kinematic mapping of the operator’s manipulations enhances the instrument’s motions, while the precise kinematic control of the instrument’s end-effector is achieved by removing the unintentional tremor and scaling of the operator’s movement. Wrist articulation, \textit{i.e.} the fact that instruments can imitate the human wrist’s motion, is, also, introduced due to the enhanced flexibility of hyper-redundant, robotic instruments. From the imaging perspective, stereoscopic cameras immerse the operator inside the operating site while providing augmented reality capabilities by displaying pre-operative or intraoperative patient specific information.

![Diagram of Computer Assisted Robotic Surgery.](image)

Figure 2.9: General schematic of Computer Assisted Robotic Surgery. Adapted from [2]
2.5.1 Historical Developments

The first recorded robotic surgical procedure was a CT-guided brain biopsy that took place on 11 April 1985, at the Memorial Medical Centre, Long Beach, CA, USA. A lightly customised industrial robot, a Unimation PUMA 200 industrial robot, was used to place a needle for brain biopsy [133]. The rationale behind this work was to use machine precision to assist positioning the needle in such way that the latter can access deep brain areas without damaging vital structures. The trajectory to be followed by the needle was defined by the surgeon through CT guidance. This first attempt summarises the overall idea behind robotic surgery: “machine precision under human control with the assistance of imaging and sensing”. In parallel with the first neurosurgical applications, surgical robotic systems began to appear in orthopaedic surgery due to the rigidity of the bone structure which assists the registration of the robot to the anatomical scene. Specifically, “ROBODOC”, developed by Taylor and Kazanzides [134], could perform bone milling tasks autonomously based on pre-operative plans for positioning implants in total hip replacement (THR) surgery, while “Acrobot” (Active Constraint Robot), developed by Davies et al. [135], introduced the active constraints framework where the surgeon can manipulate freely the robot’s end-effector only inside a predefined allowable site. In contrast with the previous systems, “PROBOT” was used in 1988 by Sir John Wickham and Brian Davies at Imperial College London to demonstrate the feasibility of performing a robotic-assisted Trans-Urethral Resection of Prostate (TURP), becoming the first surgical robot to operate in soft-tissue [136].

2.5.2 Rigid Teleoperated Robotic Systems

Teleoperated systems appeared in the early ’90s and represent the most common surgical robotic systems currently in use. Such systems feature a master-slave architecture that physically separates the surgeon from the surgical field. Usually, the surgeon operates the master console to control the slave robotic arms, which perform the surgery at the operating table. This control paradigm allows the surgeon to perform the procedure remotely from the surgical site (distance is generally limited to 0.5-3 m). Teleoperated systems
were initially conceptualised for military and space applications as experienced surgeons could operate on injured soldiers far from the battle scene or on remote astronauts. Similarly, teleoperated systems can even be used to protect the surgeon from a hazardous environment, e.g. external imaging using radiation, by allowing a safe distance between the slave and the master.

Since the first introduction of industrial robots in the operative room, a wide range of teleoperated surgical robots have been developed and used in clinical applications. The first robotic device approved by the Food and Drug Administration (FDA) in 1994 was the Automated Endoscopic System for Optimal Positioning (AESOP) [137] developed by Computer Motion Inc.. Cognition was already apparent in this early robotic device as it provided hands-free control by voice commands. The AESOP robot was then included in the ZEUS platform [138], as an endoscope holder, additionally to the other two remotely operated robotic arms of the system. The ZEUS platform became also popular through “Operation Lindbergh” which was the first teleoperated transatlantic laparoscopic cholecystectomy performed between Strasbourg (France) and New York (U.S.A.) [139]. A similar attempt, towards exploring the versatility and connectivity of current teleoperated robotic systems, was performed in Plugfest 2009 where nine research teams from universities and research institutes operated from different locations around the world using the same transmission protocol [140].

The same concept of remote teleoperation was followed in parallel by the most successful, commercially available, master-slave robotic surgical system to-date; the Da Vinci® robotic system (Intuitive Surgical Inc., Sunnyvale CA., USA) [8] (see Figure 2.10). It was introduced in 1999 and it was based on the SRI Green Telepresence System developed by DARPA for military applications. Currently there are three generations of Da Vinci systems in the market (Da Vinci S, Si (HD), Xi) with an installed base of 3745 units as of July 2016 [141]. The benefits introduced by the robot are significant as it introduces a digital, rather than mechanical, intermediate level between the surgeon and patient. This level includes:

- a master console (see Figure 2.10 (left)) that the surgeon sits at and manipulates the
arms of the robot through a 3D visualisation environment, displaying a real-time, magnified view of the operating site through a stereoscopic endoscope, and

- a slave robotic system (see Figure 2.10 (centre)) incorporating four robotic arms that position the EndoWrist™ instruments allowing 7-DoF natural motion inside the patient with enhanced dexterity and wrist articulation.

All these features contribute to the advantages of the Da Vinci system over conventional laparoscopic MIS as surgeons operate in an immersive 3D environment, with a direct view inside the patient, while manipulating instruments that emulate their hand motions and are further enhanced by motion scaling and tremor removal algorithms. The EndoWrist™ instruments, initially based on the seminal work of Madhani et al. [142], has been improved over the next generations, adding features such as increased articulation or energy delivery capabilities. The clinical uptake of the da Vinci surgical system is significant as it has been used in a wide range of procedures (over 3 million procedures as of 2016 [141]), such as coronary bypass and hernia repair [143, 144]; however, its primary use remains limited to urologic surgery and in particular to radical prostatectomy [145] and gynaecological procedures [146].

Despite the advantages of the Da Vinci platform, the adoption of the system is hampered
by its high cost, including significantly high maintenance costs and additional training costs. A lack of versatility is also prominent due to the elongated rigid shaft surgical instruments, which restrict the port placement and reduces the reachable workspace. This is mostly noticeable in paediatric robotic surgery where the surgeon has to cope with a limited workspace [147]. In the literature, attempts with the Da Vinci robot to increase the workspace by incorporating a single incision laparoscopic device showed promising results [148], while introducing further problems such as limited triangulation and instrument clashing [149].

Additionally, the lack of haptic feedback is a key limitation of the current da Vinci system when compared with conventional laparoscopy, although the company has already filed several patents in this field [150]. This capability also features in numerous research publications and some newer surgical systems, mainly in the form of strain gauges and optical sensors. The multi-port TransEnterix ALF-X system (originally developed by SOFAR SpA, Italy) is one of the upcoming commercial systems that includes force sensing capabilities using sensors at the proximal base of the rigid instruments, but the latter do not provide any wrist articulation [151]. Another system that introduced the concept of robotic assisted microsurgery is the RAMS system developed by the NASA Jet Propulsion Laboratory (JPL) [152]. The system consists of two 6-DoF robotic arms in the slave part capable of 10 µm positioning accuracy, equipped with 6-DoF tip-force sensing at the tip, capable of scaling the motion of the hands in a scale of 1 : 100 and achieving the required frequency of tremor filtration (8 – 14 kHz) for eye surgery. For neurosurgical operations, Sutherland et al. developed the NeuroArm (University of Calgary, Canada), a remotely-operated, MR-compatible surgical robot [153]. The slave part of the system includes a SCARA-based 6-DoF robotic arm with integrated 3-DoF optical force sensors in the end-effector of the tool to provide haptic feedback input to the surgeon. Using motion scaling and advanced piezoelectric ceramic actuators, NeuroArm’s instruments can be manipulated with 3 µm accuracy through two PHANTOM Omni haptic devices with a custom stylus integrated in the end-effector of the haptic devices. The surgeons are further assisted by intraoperative registration algorithms that register the pre-operative patient specific data to the intraoperative scene and the robot’s position. A detailed review on force sensing medical robotic systems can be found in the work of Trejos et al. [154].
A notable robotic system that has not been commercialised as yet, is the DLR MiroSurge, developed by the German Aerospace Centre, consisting of three lightweight robotic arms, called MIRO, guided by either classic position control methods or force-torque impedance control. Two of the arms are equipped with custom instruments, providing enhanced articulation and force-torque sensing, while the third arm acts as a camera positioner with stereoscopic capabilities. The developers have also implemented a variety of software enhancements such as semi-autonomous control algorithms to allow hands-on interaction with the user and adaptive motion tracking [155].

Finally, a different approach towards teleoperated surgical systems was followed by the Biorobotics Laboratory at University of Washington DC with the RAVEN system [156], now commercialised by Applied Dexterity (Seattle, USA). The system is tendon-based with lightweight instrument arms, mounted directly on the operating table, and includes 7-DoF wristed instruments. Notably, the developers provide the system as an open platform to other research institutes worldwide and various modifications of the previous system are developed, such as a multi-bending flexible imaging device that has been incorporated in the design to enhance the capabilities of vision [157]. The latter flexible robot enhanced the capabilities of the surgical robot by improving its dexterity and versatility, as it will be described in the next section.

2.5.3 Highly Articulated and Flexible Robotic Systems

Recently, further research has been focused on minimising to a greater extent the invasiveness of current MIS procedures compared to rigid-based robotic systems. Therefore, recent developments consist of flexible-based, robotically-assisted surgical devices that can navigate through more complex anatomic structures and perform single incision laparoscopic surgery (SILS) or natural orifice procedures such as transluminal (NOTES) or endoluminal. The technical challenges arising from these operations are related to the need for increased flexibility of instruments while maintaining stable operation and the desire, from the surgeon’s side, for instrument triangulation. In particular, surgeons during this type of procedures access the operating site through curved, narrow, anatomical pathways with sensitive, surrounding organs and should be able to have a clear and broad
<table>
<thead>
<tr>
<th>System</th>
<th>Company</th>
<th>Use</th>
<th>Procedure</th>
<th>Cost ($K)</th>
<th>Approval</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALF-X</td>
<td>Sofar SpA, Italy</td>
<td>Teleopated Multi-port system</td>
<td>Laparoscopy</td>
<td>1400</td>
<td>CE</td>
</tr>
<tr>
<td>Amigo</td>
<td>Catheter Robotics Inc., USA</td>
<td>Catheter robot</td>
<td>Endovascular</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>ARTAS</td>
<td>KB Medical, Switzerland</td>
<td>Stereotactic robot / positioner</td>
<td>Neurosurgery</td>
<td>-</td>
<td>CE</td>
</tr>
<tr>
<td>ARTAS</td>
<td>Restoration robotics, USA</td>
<td>Hair restoration</td>
<td>Hair transplant</td>
<td>200</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Artis Zeego</td>
<td>Siemens, Germany</td>
<td>CT scanner</td>
<td>Imaging</td>
<td>2000</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>CorPath200</td>
<td>Corindus, USA</td>
<td>Catheter robot</td>
<td>Endovascular</td>
<td>300</td>
<td>FDA</td>
</tr>
<tr>
<td>Cyberknife</td>
<td>Accuray, Inc., USA</td>
<td>Radiosurgery</td>
<td>Neurosurgery</td>
<td>4200-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>DaVinci</td>
<td>Intuitive Surgical Inc, USA</td>
<td>Teleopated Multi-port system</td>
<td>Laparoscopy</td>
<td>600–2500</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Dex Robot</td>
<td>Dextérité Surgical, France</td>
<td>Handheld MIS tool</td>
<td>Laparoscopy</td>
<td>-</td>
<td>CE</td>
</tr>
<tr>
<td>Epoch</td>
<td>ERA endoscopy S.R.L, Italy</td>
<td>Endoscope holder</td>
<td>Laparoscopy</td>
<td>-</td>
<td>CE</td>
</tr>
<tr>
<td>Epoch</td>
<td>Stereotaxis Inc., USA</td>
<td>Robotic navigation system</td>
<td>Endovascular</td>
<td>1200</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Flex</td>
<td>Medrobotics, USA</td>
<td>Single-port</td>
<td>TORS</td>
<td>700</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Freehand</td>
<td>OR Productivity PLC, UK</td>
<td>Endoscope holder</td>
<td>Laparoscopy</td>
<td>-</td>
<td>CE / 510(k)</td>
</tr>
<tr>
<td>Gamma Knife</td>
<td>Elekta , Sweden</td>
<td>Radiosurgery</td>
<td>Neurosurgery</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Invendoscope</td>
<td>Invendo Medical, Germany</td>
<td>Endoscope holder</td>
<td>Laparoscopy</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>iSYS 1</td>
<td>iSYS Medizintechnik GmbH, Austria</td>
<td>Placement and biopsy robot</td>
<td>Neurosurgery</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Makoplasty</td>
<td>Stryker Inc., USA</td>
<td>Milling robot for TKA / THA</td>
<td>Orthopaedic</td>
<td>-1000</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Maxio / Robio</td>
<td>Perfint Healthcare Pvt. Ltd., India</td>
<td>Needle placement</td>
<td>Biopsy</td>
<td>1000</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Maxio / Robio</td>
<td>Bluebelt Technologies , USA</td>
<td>Handheld milling robot</td>
<td>Orthopaedic</td>
<td>400-450</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Neuromate</td>
<td>Renishaw, UK</td>
<td>Stereotactic robot</td>
<td>Neurosurgery</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Novalis Rad.</td>
<td>Brainlab AG , Germany</td>
<td>Radiosurgery</td>
<td>Neurosurgery</td>
<td>-</td>
<td>(Recalled)</td>
</tr>
<tr>
<td>Radius System</td>
<td>Tuebingen Scientific Medical GmbH, Germany</td>
<td>Handheld articulated robotic tool</td>
<td>Laparoscopy</td>
<td>-</td>
<td>CE</td>
</tr>
<tr>
<td>RAVEN</td>
<td>University of Washington, USA</td>
<td>Teleopated Multi-port system</td>
<td>Laparoscopy</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Renaissance</td>
<td>Mazor Robotics, Israel</td>
<td>Automatic placement of implants</td>
<td>Spine surgery</td>
<td>700</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Robodoc</td>
<td>Think Surgical, USA</td>
<td>Milling robot</td>
<td>Orthopaedic</td>
<td>600</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Rosa</td>
<td>Medtech, France</td>
<td>Stereotactic robot</td>
<td>Neurosurgery</td>
<td>330</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Soloassist</td>
<td>AKTORmed GmbH, Germany</td>
<td>Endoscope holder</td>
<td>Laparoscopy</td>
<td>60</td>
<td>CE</td>
</tr>
<tr>
<td>Sensei X</td>
<td>Hansen Medical CA, USA</td>
<td>Catheter robot</td>
<td>Catheter</td>
<td>600</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>SYMBIS</td>
<td>Deerfield Imaging Inc., CND</td>
<td>MRI compatible robot</td>
<td>Neurosurgery</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
<tr>
<td>Viky</td>
<td>EndoControl, Inc, France</td>
<td>Endoscope holder</td>
<td>Laparoscopy</td>
<td>-</td>
<td>CE / FDA</td>
</tr>
</tbody>
</table>

Table 2.2: List of commercially available surgical robotic systems (as of September 2016)

68
view of the operating site, without compromising the safety of the patient and without losing the increased dexterity provided in other MIS procedures. Up to now, similar flexibility for endoluminal operations was provided only by traditional endoscopes, which do not include any robotic actuation and are manually driven, directly by the surgeon. For this reason, researchers steered towards the development of either single-port, highly articulated or flexible, endoscope-based robotic platforms with robotically actuated segments and increased flexibility. In this section, a brief overview of the most important systems will be provided; for more information the reader is advised to read the detailed review publication by Vitiello et al. [158].

Despite the fact that flexible robots for general purposes have existed for more than 50 years, medical flexible robots were introduced with the publication of Ikuta et al. [159] that used smart memory alloys (SMA) to actuate a five-segment endoscope. After this publication, other works followed similar snake-like approaches; Choset et al. [160] developed the Highly Articulated Robotic Probe (HARP) (later called CardioARM) with the aim initially to navigate to the entire intra-pericardial space through only a single (sub-
xiphoid) port [161]. Recent additions to the platform included a shape-locking over-tube mechanism and two lateral flexible ports for instrument triangulation [162]. This is the base of the commercial edition of the robot, called Flex® robot (distributed by Medrobotics Corporation) which is currently FDA-approved for Transoral Robotic Surgery (TORS) [163].

The concept of snake-like robots was followed by Yang et al. [164] that proposed a bio-inspired 7-DoF articulated robotic device, called i-Snake®. Its joint design is based on universal joints and they are actuated through embedded micro-motors accompanied by a tendon-driven mechanism. In its body, two internal channels are incorporated, allowing the deployment of a standard endoscopic camera for visualisation and various endoscopic interventional instruments. Furthermore, the same robotic device demonstrated enhanced imaging capabilities as the authors integrated a multispectral imaging system with force adaptive control [165] and a pCLE probe in order to accomplish in situ cellular level imaging for real-time tissue characterisation and positive cancer margins identification. The key limitation of the aforementioned platform is the limited tissue manipulation capabilities due to the deployed instruments lack of triangulation. To overcome this, Yang et al. [166] developed another version of the i-Snake robot, which incorporates an actuation pack at the proximal end, with the i-Snake joint units for a 3-DoF articulated “head” and two flexible robotic arms.

The concept of flexible arms in combination with a rigid shaft for enhanced visualisation is adopted by Simaan et al. [167, 168] in the platform called Insertable Robotic Effectors Platform (IREP), recently licensed to Titan Medical Inc. as SPORT™ system. The platform provides two snake-like continuum arms of 6-DoF each, with push-pull actuation, an 1-DoF gripper, along with a 3-DoF (pan, tilt and zoom) stereo vision camera, while two 2-DoF five-bar mechanisms allow the arms to fold-unfold. Experiments with the platform were performed only in a laboratory setting, presenting however promising results for delicate pick and place tasks and knot tying.

A notable concept of a highly articulated single-port robot has been developed by Intuitive Surgical, called Da Vinci SP. It has been already FDA-approved for clinical use and it is likely to be marketed in 2016. There are not currently many technical details available, but
it consists of three 8-DoF highly articulated, snake-like, 6 mm instruments (1 joggle joint for triangulation and 7-DoF instruments) along with a flexible stereoscopic camera, all mounted in a cart on the patient-side [169]. The flexibility of the system is a step forward compared to the rigid Da Vinci Xi system and allows new operations to be performed robotically, more efficiently and safely [170, 171, 169].

Another novel bi-manually controlled robotic system has been developed by Dario et al. [172, 173] through the ARAKNES consortium called SPRINT (Single-Port laparoscopy bImaNual roboT). The system consists of two 6-DoF robotic arms of 18 mm diameter, that can be individually inserted, and a stereoscopic camera, which are introduced through a cylindrical tube of 30 mm diameter. Despite the large dimensions of the slave side, the surgeon’s console presents enhanced capabilities as it is equipped with a 3D full-HD monitor, incorporating augmented reality and eye-tracking capabilities, and two Omega 7 haptic devices. An alternative approach to highly articulated robotic systems is proposed by the ANUBISCOPE® (iCUBE STRAS) robot, developed at IRCAD-Strasbour, that is based on the Karl Storz® Anubis® system, a 18 mm diameter single-access endoscope with a flap-based distal tip that opens to reveal two 3.6 mm triangulating flexible instruments. The extremely complex surgical control of the manual system and the ergonomics of the system are greatly improved by motorising the control handles and adding teleoperation capability. Up to now, the system has been already clinically used for transvaginally and transgastrically NOTES cholecystectomy and also for endoscopic submucosal dissection (ESD) operations [174, 175].

Finally, a special category of flexible surgical robots are the endovascular, catheter-based robots such as the Sensei® X2 Robotic Navigation System and the Magellan® Robotic System, developed and commercially available by Hansen Medical Inc. [176, 177]. These differ from all the previous systems as the robot consists of a thin (< 10 Fr diameter) robotically controlled catheter, improving its dexterity and reachability, but they do not offer the option of multiple instruments or triangulation. These catheters are mainly for endovascular navigation and catheter-ablation for electrophysiological operations. The slave side of the systems comprises a single robotic arm, fixed on the patient table, for gross positioning and a 3D steerable catheter manipulator with one or two bending sec-
tions. The master console consists of a parallel manipulator allowing 6-DoF hand motion and a graphical user interface that displays pre-operative data, intraoperative angiographic images and the corresponding 3D position of the catheter, simplifying the navigation for the surgeon. In the Sensei system, surgeons perform the procedures at the master console assisted by a technology called IntelliSence® which is the Sensei’s built-in pressure sensor technology that measures proximally the forces applied on the catheter’s shaft due to tissue contacts, delivering visual force feedback to the surgeon.

All the aforementioned systems demonstrate improved versatility and reachability and are a great step forward towards the elimination of the surgical incisions and the patient trauma while keeping the safety features and the ergonomic manipulations of the rigid, master-slave, robotic systems.

### 2.5.4 Hands-on and Handheld Surgical Robotic Systems

An essential requirement of surgical robots is to closely collaborate with surgeons, enhancing their abilities and assisting them during the operation. All the aforementioned systems, despite fulfilling to some extent their role, still have a relatively large footprint in the operating theatre, which is controversial compared to the idea of miniaturisation in the era of digital mobile computing. In parallel to the previous developments, an increased interest has been exhibited towards mechatronically-enhanced, hands-on and handheld surgical instruments, that can collaborate with surgeons directly, providing a certain degree of integrated intelligence and autonomy. These instruments provide clear benefits over other large robotic systems by providing a more natural and intuitive operation, with no need of extensive training, easier integration with other computer-assisted surgical (CAS) systems, real-time adjustment of the interaction with the operating field in terms of forces applied and perturbations, and lower costs.

From the early tool developments, precise tissue manipulation, increased dexterity and enhanced sensing capabilities, such as haptic and tactile feedback, were the main goals of mechatronically enhanced instruments. In this category, we can observe two different trends in instrument developments; one that has instruments with a grounded frame of
Figure 2.12: Exemplar cooperatively-controlled and handheld instruments for orthopaedic surgery (a) Acrobot ©2006 IEEE, (b) Stryker Mako Surgical Arm ©Stryker Corp. and (c) Navio\textsuperscript{TM} PFS system ©BlueBelt Technologies, Inc.

reference, \textit{i.e.} attached to a robotic arm and the surgeon uses a cooperatively-controlled (hands-on) approach to manipulate them, and one that has instruments totally handheld without any grounded reference [178].

The first robot to appear in this category is the Acrobot (previously mentioned) as the surgeon could manipulate the robot by holding a handle while the robot’s motion is constrained in a pre-designed workspace. Providing also active guidance, the RIO\textsuperscript{®} MAKOplasty\textsuperscript{®} system (recently acquired by Stryker Corp.) is another exemplar system that was developed for reconstructive orthopaedic surgery and in particular partial knee and total hip replacement. The system includes a robotic arm that the surgeon holds freely and guides actively through the use of tactile resistance to remove the appropriate bone area based on pre-operative imaging [179]. An alternative approach in precision bone cutting is offered by the commercially available Navio\textsuperscript{TM} PFS system (BlueBelt Technologies, Inc). In contrast to the previous two robots, this is a handheld instrument, with marker-based optical tracking incorporated, that guides the surgeon by controlling the depth of the cutting tool to achieve precise bone cutting.

Furthermore, a surgical area that cooperatively-controlled and handheld instruments excel
at is microsurgery. Developments in tremor suppression and force feedback combined with the inherently safe and compact nature of these instruments have encouraged their use over classic teleoperated systems. The Steady-Hand Eye robot, developed at John Hopkins University by Taylor et al. [180], is a cooperatively-controlled robotic assistant intended to be used for retinal microsurgery. The system consists of a table-mounted 7-DoF robotic arm, manipulated in a hands-on fashion through a virtual fixtures control law with audio feedback; the arm, therefore, guides the surgeon’s hand motions from hitting delicate structures by introducing virtual walls, while eliminating the physiological tremor from the surgeon’s hand. The Steady Hand consists of XYZ linear stages for translation, a tilting mechanism with a mechanical remote-centre-of-motion (RCM) mechanism, a rotary stage for rolling and a tool adaptor incorporating a force sensor, which prohibits the operator from exerting large forces both on the target and the RCM area. Clinically, the Steady Hand robot has successfully performed retinal vein cannulation, a very demanding operation requiring peeling of the epiretinal membrane and insertion of a needle into the lumen of a retinal vein [181].

Developed also for vitreoretinal surgery, one of the most mechatronically-advanced, hand-held, tremor-suppression, robotic instrument is the Micron [3], developed by Riviere et al. The main aim of this “smart” instrument is to actively compensate for the surgeon’s tremor, while providing real-time guidance for microsurgical tasks. In the first version of the instrument it consisted of six inertial sensors (three accelerometers and three gyroscopes) that monitor the tip’s motion and accordingly provide feedback to the actuation system. The latter included a 3-DoF manipulator based on a Stuart platform actuated by piezoelectric actuators and could attenuate tremor movement. Tremor was suppressed by sophisticated algorithms in the frequency domain and removed from the intended motion components. The experiments accompanying this work showed 37% attenuation of the involuntary 3D motion of the hand tremor [182]. In a second iteration of the instrument, the inertial sensing was replaced by the ASAP (Apparatus for Sensing Accurate Positions) optical position tracker based on position sensing devices (PSD). This version incorporated a modified flexure-based parallel manipulator providing a 400 µm range.
Figure 2.13: Exemplar motion-compensation instruments for micro-surgery (a) Steady-Hand robot ©2010 IEEE, (b) Initial prototype of Micron @CMU [3], (c) ITrem - Latt et al. @2009 IEEE, (d) SMART instrument ©2012 OSA, with permissions from the author (e) Micron - MacLachlan et al. @2012 IEEE and (f) Micron - Yang et al. @2012 IEEE.

of motion with 1.8 $\mu m$ RMS position error and 46% reduction in the undesired tremor movements [183]. This device has also recently been enhanced with visual-servoing using the images from a microscope [184], which are then used to provide active guidance in terms of virtual fixtures to the surgeon [185]. The latest version of Micron implements a 6-DoF parallel manipulator based on miniature piezoelectric linear actuators in the tip of the instrument which further reduces the forces exerted during MIS eye surgery as it can incorporate a RCM capability and also stabilise the whole tool pose with easier kinematic reactions [186]. Recently, both the Steady Hand and the Micron system were tested together in the same procedure (membrane peeling) to evaluate their performance and the study demonstrated that both systems had similar performance in terms of tremor reduction and consistency [187]. Also, the Micron system was improved in the latest publication of Wells et al. [188] where a hybrid force / position control algorithm was implemented, allowing it to control forces of less than 5 mN while peeling a membrane.

A different approach was followed in Song et al. [189, 190] with the SMART surgical tool for tremor suppression which integrated a common path, swept source optical co-
herence tomography-based (CP SS-OCT) system. A piezoelectric motor is coupled with the fibre optic CP-OCT sensing, which acts as a high-speed distance sensor, measuring the distance of the needle tip from the sample surface. Clinical tests in a chicken embryo model demonstrated attenuation of hand-tremor of around $0\text{–}15$ Hz. However, this implementation is limited to one-DoF sensing and therefore the authors are currently investigating a 3-DoF solution. Similar performance was presented by Saxena et al. [191] who incorporated Ionic Polymer Metallic Composites (IPMCs) to create a lightweight tremor suppression instrument in 1-DoF.

Bridging the gap between manual rigid laparoscopic instruments and costly master-slave robotic systems, various handheld instruments have been developed to introduce wristed robotic articulation in the form of laparoscopic tools. Dario et al. [192] developed a cable-actuated multi-joint mechatronic handheld tool for arthroscopy that was equipped with position and force sensing, allowing real-time registration of pre- and intraoperat-
ive imaging navigation, contact sensing, and collision avoidance for predefined delicate anatomical structures, such as cartilage and ligaments [193]. Recently, a 0.9-mm diameter, intraocular, dexterous robotic instrument was developed by Iordachita et al. [194], demonstrating miniaturisation of even articulated instruments. In general laparoscopic surgery, Matsuhira et al. demonstrated that wristed articulation provides better reachability compared to normal laparoscopic tools in tasks such as suturing due to increased dexterity [195]. Also, researchers attempted to reduce the overall weight of the motors on the instruments by encapsulating part of the actuation at the surgical trocar [196]. Regarding the increased dexterity, other approaches introduced multiple DoFs at the end of the robotically actuated instruments, demonstrating improved ergonomics with simple user interfaces [197, 198].

A specific category of handheld instruments that could be of particular interest for optical imaging, and for confocal endomicroscopy in particular, which requires constant contact and consistent forces exerted on the soft tissue surface, is force controlled and force feedback (haptic feedback) enabled handheld instruments. In the work of Yuen et al. [199] dynamic force control is incorporated into a handheld instrument used for surgical manipulations on the beating heart. The instrument applies precise forces on the tissue by combining the force measurements acquired by a force sensor at the tip of the instrument and a 3D ultrasound probe outside of the patient. This instrument was later improved by
adding an active inertia cancellation mechanism that eliminates the reaction forces generated by the mechanism and provides increased force sensitivity together with the motion compensation [200]. The work of Latt et al. [201, 202] showed the capabilities of force controlled handheld instruments in endomicroscopy (further information will be provided in section 2.6).

Additionally, handheld devices have been developed to provide “super-human” senses and enhance the existing tool-tissue contacts, essential in many microsurgical operations. In particular, Yao et al. [203] with the MicroTactus instrument could provide vibrotactile and auditory feedback by sensing the tool-tissue interactions with an accelerometer while probing the surface of the cartilage to identify small cuts during arthroscopic interventions. A series of force amplifying instruments have been developed by Stetten et al. [204] that sense tool-tissue interactions and can provide a bi-directional magnified force to the operator. In parallel to the previous work, Payne et al. showcased various compact handheld force amplifying devices using strain gauges [205, 206], piezoresistive-based force sensors [207] and optical tracking [208] to improve human force perception (in the mN range) and safety during tool-tissue interactions.

2.6 Robotically-Assisted Endomicroscopy

As can be seen in the review of surgical robotics systems above, there have been significant advances in terms of safety, dexterity and human-robot interaction, but very few of them integrate any type of microscopic imaging. Most of the systems rely on preoperative imaging for guiding the surgeon and intraoperative macroscopic imaging from the endoscopic camera view. On the other side, microscopic optical imaging, as described in the first part of the chapter, is still performed through manually controlled endoscopes with reduced stability, accuracy and limited FoV. As a result, robotically-assisted endomicroscopy seems an ideal candidate to alleviate these issues and help translate the advantages of optical imaging into the operating theatre.

In the last few years, an increased number of studies have demonstrated the potential of
Figure 2.16: Exemplar robotic rigid approaches for endomicroscopy (a) Latt et al. ©2011 IEEE, (b) Latt et al. ©2012 IEEE, (c) Giataganas et al., (d) Rosa et al. ©2012 IEEE, (e) Rosa et al. ©2011 IEEE and (f) Rosa et al. ©2014 IEEE.
robotic actuation for intraoperative microscopic imaging using probe-based confocal laser endomicroscopy (pCLE). Initial research in robotically-assisted pCLE was performed by Latt et al. [201, 202] by developing handheld rigid instruments with force sensing capabilities. Since pCLE is a contact-based imaging modality, forces applied to the tissue affect both deformation and imaging quality [209], especially when manipulated in a handheld manner. Latt et al. used load cells to sense axially the forces applied to the tissue (distally in this work [201] and proximally in this work [202]) and in order to maintain consistent contact (at 100 μN) linear actuation was used proximally (voice coil [201] and DC-servomotor [202]). Both works showcased the effects of forces in image quality of individual images but not in terms of mosaicking and large area scanning. The latter was the subject of an early work of the author of this thesis [210, 211] where a cooperative robotic manipulator (KUKA LWR 4+) was used in combination with the force adaptive instrument from Latt et al. [202] to automatically perform large area, spiral mosaics and simultaneously perform a 3D surface reconstruction using the force measurements when the instrument was in contact with tissue. In this work two novel schemes were introduced; spiral mosaics for reduced tissue deformation and machine learning through the learning by demonstration approach to provide expert knowledge through demonstrations from surgeons.

In parallel, a series of robotically-assisted pCLE devices were presented by Erden, Rosa et al. as part of the PERSEE project and in collaboration with Mauna Kea Technologies. In particular, initially Rosa et al. [212] developed a 5.45 mm diameter, micro-positioning robotic device based on hydraulic actuation with three balloon catheters. The device includes an outer tube for tissue stabilisation and is attached to a 6 mm bendable laparoscopic tool. It does not feature any force sensing capabilities but only mechanical stabilisation. The authors could perform small area mosaics but actuation inconsistencies was a limit on the performance for large area mosaics. This was alleviated by the development of an under-actuated device based on a conic mechanism for the performance of spiral mosaics. The conic structure of the microscopy probe was used in conjunction with a linear cam mechanism to perform spiral trajectories. Erden et al. presented, initially, a large 5 : 1 scale prototype [213] and later demonstrated the final 5 mm device [214]. This approach uses a spiral path, in contrast to previous raster scanners, due to the fact that in
raster scans there are sudden changes in the directions of the motion resulting in increased tissue deformation in soft tissue scans, a phenomenon called “slip-stick” by the authors [215]. To improve further the performance of spirals, a visual servoing controller was implemented based on the microscopic images [216]. The controller showed promising results ex vivo using a large industrial manipulator and later improved the accuracy of the micro-scanners. As a result of these developments, the authors could perform large area mosaics of 3 mm² intraoperatively. On the other side, the device still features actuation problems, especially in the centre region, and stabilisation of the tissues is still performed using an outer tube (a mechanical stabiliser). Also, in terms of the endomicroscopy system, the authors used the commercial Cellvizio system by Mauna Kea Technologies which could perform confocal fluorescence imaging with 240 µm FoV microscopic images only at a rate of 12 fps, which is a limiting factor for robotic applications as we will see in the next chapters.
As presented in the previous sections, the introduction of robotics in surgery improved the dexterity of surgeons in difficult to reach areas inside the patient and this was also the case for the developed robotic assisted endomicroscopy devices. This was initially demonstrated in the work of Newton et al. [217] with the adaptation of the i-SNAKE platform (a 7-DoF robotic articulated endoscope) where the authors added a force sensitive linear servoing mechanism proximally to the endoscope, with a 6-DoF force sensor attached directly to the imaging probe measuring axial forces and a DC motor and a rack and pinion mechanism compensating for force changes. Despite the inaccuracies of force sensing due to the frictional effects with the working channel along the imaging probe, the authors could generate approximately 2 mm area mosaics with consistent image quality.

A different approach for tissue stabilisation and improved dexterity was pursued by Zuo et al. [4, 218, 219] where the authors used balloon-based stabilisation and a 2-DoF mechanism for better reachability. In particular, an intraoperative breast scanning device [4, 218] was developed with a balloon attachment in the tip of the instrument that can be inflated inside the breast cavity to stabilise the surrounding tissue. Inside the balloon, there is a 2-DoF scanning mechanism based on a miniature gear that is actuated proximally. To alleviate issues of irregular cavity surfaces, there is a passive, spring-based, axial mechanism to adapt the tip of the instrument. The instrument was further miniaturised in [219] to a 4 mm diameter device with a 2.2 mm working channel for probe interchangeability using a hybrid gear and linkage-based approach, in contrast to the purely gear-based approach in the previous instrument, and the authors claimed scanning of large areas of 25 – 35 mm².

An interesting approach towards increased dexterity is the handheld mechatronic tool by Payne et al. [220]. In this work, the authors demonstrated a highly articulated 1-DoF tool of 4.2 mm diameter, that can retroflex (bend backwards) and is used for diagnostic arthroscopy, a notable operation with confined spaces. A schott fibre-bundle is integrated in the tip allowing optical biopsies in very tight workspaces using either epi-fluorescence or confocal microscopy. Additionally, this tool has an integrated monocular camera and a light source in the tip.

Recently, there has been increased interest by surgeons in using microscopic imaging
probes in conjunction with the da Vinci robot for intraoperative tissue identification. Brahmdhatt et al. [221] and Lopez et al. [222] used pCLE for microsurgery and radical prostatectomy respectively, manipulating the probe directly with an Endowrist® needle driver, while in the work of Patsias et al. [223] a high resolution microendoscopy (HRME) optical probe was inserted through a 5FR introducer tool, which is a 5 mm articulated robotic tool, to perform transoral robotic surgery. All the latter approaches did not use any bespoke robotic tool and were assisted by the motion scaling and improved dexterity of the da Vinci robot. However, the manipulation of the probe with a needle driver is not stable for all operating sites and using one tool just for microscopy is not cost and time efficient as the tools have limited uses. Additionally, the da Vinci robot lacks haptic feedback and therefore large area coverage is hindered by contact forces and tissue deformation.

Finally, additionally to confocal and microendoscopic approaches, the use of other optical biopsy techniques with robotic surgery has been explored. In the very recent work of Gorpas et al. [224], FLIM imaging is used to augment the surgeon’s macroscopic FoV by superimposing real-time (30 fps) tissue bio-chemical information, using pseudo-color masks, while scanning with a motorised stage. With this approach, features of more than 200 µm could be distinguished, a fact that is helpful for initial delineation of tumour margins but cannot be compared to histology’s resolution. The authors are currently exploring the use of this technology with surgical robotic systems. Furthermore, tissue classification through a point-based Raman sensor was explored together with the SPRINT robot in the work of Ashok et al. [225]. The authors presented a proof-of-principle study by using a rather large probe with a disposable sleeve to assess different tissue samples, demonstrating 95% accuracy. Another promising piece of work with spectroscopic imaging is presented by Avila et al. [226] with the “Endodrone”. The “Endodrone” is an endoscope attachment that allows optical probes to be scanned across the colon wall to acquire complete spectroscopic data about a region of the colon for screening purposes.
2.7 Discussions and Conclusions

Great advances in optics, mechatronics and computer science have enabled the development of endomicroscopy technologies over the last few decades, opening a new field of \textit{in vivo} microscopic disease detection. In this chapter, the evolution of real-time cancer detection imaging modalities is discussed. In addition to the principles of the imaging modalities, an overview of the scanning methods used to create images through optical fibres was presented. The FoV of these scanning methods, however, is usually less than a millimetre, presenting hurdles for large area coverage and for the creation of histology-like images, i.e. meaningful images of sufficient quality and size for the clinicians to interpret. Additionally, the flexibility of these probes, necessary to allow access to sites
within the patient, creates issues when manipulated either manually by hand or by manual endoscopes.

Robotic manipulation could potentially enhance the usability of these optical probes and improve their diagnostic yield. For this reason, in the second part of this chapter, an overview of robotic surgery systems has been provided, with a focus on different features that can potentially be used in conjunction with microscopic imaging. Master-slave robots provide motion scaling and improved dexterity to reach complex anatomical structures, while handheld robotic systems offer a compact, low cost alternative, integrated with tremor suppression and force control. The common denominator of every robotic system was the improved stability and manipulation which is essential for performing large area coverage with microscopic imaging. In the next chapter, endomicroscopic image enhancement techniques will be introduced with the aim of optimally producing histology-like images intraoperatively for large area surveillance.
Chapter 3

Microscopy Image Enhancement, Mosaicking and Multi-modality Image Fusion

3.1 Introduction

As discussed in the previous chapters, in vivo optical microscopy typically involves the use of fibre bundles to deliver high resolution microscopic images for optical biopsy. Light is relayed through compact, flexible fibre bundles to and from the tissue, while still allowing high resolution video-rate microscopy imaging [227]. In proximal scanning fibre-based approaches, technical challenges are associated with the acquisition of images through multiple fibres and the extraction of useful histological information (see Figure 3.1). As the light is scanned proximally on the surface of the fibre bundle, a

1Content from this chapter was published as:
honeycomb-like pattern artefact is present, affecting the lateral image resolution. The scanning amplitude and signal sampling frequency of the proximal mirror have been adjusted to spatially oversample the proximal fibre surface and, as a result, the honeycomb fibre pattern appears. Also, in implementations that use leached fibre bundles, such as in endocytoscopy [227], there is a possibility of broken fibres that result in dark or white spots in the image, introducing structural artefacts.

In addition, the sub-mm FoV of endomicroscopy probes may make it difficult for large area surveillance, and thus require the development of mosaicking algorithms [114]. Mosaicking is an image processing technique for stitching together adjacent image frames as the probe is moved across the tissue. In this way, larger microscopic images can be generated. In this chapter, algorithms for image processing and mosaicking will be discussed for both monochrome and coloured microscopic images with a particular focus on real-time processing of high frame rate endomicroscopy systems.

The aim of this chapter is to lay the foundation of endomicroscopy image creation, from the initial acquisition of raw fibre-based images to the generation of high resolution reconstructed microscopy images and then using them to create high resolution histology-like mosaics that can be fused with macroscopic views of the surgical scene. All this knowledge will offer the surgeon much more comprehensive information and greatly improve the prospects for intraoperative margin evaluation.

The key features of the developed image processing framework include the following:

- The generated images provide histology-like characteristics over a wide area of tissue for improved diagnostics.

- Mosaic images provide intraoperative, *in vivo* image guidance with cellular level resolution.

- The developed image reconstruction algorithm improves the visualisation of each individual image by eliminating artefacts due to the honeycomb-pattern and broken / damaged fibres.
Figure 3.1: Challenges associated with fibre bundle based technology that necessitate the existence of an image reconstruction algorithm; (a) honeycomb-like pattern artefact, (b) non-uniform circular image and (c) damaged or broken fibres.

- The developed mosaicking algorithm has superior speed performance, making it suitable for use with high frame rate microscopy systems (up to 120 fps).

- It is compatible with various types of endomicroscopy systems such as confocal, widefield, dual-wavelength and endocytoscopes.

- Preliminary work on intraoperative fusion of macroscopic views with microscopic images demonstrates the potential for improved intraoperative diagnostics and guidance.

3.2 Image Reconstruction

3.2.1 Materials and Methods

In this section, the honeycomb artefact is tackled through a local interpolation technique between the intensities of neighbouring fibre cores. The algorithm follows a similar path to the one described by Le Goualher et al. for confocal fluorescence endomicroscopy [104]. This approach can offer real-time performance to assist the surgeon with enhanced
high-resolution images during an optical biopsy process and can be applied to other fibre-based optical imaging techniques such as endocytoscopy. The aim of this process is to eliminate the intrinsic autofluorescence of the fibre and restore the true biological sample fluorescence, given the raw data transmitted by each fibre.

The overall steps of the proposed image reconstruction – enhancement algorithm include:

1. an offline step with an acquisition of a bright-field image to identify the individual fibre cores and of a dark-field image that represents the background signal;

2. an offline step to create a calibration matrix for each pixel of the microscopic image that includes forming a Delaunay triangulation over the fibre cores and expressing each pixel in barycentric coordinates with respect to the enclosing triangle;

3. an online step that subtracts the background signal from every image obtained;

4. an online step that filters the overall image with a Gaussian filter to reduce noise (optional);

5. an online step that interpolates each pixel’s intensity value based on local neighbour fibre core intensities calculated during the offline steps and

6. an online step that corrects for intensity variations across the fibre.

Representative figures of raw images with the honeycomb pattern artefact are depicted in Figure 3.2.

During the initial offline calibration stage, an image $I_b$, representing the background signal (e.g. fibre bundle artefacts, autofluorescence), is obtained. For confocal microscopy, the image should correspond to a non-fluorescent target (e.g. acquiring an image as the fibre is in the air), while for widefield microscopy and endocytoscopy, the fibre should be in a dark environment and away from fluorescent targets (to eliminate the out-of-focus light). In parallel, a bright calibration image $I_r$ with the core pattern visible is acquired. In confocal and widefield microscopy, the background image $I_b$ can be used instead of the bright images providing it clearly shows the core pattern. Alternatively, the average
of a sequence of images can be used as the bright calibration image. For reflectance microscopy, a strongly reflecting, uniform target should be used. If this is a colour image, then the plane with the highest intensity is selected. The aim of this stage is to identify offline, prior to the reconstruction operation, each of the individual cores of the fibre bundle and map their position.

To identify the boundary of the fibre bundle, a convex hull algorithm is applied, which
Figure 3.4: Core pattern removal process for an endocytoscopic image; the same process if followed with any fibre-based imaging approach.

fits a circle to the calibration image after the latter has been transformed into a binary image with an automatic threshold approach based on Otsu’s method [228]. With a regional maxima algorithm, implemented in Matlab® (Mathworks, Massachusetts), the core centres can be identified as they present the highest pixel intensities among their neighbour pixels (see Figure 3.4). Cores with saturated or low intensity values are assumed broken and removed from the record. The core position information is used as an input to a Delaunay triangulation mesh algorithm that creates triangles between neighbour points and ensures that no interior points exist in each of the triangles.

For each pixel in the image, the algorithm identifies the enclosing triangle, converts the pixel’s location to triangular barycentric co-ordinates (a measure of its distance from each vertex) and stores this information in a look-up table. Barycentric coordinates are useful to interpolate linearly between the intensity values at the vertices of the triangle. In particular, barycentric coordinates define the weights that each pixel inside the triangle should be assigned depending on the triangular areas calculated between the pixel and
Figure 3.5: Triangle definition of barycentric coordinates (left) and interpolated result of a monochrome image for one exemplar triangle (right).

the vertices of the triangle, as shown in Figure 3.5. Therefore, to reconstruct subsequent images, each pixel is assigned an intensity value, $I_p$, by linear interpolation between the intensities of the cores lying on the three vertices of the enclosing triangle.

$$I_p = b_1 I_{c,1} + b_2 I_{c,2} + b_3 I_{c,3}$$ (3.1)

where $I_{c,i}$ is the intensity values of the core at vertex $i$, and $b_i$ is the corresponding pre-calculated barycentric co-ordinate for that reconstruction pixel.

Specifically, each $b_i$ is calculated from:

$$b_1 = \frac{A_{CP_2P_3}}{A_{P_1P_2P_3}}$$

$$b_2 = \frac{A_{CP_1P_3}}{A_{P_1P_2P_3}}$$

$$b_3 = \frac{A_{CP_1P_2}}{A_{P_1P_2P_3}}$$ (3.2)

where $A_i$ are the triangular area values of the corresponding points. The weight values for each pixel are then stored in a calibration matrix with the corresponding triangle points and all of them are used in the online steps for each sequential image that is acquired. In colour images, the reconstruction is performed separately for each of the 3 colour planes.

To account for variations in illumination intensity and coupling efficiency of the different fibres, an intensity correction method is also implemented. This is performed by first reconstructing the image used for the initial calibration $I_r$. As this image should be of uniform intensity, it is possible to calculate a 2D intensity correction map $I_{intensity}$ which, when multiplied by the reconstructed image, renders it uniform (see Figure 3.6). Each
pixel \((i, j)\) of the 2D intensity correction map is calculated as:

\[ I_{\text{intensity}}(i, j) = \frac{\bar{I}_s}{I_s(i, j)} \]  \hspace{1cm} (3.3)

where \(I_s\) is the image of uniform intensity that is calculated from the initial calibration image and \(\bar{I}_s\) is the mean intensity value of the constant fluorescence image. Before calculating the 2D intensity correction map, a background subtraction is performed on the initial calibration image \((I_s = I_r - I_b)\) as well as all the subsequent images \((I_c = I - I_b)\) in order to eliminate any background signals generated by the fibre bundle (see Figure 3.3). If the background image is used as the bright calibration image, then former subtraction is not made, \(i.e. I_s = I_b\).

Finally, the true intensity is calculated by multiplying the captured image intensity, \(I\), with the calculated 2D intensity correction map:

\[ I_{\text{true}} = I \cdot I_{\text{intensity}} \] \hspace{1cm} (3.4)

Figure 3.6: Intensity correction process; all core pattern-free images are multiplied by a 2D intensity map to reconstruct a uniformly illuminated microscopic image. The top images correspond to the reconstruction of the calibration image that when is multiplied by the intensity map, it renders it uniform.
3.2.2 Experimental Results

During online image acquisition, based on the calibration matrix that is represented as a look-up table, all the pixels are assigned a value based on a weighted sum of the intensities of the three cores at the vertices of the enclosing triangle. This allows the image reconstruction approach to be used online. The additional delay is up to 40 ms for monochrome images and up to 160 ms for colour images, which is a permissible value for online processing for systems with a frame rate of 12 fps. This is also strongly dependent on the size of the microscopic images.

For high frame rate systems of 30 fps and more, the aforementioned approach is used for post-processing of the acquired video sequences and a simple Gaussian filter over the whole image is used for online visualisation and real-time mosaicking. The Gaussian filter cannot compensate for broken or damaged fibres but it is still useful for online visualisation at high frame rates.

As can be seen from the colour images in Figure 3.7, the results correct almost all image artefacts and the diagnostic yield of the endomicroscopic images could be expected to improve significantly. Additional exemplar results of the developed image reconstruction
algorithm are presented in Figure 3.8.

### 3.3 Building Large Area Mosaics from Microscopic Images

Another limitation arising from the use of fibre bundles is the trade-off between lateral resolution and the FoV. Fibre bundles contain a finite number of cores (from 10,000 to 100,000), that essentially act as pixels in the image. Whilst high-resolution endomicroscopy is possible through the use of non-unity magnification optics at the distal tip of the fibre bundle, the result is always a proportional reduction in the FoV. High-resolution fibre bundle endomicroscopes are limited in FoV to the order of 240 µm, making surveillance of large areas of tissue difficult. Also, due to this limited FOV it is challenging for clinicians to correctly interpret and assess endomicroscopy images. This fact was clearly identified in the work of Patsias et al. [223] where the clinicians found it challenging
to obtain a broad sense of tissue morphology. Clinical applications have highlighted the importance of wide FoV coverage through video mosaicking, allowing the user to move from point-based “optical biopsy” to regional analysis. Video (or image) mosaicking algorithms stitch together multiple sequential video frames in order to generate a wider environment map of a scene. Several recent publications have been focused on video mosaicking of confocal images, which are mainly monochromatic greyscale images.

The following sections provide a more detailed description on how to generate large area mosaics by stitching together sequential microscopic images. The novelty of the proposed algorithm is the real-time performance in conjunction with the high frame rate endomicroscopy system (Appendix A.1). The algorithm has also been extended to colour microscopic images, including images acquired by a dual wavelength CLE system or an endocytoscopy system. It provides a solution for motions that are not only 2D translational movements but also rotations between successive image frames. All these additions create a versatile framework for mosaicking microscopic images.

3.3.1 Materials and Methods

3.3.1.1 Basic Video Mosaicking Algorithm

The basic concept of video mosaicking is to stitch together consecutive images and for this a low-complexity but efficient algorithm is implemented based on pair-wise normalised cross correlation (NCC) coupled with a template-based image registration technique. This has been already introduced in various works for endomicroscopic mosaicking but here it will be extended to high frame-rate and colour systems. The algorithm follows the implementation of fast normalised cross correlation algorithm based on the seminal work of Lewis \[229\]. In particular, the approach is as follows:

For successive pairs of images, \(I_1\) and \(I_2\) (dimensions \(M_{Ix} \times M_{Iy}\)), a central template \(t\) of smaller dimensions (\(N_{tx} \times N_{ty}\)) is extracted from \(I_1\). The 2D cross correlation between the template and the image \(I_2\) is calculated to find the NCC value for each displacement \((u, v)\) of the template over the image \(I_2\):
Figure 3.9: (left) Normalised cross correlation algorithm outputs an array with dimensions equal to the combined dimensions of the source image (e.g. 400 pixels here) and the template image (e.g. 50 pixels here); the intensity of each pixel is determined by the NCC value; here the peak in (165,225) is depicted with $NCC = 0.9775$; the distance between this pixel and the centre of the image corresponds to the shift from the previous image to the current image, (right) illustration depicting the overall mosaic image when stitching two images as well as the shift from the previous image to the current and the corresponding template images.

The normalised cross correlation value is defined as:

$$NCC(u,v) = \frac{\sum_{x,y}[I_2(x,y) - \overline{I_2(u,v)}][t(x-u,y-v) - \overline{t}]}{\sqrt{\sum_{x,y}[I_2(x,y) - \overline{I_2(u,v)}]^2 \sum_{x,y}[t(x-u,y-v) - \overline{t}]^2}}$$  \hspace{1cm} (3.5)

where $\overline{I_2(u,v)}$ is the mean value of the image $I_2$ within the area of the template $t$ which is shifted by $(u,v)$ and is calculated by:

$$\overline{I_2(u,v)} = \frac{1}{N_xN_y} \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} I_2(x,y)$$  \hspace{1cm} (3.6)

and $\overline{t}$ is the mean value of the template image.

Then, based on the maximum cross correlation value, the location of this peak $NCC_{\text{max}}(i,j)$ is used as an estimate of the shift $(i,j)$ between images $I1$ and $I2$ (see Figure 3.9).

In this approach, we assume that non-rigid tissue deformations can be ignored and that motion distortions are negligible. These assumptions are realistic in the applications that
will be presented in this thesis because: 1) The high frame rate endomicroscopy system used is less susceptible to motion-induced distortion than lower frame rate, point-scanning systems. 2) In force-controlled robotic applications, the forces applied to the tissue are significantly less than handheld or endoscopic manual manipulations of microscopic probes; as a result, morphological changes are eliminated between consecutive frames. Based on these assumptions, we can perform rigid registrations of the input frames.

The algorithm is enhanced by a two-way approach as both forward ($I_1$ to $I_2$) and reverse directions ($I_2$ to $I_1$) are used to find the location of the cross correlation peak, with the result of the highest cross correlation value taken as the best estimate of the shift. Furthermore, if the confidence is low, i.e. less than 85% (this value has been found empirically to provide optimal results in tissue samples), then the shift is considered as the mean value between the forward and the backward direction. To improve the robustness and avoid local minima the shift values are compared to previous shifts and in the case of an excessive shift, then the result is instead taken as the mean value of the previous 2 steps.

The position of the image template is selected to be in the centre of the image as it allows 2D movements around the centre of the image and it is more probable to find this template in the next video frame. Three different template sizes are used in NCC (30% - 50% - 70% of the original size dimensions) and the template size with the highest cross correlation value is used for the final result and the reverse direction. In early implementations of mosaicking techniques, the template size was a manually tunable parameter that depends on the probe scanning velocity and the image overlap; high speed scanning with small overlap leads to errors with large templates, e.g. if the template of image $I_1$ is large and the overlap small then image $I_2$ will not have sufficient features to correlate the two images; on the other side, large templates and overlaps improve significantly the robustness and accuracy of the shift result.

The mosaic is reset if the cross correlation falls below a pre-defined threshold, or if the intensity of the image is below a user-specified minimum value (meaning that there is no image present), or if the mosaic runs outside of the image size. This also addresses
partially the issue of non-rigid registration as the latter will lead to poor cross correlation and therefore the mosaic will reset.

The mosaicking algorithm was initially tested and implemented in the Matlab environment, where the achieved frame rate was 48 \( \text{fps} \) for the simple pairwise registration. In the high frame rate line-scanning endomicroscopy system, the mosaicking algorithm is implemented in Labview in the form of a DLL, written in C++ using the openCV library. This improves the performance of the algorithm significantly and a frame rate of 120 \( \text{fps} \) acquisition rate can be achieved. It should be mentioned, also, that only one template size is used in real-time mosaicking in order to keep the update rate high.

### 3.3.1.2 Rotation-Invariant Video Mosaicking Algorithm

Robotic mechanisms such as the robotic large area scanner for intraoperative breast endomicroscopy proposed by Zuo et al. [4] result in rotation, as well as translation, of the probe due to their adopted mechanical design. The aforementioned NCC algorithm can fail when there is successive rotation between frames. This means an additional DoF has to be considered. Consequently, two simple, offline mosaicking algorithms have been implemented in Matlab\(^\text{®} \) (Mathworks, Massachusetts) to include the handling of rotational motion.

For initial experiments, the calculation of the rotation angle between successive image frames is identified by calculating the Radon transform [230]. The Radon transform projects the image intensities along a specified angle and the resulting projection is an integral of the pixel intensities along the line oriented on this angle. This inherent property of the Radon transform provides useful characteristics to capture the directional information of the images. Specifically, the radon transform of a 2D function is an integral defined as:

\[
R(r, \theta)[f(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(r - x \cos \theta - y \sin \theta) dx dy
\]  

(3.7)

where \( r \) represents a line’s perpendicular distance from the origin and \( \theta \) is the corresponding angle between the y-axis and the line. If the Radon transform is combined with
a Fourier transform then the result provides the polar coordinates of the image that incorporates the angular characteristics of the image. Therefore, by comparing the results from the previous and the current image, the rotation angle can be identified and used as the most likely rotation angle. This implementation uses the Radon transform algorithm already integrated within Matlab®.

However, the Radon transform when applied to noisy images can generate false positives and therefore local exhaustive search, which is more robust but more computationally expensive, is chosen for offline processing. In particular, the software is set to rotate each image frame by a number of trial angles (between 0 and 8 degrees in steps of 0.5 degrees) and the normalized 2D cross correlation between each rotated image and the previous frame is computed. The cross correlation peaks can then be compared and the highest value is selected as the most likely rotation angle and lateral shift. If the cross correlation peak falls below a threshold (manually tuned depending on the tissue type), the shift and rotation are instead assumed to be the same as for the previous pair of frames, which is a reasonable assumption for robotic scanning devices.

### 3.3.1.3 Blending Algorithm

For improved visualisation, a blending algorithm is applied between the mosaic image and the overlaid image due to artefacts that may appear on the edges of overlapping images caused by non-uniform light intensity (see Figure 3.10). In this implementation, distance weighting is used to account for these differences. Both the current mosaic \(I_{mc}\) and successive frame \(I_i\) are converted into a binary image and a distance transform is performed on each image. The resultant values are then used as weights and the weighted mean intensity is calculated from both images, for each pixel of the mosaic.

In particular, the resulting mosaic image \(I_m\) will be:

\[
I_m = aI_{mc} + (1 - a)I_{mc}
\]  

(3.8)

where the constant \(a\) is calculated as:

\[
a = \frac{d_{trans1}}{(d_{trans1} + d_{trans2})}
\]  

(3.9)
Figure 3.10: Mosaic images with and without blending processing.

3.3.2 Experimental Results

3.3.2.1 Basic Video Mosaicking Algorithm

Exemplar mosaics using the NCC mosaicking approach are presented in Figures 3.11 and 3.12. The mosaics of Figure 3.11 are generated using the rigid spiral scanning instrument that will be showcased in chapter 5 and demonstrate mosaics of lens cleaning paper and porcine fat tissue.

The mosaic of Figure 3.12 is generated by a prototype 3 DoF instrument for colorectal surgery (more details in appendix C) and presents a large area mosaic of lens cleaning paper of 21.6 mm$^2$. 

101
In the developed endomicroscopy systems, the mosaicking algorithm is implemented in native Labview Vis except the NCC algorithm which was written in C++ using the openCV library and then converted into a DLL. This improved the performance of the
algorithm significantly and mosaicking with 120 $fps$ acquisition rate can be achieved.

### 3.3.2.2 Video Mosaicking from Colour Microscopic Images

Figure 3.13 shows an exemplar mosaic, 3.8 $mm$ in length, created after manually scanning over *ex vivo* porcine colon mucosa stained with toluidine blue 0.1%. The results will, of course, depend on the speed and consistency with which the probe is moved across the tissue, as well as the heterogeneity of the images.

From the results presented, it can be seen that the generated mosaic has pixel-level accuracy when the registration between the images is good. In cases where the registration between the consecutive frames is poor, either due to poor video quality or fast movement, spatial information derived from the robotic kinematic chain can be used to improve the consistency of both the mosaicking and tracking processes.

### 3.3.2.3 Video Mosaicking Incorporating Rotational Motion

The radon transform algorithm was initially developed as a rotation-invariant mosaicking algorithm. An exemplar mosaic from a simulated trajectory is presented in Figure 3.14(a). In order to create the simulated image dataset, a large pseudo-histological image was used; a spiral trajectory was defined inside the image and in each trajectory point (used as the
frame centre) each frame was cropped. In order to simulate the rotational motion, in each trajectory point the obtained image is rotated by the desired angle.

The aforementioned mosaicking algorithm that is independent of rotational motions is demonstrated using a fabricated prototype of a large area scanner for breast endomicroscopy, reported in [4] and not discussed further in this thesis. The prototype has the capability to scan over an area using various trajectories, such as linear scans, concentric circles and spiral scans. This was demonstrated by scanning a non-overlapping spiral pattern over a piece of lens tissue paper stained with acriflavine. The resulting mosaic is shown in Figure 3.15.

Since the position and angle of each frame is determined by comparing it with the previous frame, this method accumulates errors during the scan. This would prevent a gap-free image from being generated, as it is unlikely that the loops of the spiral would correctly align. The approach using the radon transform is also prone to accumulating errors. A more complex mosaicking algorithm would therefore be preferable for clinical use, one
that uses positional feedback from the motor and takes into account tissue deformations. Given the likely high computational requirements of such an algorithm, it is suggested to implement a simpler, online version for provisional, real-time visualisation, such as NCC algorithms, and a more complete, offline version for diagnostic purposes.

### 3.4 Intraoperative Fusion of Macroscopic Views with Microscopic Images

In the last section of this chapter, the first step towards bridging the gap between scales will be provided in the form of a fusion approach to combine the macroscopic stereoscopic view with the microscopic image mosaics. The preliminary work and proof of concept has been developed by the author of this thesis for transanal surgery using a prototype 3-
DoF instrument for colorectal surgery (more details in Appendix C). This work has been recently improved with the collaborative work of Mr. Lin Zhang and the integration of the framework with the Da Vinci robotic system using the da Vinci Research Kit (dVRK) and is presented in Appendix B.

3.4.1 Materials and Methods

The current study involved a method for augmented reality visualisation to fuse high resolution microscopic images of the cancerous region, acquired intraoperatively, with the endoscopic views in a three-dimensional representation of the surgical field in the operation of Transanal Endoscopic Microsurgery (TEM). A Wolf stereo endoscope (Richard Wolf, Germany) is used for imaging. The endoscope has a tilted lens and observes the surgical field at an angle of 50°. A custom bracket secures two Storz HD (1920x1080i) camera heads (Karl Storz, Germany) on the endoscope. Thus, during endoscope manipulation, the relative camera positions and the stereo baseline are held constant. The long length of the endoscope allows operations in the majority of the lower rectum. Together with the instruments, it is inserted through a Wolf TEM rectoscope tube and faceplate (Richard Wolf, Germany). In this study, the endoscopic images are sub-sampled by a factor of four to achieve real-time performance. Microscopic images are captured using a widefield optical microscopy probe of an established design (Appendix A.2). The instrument is actuated under robotic control for large-area scanning (see Appendix C). A checker-board pattern (KeyDot®, Key Surgical, Minnesota) is securely attached at the distal end of the endomicroscopy probe, thus having a fixed, known relationship between the imaging tip and the probe shaft. The marker remains visible in the stereo endoscopic images during scanning (see Appendix C).

The workflow of the algorithm is as follows. First, camera calibration is performed offline [231]. Online, the stereo images are rectified based on the calibration parameters. Subsequently, stereo reconstruction using semi-global block matching [232], as implemented in Matlab® (Mathworks, Massachusetts), is carried out to extract the three-dimensional structure of the FoV. Semi-global block matching enforces consistent disparities among
neighbouring pixels in the scene [232], alleviating several of the problems encountered due to a lack of high frequency texture in endorectal images. The endomicroscopy probe is actuated and the obtained images are processed offline to remove the honeycomb pattern due to the fibre-bundle cores, and then the images are mosaicked. During scanning, the checkerboard marker at the distal end of the probe is tracked in the stereo images using a checkerboard pattern detection algorithm based on corner extraction [233]. The calculated disparity, back-projection, and triangulation, result in localisation of the probe in the three-dimensional field-of-view [234]. The relative pose of the probe’s tip to its calibration pattern is known from the CAD models of the components and is fixed throughout the experiments. Also, the camera pose is fixed during both the reconstruction and tracking phases. Hence, tracking and localising the checkerboard pattern relates the microscopic image and the mosaic to the three-dimensional reconstructed FoV (see Figure 3.16). Finally, the mosaicked images are warped on the 3D reconstruction according to the results of probe localisation.

### 3.4.2 Experimental Results

The system and algorithm are evaluated with ex vivo tissue, specifically a porcine colon tissue specimen that gives rise to clinically relevant endoscopic and microscopic images. The tissue sample is brought to room temperature prior to experimentation, and is stained with proflavine fluorescent contrast agent. To simulate physiological curvature, the tissue
is placed in a curved plastic container. Figure C.2 shows the distal side of the experimental apparatus. The three-dimensional FoV is reconstructed, and, as shown in Figure 3.17, the curvature of the tissue is captured. The probe is actuated while keeping contact with the tissue to cover a linear region of approximately 10 mm. Tracking results and triangulation relate the microscopic views to the global coordinate frame of the 3D reconstruction allowing the fusion of the created mosaic with the coarse reconstructed anatomy. Hence, the presented three-dimensional map may provide information on multiple scales, ranging from high-level anatomy to the microscopic level, and allows closer identification of tumour margins and tissue characteristics (Figure 3.17).

### 3.5 Discussions and Conclusions

In this chapter, an overall image processing framework for fibre-based endomicroscopy imaging is described. Due to the use of a fibre-based system, raw images are reconstructed based on a triangular linear-based interpolation approach in order to reject the honeycomb-pattern artefact. Similar approaches have been explored in the literature.
However, with the proposed algorithm, a technique to suppress broken / damaged fibres is presented both for monochromatic and coloured imaging systems, as well its real-time implementation is discussed. Regarding the speed, the current image reconstruction implementation can accomplish up to $12 \, f ps$, which means that it can be applied to video-rate imaging, but for high frame systems the simpler Gaussian filtering approach is selected. This can be improved in future work by implementing the algorithm in C++ and using processing with parallel computing techniques.

The small FoV of endomicroscopy probes makes it challenging to evaluate larger areas of tissue, a problem that has motivated the development of mosaicking algorithms for fluorescence confocal endomicroscopy. The creation of mosaics has been demonstrated for all the in-house monochromatic and colour endomicroscopy systems, demonstrating the potential to provide a much larger FoV for optical biopsy than has previously been possible. The novelty of the mosaicking algorithm relies on its implementation with the high frame rate endomicroscopy system as it can achieve $120 \, f ps$ without any compromise in terms of robustness. Additionally, a mosaicking algorithm that is rotation invariant, a feature that facilitates the design of certain mechatronic instruments, is described and evaluated. Both algorithms are based on pairwise registration between sequential frames. This, however, leads to potential cumulative errors when very wide areas are mosaicked. Therefore, global registration algorithms are needed as a post-processing step to reduce these errors, as described in section 2.4. An alternative approach to eliminate potential mosaicking errors, specifically applicable to robotic applications, is to use a hybrid approach that combines both imaging and position information as it has been presented in the works of [118, 211].

In the last part of this chapter, results for data fusion for three-dimensional macroscopic images are presented. Our early method was completely offline, which hinders its use for intraoperative applications. However, it suggests that a potential fusion of multi-scale images can significantly improve surgical guidance intraoperatively as the surgeon can operate while looking at cellular scale histology-like information. This also assists with pathological site re-targeting as the biopsy sites can be tracked intraoperatively. Recent work has attempted to alleviate these challenges by implementing an online fusion al-
algorithm that displays in the same operating scene, both the 3D reconstructed site and the generated mosaics.

To conclude, in this chapter an imaging framework that offers the surgeon much more comprehensive information is developed. It greatly improves the prospects for intraoperative margin evaluation. These algorithms are crucial for taking full advantages of the robotic scanning systems described in the remaining chapters of this thesis.
Chapter 4

Robotic Pick-up Probes for *in vivo* Endomicroscopy\(^1\)

4.1 Introduction

As presented in Chapter 2, a disadvantage of using fibre bundles for endomicroscopy is that the resolution and FoV are limited by the finite number of fibre cores. In order to obtain clinically useful resolutions (typically less than 10 \(\mu m\) for cellular imaging), the FoV has to be limited to less than half a millimetre. Despite the advantages of endomicroscopic imaging, intraoperative interpretation of the images acquired can be challenging. This fact was clearly identified in the work of Patsias *et al.* [223], where the clinicians found it challenging to obtain a broad sense of tissue morphology. The use of mosaicking to increase the effective FoV, as presented in the last chapter, is only a partial solution to this problem, as it remains difficult to study large, complex, 3D areas. In current endo-

---

\(^1\)Content from this chapter was published as:
microscopic applications, probes are almost exclusively controlled manually, leading to inconsistent area coverage.

Robot-assisted probe manipulation may provide a solution to this problem. In particular, the Da Vinci® robot incorporates motion scaling along with fine instrument handling and reduced tremor for consistent tissue handling and manipulation. All these features provide the necessary precision, accuracy and stability so that an endomicroscopy probe can be manipulated consistently [223].

The aim of this chapter is to describe a novel method of robot-assisted probe manipulation using existing state-of-the-art surgical systems, providing a simplified way to translate this technology in the normal surgical workflow. Three different pick-up probes for existing surgical robotic systems, allowing integration of standard endomicroscopy probes, are presented. A simple pick-up instrument that integrates an endomicroscope probe to facilitate scanning of large areas will be briefly described along with examples of its use. This is then followed by a short description of a pick-up probe that includes both OCT and endomicroscopy probes for multi-modality optical imaging. Finally, a force adaptive pick-up probe that can be used in conjunction with the Da Vinci® instruments for intra-operative endomicroscopy imaging is described. The device uses a novel low-friction air bearing with passively adaptive axial force to maintain constant contact between the tissue and the imaging probe, facilitating microscopy scans over complex and curved surfaces such as those encountered during in vivo tumour margin delineation. Detailed ex vivo user experiments will be presented to demonstrate the effectiveness of the technique.

4.2 Single Modality Pick-up Design and Employment

In order to demonstrate the benefits of robotic manipulation in optical imaging, a simple pick-up probe has been designed that has a slot which can be grasped by the end-effector of robotic surgical devices (see Figure 4.1(a)). This concept has been tested both with the Da Vinci® Endowrist® instrument as well as with the Raven surgical robot. The probe has a through hole of \( \phi 2.7 \text{mm} \) that a \( \phi 2.6 \text{mm} \) endomicroscope fibre bundle probe (e.g.}
Cellvizio UHD probes) can be inserted into. The fibre bundle is fixed by a grub screw from the side. The pick-up also features a marker (KeyDot®, Key Surgical, Minnesota) which is attached to the probe for accurate pose estimation of the fibre bundle and its respective image position in 3D space. The probe is fabricated with Veroclear material using a rapid prototyping machine (Stratasys Objet500 Connex3, Objet Geometries Ltd., USA) and the overall dimensions allow the probe to be inserted through a common 10 mm surgical trocar port.

This probe was used in the work of Zhang et al. (see Appendix B) to automate the scanning process with the Da Vinci® system. In the same work, in order to assist the users who desire to observe not only the scanning process but also the endomicroscopy image feed, the endomicroscopy images are overlaid over the marker on the probe, in a form of augmented reality, as shown in Figure 4.1 (c) and (e).
4.3 Multi-modality Pick-up Design and Employment

Using the probe described in the last section, it became apparent that the lack of haptic feedback can result in unregulated surface contact forces. This hinders the operator from maintaining constant tissue-probe forces and therefore makes it difficult to scan over uneven or curved surfaces. Furthermore, if excessive pressure is applied, this tends to lead to severe tissue deformation and micro-structure damage, which in general also affects the quality of images and mosaics [235, 209].

To cater for tissue deformation and maintain consistent contact forces, the idea of a multi-modality pick-up probe for robotic systems is explored (see Figure 4.2(a)). This probe includes the endomicroscopy fibre-bundle that provides superficial cellular level information and an OCT probe that outputs a high resolution cross-sectional visualisation of the tissue morphology underneath the scanning area. Simple processing of an OCT image can identify the top surface of the tissue and therefore it can be used to calculate, sensor-less and with high accuracy, the distance of the probe from the tissue (see Figure 4.2(b)). Apart from identifying the tissue surface distance, the OCT image can be used as an additional optical biopsy technique and therefore, similarly to endomicroscopy, it can be scanned across the tissue surface to generate a 3D OCT volume on the fly and facilitate intraoperative tissue diagnosis.

The mechanical design of the probe features two through holes, Ø2.7 mm and Ø2.8 mm respectively, into which the endomicroscopy and an in-house OCT probe can be inserted. The dimensions of the probe allow it to be inserted through a common 10 mm surgical trocar port. The imaging probes are fixed independently by two grub screws at the side of the probe. Similarly to the previous probe, it features a marker (KeyDot®, Key Surgical, Minnesota) which is used for pose estimation. The probe is fabricated with stainless steel through Direct Metal Laser Sintering process (Concept Laser Mlab Selective Laser Sintering (SLS), ES-Technology, UK) for improved durability. This concept has been tested with the Da Vinci® Endowrist® instrument but its employment is compatible with any surgical system that has a grasping instrument. This probe was used in the work of Zhang et al. (see Appendix B) for automated 2D scanning of CLE/OCT probes, providing
Figure 4.2: (a) CAD design illustration of the CLE/OCT pick-up probe; (b) demonstration of the correlation between the distance to the tissue surface measured in the OCT image (bottom row) and the quality of pCLE images (top row). As the surface moves closer to the probe, from left to right in the figure, the probe-to-tissue distance measured by the OCT probe becomes smaller while the pCLE image quality improves.

large area tissue coverage. The diameter of the holes could be adjusted to suit other imaging probes, as required.

4.4 Adaptive Pick-up Probe for in vivo Endomicroscopy

As described in section 2.6, the two main hardware-based schemes that can facilitate the application of constant force during probe-tissue interaction, and hence help to obtain consistent microscopy images, are the active and passive force control approaches. In particular, active control techniques, as the name implies, actively control the forces exerted on the tissue at the contact point so as to maintain them at a specific value. Passive force control approaches, on the other hand, maintain constant tissue forces using a passive physical system, such as a spring mechanism. As presented in section 2.6, several instruments demonstrate improved acquisition of endomicroscopy images with the use of a force control scheme, either with an active control scheme [201, 202, 217] or with a passive, spring-based control [4, 218, 219]. Such devices have shown promising results for robotically controlled endomicroscopic imaging, but they are large, bespoke devices requiring complex motor actuation and sensing. They are designed as a stand-alone sys-
Figure 4.3: (Left) Computer-Assisted Drawing (CAD) model of the proposed drop-in probe and (right) fabricated instrument handled by the da Vinci Endowrist® Needle Driver.

tem rather than to cooperate with existing MIS platforms. Additionally, passive control approaches, that have smaller footprints and require simpler mechanisms, result in non-linear behaviours, especially in the limits of the motion, reducing their effective range.

Considering these arguments, in parallel with the previous multi-modality software-based approach for tissue distance estimation, a mechanical approach with an open loop force control pick-up probe is developed. This can also be used by robotic instruments and assist in the maintenance of consistent instrument-tissue contact. The adaptive mount is designed to hold a single endomicroscope probe and aims to minimise tissue deformation during scanning. One of the significant challenges associated with adaptive force control with micro-force sensitivity is the sensitivity of the actuator’s friction control. The new approach uses a low-friction, air bearing with open loop adaptive axial force control. This facilitates microscopic scans over complex and curved surfaces, such as those that would be encountered during in vivo tumour margin delineation.
4.4.1 Methods and Materials

4.4.1.1 Mechanical Design

The fabricated instrument, along with the CAD model, is presented in Figure 4.3. The concept is based on the fact that an air cushion can be created between two cylinders, thus reducing the friction during their motion while acting as an opposing force to elongate the probe when it is not in contact with the tissue. During operation, the air pressure and supply are kept constant in order to maintain a constant force at the tip of the probe. The operation is similar to a spring-based mechanism but the difference of this mechanism is that the air cushion keeps the contact force constant (rather than increasing with shaft displacement), whilst at the same time the air minimises the friction of the mechanism. The results suggest that the passive force adaptive probe holder, combined with a high frame rate endomicroscopy system, allows large area mosaicking with minimal tissue deformation.
To simplify clinical deployment, the fabricated probe is used as a drop-in pick-up device delivering the imaging probe, and can fit inside a common 10 mm surgical trocar port. It can be dropped through the port at the operating site and then picked by an Endowrist® grasper or needle driver.

At the lower part of the device, the imaging probe is fixed and a marker for probe tracking is attached. The tracking will not be further discussed here but it can be used in the same manner as with the previously mentioned pick-up probes. The outer tube of the proposed device has an outer diameter of 5 mm and inner diameter of 4.1 mm. The corresponding values for the inner tube are 4 mm and 3.2 mm, respectively. Both tubes are made of machined stainless steel, while the part that fixes the imaging probe and the other part that includes the pick-up handle were made with rapid prototyping stainless steel through Direct Metal Laser Sintering process for improved durability. Constant compressed air is supplied by an air-compressor at 5 PSI, a value that is determined through a trial-error process to ensure less than 100 mN forces, with an air tank for smoothing out the flow, through an 8 mm diameter flexible plastic tube. The range of the pick-up probe is selected at 8.6 mm with the aim to provide a safe range for tissue motion compensation and reaction from the operator. The range was also selected to be small enough not to create burden for the operator during manipulation; the larger the overall length of the probe, the more uncertainty there is during probe manipulation with the robotic grasper. Additionally, in the case where the range limits are exceeded, the probe acts as a rigid pick-up mount and any further pressure exerted is transferred to the tissue.

4.4.1.2 Imaging System

The endomicroscope probe used in the following experiments in conjunction with this probe is a Gastroflex UHD fibre bundle probe (Cellvizio, Mauna Kea Technologies) fitted to the in-house, high speed laser line-scanning system (see Appendix A.1). The probe is designed to be placed in direct contact with the tissue under study, having a working distance of only a few 10s of microns. It incorporates an approximately x2.5 magnification distal micro-objective, resulting in a FoV of 240 microns with a resolution of approximately 2.4 µm. The raw image was streamed directly to disk at 120 fps, while an
on-screen preview was shown at 20 frames per second. This preview was fed to the Da Vinci® master console, and displayed to the user in the upper-right corner of the normal 3D endoscope view (picture-in-picture view). For this study, all mosaicking was performed off-line. The image processing algorithm used in this work is described in more detail in chapter 3.

### 4.4.1.3 User Validation

A pilot ex vivo user study was conducted to perform an initial assessment of the clinical potential of the proposed device. Five users range from novice to experienced, both in the use of the Da Vinci® system and in their knowledge of endomicroscopy and mosaicking, were recruited. Details about each user are presented in Table 4.1.

A phantom was fabricated using rapid prototyping technology, which incorporates a tubular structure and a curved surface in the centre. A piece of porcine colon tissue was stained with a fluorescent dye (proflavine) and laid over the phantom. Colon tissue was chosen for these initial experiments as it has strong features for mosaicking, particularly the mucosal crypts. Also, previous studies have shown that the probe-tissue contact force alters the colonic crypt morphology seen in the endomicroscopy images, making this an ideal candidate for assessing the consistency of the contact force [209].

### 4.4.2 Experimental Results

#### 4.4.2.1 Pick-Up Probe Demonstration

Figure 4.5 shows the results of a simple but illustrative demonstration of the force adaptive system. The pick-up probe, with the probe inserted, was grasped by the da Vinci Endowrist® instrument, and driven across the surface of a scientific weighing scale for 60 s, showing an approximately constant force averaging to 97 mN (standard deviation: 21 mN). The same experiment was conducted with a rigid probe for comparison, with the user attempting to maintain a small and constant contact force. The latter experiment showed an average force of 613 mN (standard deviation: 498 mN). This illustrates that
the passive probe is indeed able to maintain a small and constant probe-sample contact force in comparison to a rigid probe.

![Scan Task Over A Force Sensor](image)

Figure 4.5: (a) Graph of a 60 s sample recording of the scale demonstrating that the adaptive probe proposed can apply consistent low magnitude forces to the surface of a scientific scale. Force measurements were made by recording a video of the scales as the probe was driven across. Then, from this video, each measured weight value was manually recorded every 0.22 s; (b) Frames from the recorded video presenting the adaptive probe (left) and the rigid probe (right) and their respective weight values in grams.
Table 4.1: Demonstration of each user's expertise (+ = Limited experience, ++ = Intermediate experience, +++ = Expert) and total mosaic length in mm achieved by each user over a large 3D tissue surface in the period of 1 minute.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Endomicroscopy</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Da Vinci®</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>120 Frames per Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive force probe</td>
</tr>
<tr>
<td>Rigid probe</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>30 Frames per Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive force probe</td>
</tr>
<tr>
<td>Rigid probe</td>
</tr>
</tbody>
</table>

4.4.2.2 Robotic-assisted Endomicroscopy Scanning Study over Complex Surface

The results of the user study are presented in Table 4.1. Users were asked to attempt to make several linear scans over this large area, from back to front, for one minute, following an initial training period of five minutes. The users performed this experiment both with the adaptive probe, and with a rigid probe in a random sequence. Exemplar images from the Da Vinci® endoscope view, showing the two probes being grasped by the Da Vinci®, can be seen in Figure 4.6(a-b). Exemplar endomicroscopy image frames can be seen in Figure 4.6(c-d); those were chosen to highlight the increased distortion of the crypts that tended to occur when the passive force control was not used.

The resulting videos were processed to generate mosaics. All viable mosaics were extracted from each video, where a mosaic was considered to be viable if the NCC peak value between each successive pair of frames was greater than 0.85, and where the total length of the mosaic was greater than 10 frames. The videos were then down-sampled to 30 fps and reprocessed in order to assess the necessity of using a high frame-rate endomicroscopy system.

The longest mosaics generated by each user for each probe and for the 120 fps videos are shown in Figure 4.7. The increased length of the mosaics from the use of the adaptive force probe can be clearly seen, as also can a reduction in distortion of the crypts. The latter is due to the generally lower contact forces resulting in reduced tissue deformation.
Figure 4.6: (a-b) Snapshots during the tissue task where the users had to scan, with the adaptive robotic probe (a) and a rigid probe (b), a large curved 3D surface and acquire consistent microscopy images. (c-d) Indicative microscopy images sampled during the tissue task with the adaptive probe (c) and the rigid probe (d).

Statistical results relating to the mosaics are shown in Table 4.1 (total length of all mosaics from each video) and in Figure 4.8 and 4.9 (average length and standard deviation of all mosaics from each 1 minute video), for both the 30 fps and 120 fps videos. The improvement is clear for the 120 fps videos, in the sense that longer mosaics are generated by the force adaptive probe, but there appears to be little or no improvement at 30 fps. There is also a clear improvement between the 30 fps and 120 fps videos, as expected, since for a given probe velocity there is an increased opportunity for the frame overlap required for mosaicking.
Figure 4.7: Indicative mosaic results of all the subjects that participated at the tissue study, demonstrating the largest mosaics achieved with the adaptive probe and the rigid probe respectively. The mosaics depicted have been scaled with the same ratio to allow comparison of their lengths. Zoomed versions of microscopy segments are presented to assess the changes in the morphology of the cells, with and without the adaptive force probe.
Figure 4.8: Average size of generated mosaics from each user during the 1 minute allowance for the tissue task. The images correspond to the raw acquisition rate of 120 fps.

Figure 4.9: Average size of generated mosaics from each user during the 1 minute allowance for the tissue task. The images correspond to the downsampled acquisition rate of 30 fps.
4.5 Discussions and Conclusions

In this chapter, we have demonstrated the development and use of three simple yet versatile pick-up probes that can be used by robotic instruments, such as the Da Vinci® ones. The first pick-up probe carries only a single fibre bundle. Despite the fact that contact forces are not regulated in this probe, it showcases an improved design compared to other works using robotic systems, such as in [223, 221, 222], as it is easier to grasp using the ergonomic slot rather than grasping the cylindrical probes directly, which is very unstable. The probe also features a marker that can be used for position estimation. These two design approaches are followed with the other two described pick-up probes.

The multi-modality pick-up probe provides a fixed platform to manipulate simultaneously two different fibre-based probes, allowing the acquisition of different functional and structural information in the same frame of reference and therefore registered rigidly together. The attached marker allows the modalities also to be registered to the endoscopic view and as a result to the global frame of reference, providing a step towards augmented visualisation and biopsy re-targeting intraoperatively. Here, an OCT fibre probe is used together with a high frame-rate endomicroscopy system, providing not only surface cellular information but also cross-sectional cellular information, which is particularly useful for intraoperative guidance for the identification of tumour margins. The OCT images, as described, could be used as an alternative for accurate distance sensing. Since contact is not required with the OCT images, constant tissue contact can be calculated if there is a small space between the OCT probe and the tissue. This is particularly useful in autonomous scanning tasks as the robot can regulate the distance to be constantly held at a specific value during the task. The disadvantage of using this approach, however, is that the performance of the system depends on the image quality of the OCT image. Filtering or predictive control is needed if the image is of poor quality and noisy.

An alternative hardware-based approach towards maintaining constant distance, or in other words keeping constant force, between the probe and the tissue is the third proposed pick-up probe that includes a open loop force control mechanism. The use of an ultra-low friction air bearing has permitted the pick-up probe to adapt and compensate...
to complex and curved surfaces and facilitate large area image mosaicking. The proposed device requires a limited compressed air supply, which for MIS is convenient as it is readily available for maintaining pneumoperitoneum. User studies were performed to demonstrate the practical clinical value of the device compared to a rigid probe.

The results derived suggest that the passive force adaptive probe holder, combined with a high frame rate endomicroscopy system, allows large mosaics to be generated over curved surfaces. The device is able to maintain through open control a small and consistent contact force between the probe and the tissue, offering the best prospect for good quality images and mosaics. A more extensive study on the effects of the overall air pressure versus the force exerted to the tissue will, also, provide a better insight on the benefits of this device. Interestingly, for the 30 fps imaging, the benefits of the force adaption do not appear to be realised. This may be for one of several reasons. The force adaptive probe tended to make it easier to scan across the surface, possibly encouraging the user to scan at a higher speed than when using the non-adaptive holder. This higher speed may have led to insufficient overlap between frames at 30 fps for the mosaicking algorithm to calculate the shift, resulting in apparently degraded performance for the force adaptive holder. A second, related, possibility is that the users were always scanning too fast for mosaicking at 30 fps, but that without the force adaption, the probe tended to drag the tissue along with it to some extent, reducing the effective velocity and allowing mosaics to be created. Further work will be needed to confirm these effects. Moreover, from the user studies, it was identified that the large length of the force adaptive probe created some ergonomic issues when manipulated, which will be addressed in the new version of the probe.

Finally, it is important to note that the proposed probes are not limited to Da Vinci® instruments, they can also be used in conjunction with standard laparoscopic surgical instruments. The results of this work may therefore be extended to a much wider range of surgical procedures than those currently conducted using the Da Vinci®. For example, with further miniaturisation of the force adaptive pick-up probe, it may be possible to make use of this mechanism in standard diagnostic procedures where the endomicroscope is introduced through the working channel of an endoscope.
Chapter 5

High-Speed Handheld Robotic Scanning Device with Image-Based Motion Control and Interventional Capabilities

5.1 Introduction

As discussed in section 2.5.4, cooperatively controlled and handheld surgical instruments provide a more intuitive and natural operation, allowing easier integration with other computer-assisted surgical systems while reducing manufacturing and operational costs. In this chapter, a novel, handheld, robotic device for fibre-based optical imaging will be introduced, which combines high-speed robotic scanning with image-based motion control. The device can also integrate focused energy delivery, allowing targeted therapy. This device is intended, initially, to be used for breast conserving surgery where intra-operative margin assessment is crucial, but the clinical application can be extended to other open surgical procedures. Breast cancer is the second most common cancer in the world, and the second most common cause of cancer deaths of women in developed regions [236]. This global health burden has motivated the development of less invasive
treatments including breast conserving surgery (BCS) [237]. With BCS, which is indicated for small, localised cancers, only the tumour is excised, together with a surrounding margin of healthy tissue (approximately 2 – 5 mm wide). This is then followed by radiation therapy. BCS is currently the “gold standard” for patients with small-sized tumours and is considered adequate when the tumour is excised with negative histological margins. A positive margin is defined as the presence of cancerous tissues on the surface of the excised tumour, which is normally marked with ink for easy identification. This implies a potentially incomplete excision with an associated two-fold increase in the likelihood of ipsilateral breast tumour recurrence (IBTR) [238]. Margin status is the strongest predictor of local recurrence after BCS, and with re-excision rates being between 30 and 60% in USA, and 28 and 18% in England and Wales, [239], there is a clear need for better intraoperative assessment of margins to reduce the need for re-operations.

Currently, intraoperative margin assessment is performed by gross examination by palpation [240], intraoperative radiography on specimens [241] or intraoperative ultrasound guidance [242]. None of these approaches are reliable, with post-operative histology still required for definitive confirmation of negative margins. Frozen section analysis, which provides a fast alternative to conventional histological examination, can be used for intraoperative diagnosis. This, however, is expensive and can occasionally suffer from artefacts due to freezing. It is based on discrete sampling [243] and perhaps most importantly, severely disrupts the surgical workflow, since processing still takes 20 – 30 min. Alternative techniques for intraoperative margin assessment, including radio-frequency spectroscopy (RFS) [244] and rapid evaporative ionization mass spectrometry (REIMS) [245], are being developed.

While initial development has focused on diagnostic applications, the real-time nature of endomicroscopy means that it lends itself to intraoperative imaging [246] by providing cellular scale images that could previously only be obtained by histological methods. Preliminary studies have shown potential applications in neurosurgery [247], for optimising resection, in head and neck surgery [248] for early detection and resection of squamous cell carcinomas, and for diagnostic purposes in gastrointestinal, colorectal, gastric, urinary tract, ovarian and lung cancer [41]. More recently, it has been suggested that endo-
microscopy could be used for intraoperative margin assessment in breast surgery. Chang et al. [239] and De Palma et al. [249] reported the ability of surgeons and pathologists to accurately distinguish neoplastic from non-neoplastic morphological appearances from pCLE morphological images.

As already mentioned in previous chapters, significant hurdles in the applications of endomicroscopy probes is that their FoV is limited and surgeons find it difficult to maintain stability, accuracy and dexterity inside MIS cavities. Rather than attempting whole-cavity scanning, as in the work of Zuo et al. [4, 218], the focus of this work is to instead provide the surgeon with a tool for rapid assessment of tissue microstructure that can be integrated into the normal surgical workflow. By adopting a novel and highly repeatable method of scanning, a rigid probe with micrometer-scale accuracy can be achieved over a mosaicking workspace of 14 mm². The system is demonstrated together with a high frame rate endomicroscope system (see Appendix A.1) that significantly reduces scanning time compared with the current state-of-the-art. Also shown is how robust mosaic generation over deforming tissue structures can be achieved using a visual servoing algorithm based on the endomicroscope images. Finally, for the first time in robotic assisted endomicro-
scopy, energy delivery (a $CO_2$ laser in our exemplar system) can be incorporated into the scanning device, either for image-guided ablation or for marking of tumour margins.

### 5.2 Materials and Methods

#### 5.2.1 System Overview

The main component of the proposed system is a robotic instrument that incorporates a micro-manipulation scanning mechanism, two action buttons, an ergonomic casing and a 58 mm long steel tube of $\phi 3.3$ mm outer diameter as a channel for passing multiple fibres (see Figure 5.1). An imaging probe and an energy delivery fibre can be passed through the $\phi 2.7$ mm diameter bore of the tube and fixed via a locking mechanism. The same channel can be used with other types of optical imaging probes, providing to the surgeons an accurate, scanning platform for various optical biopsy techniques.

For actuation, two $\phi 6$ mm micro servomotors (Brushless DC-Servomotor 0620, Faulhaber, Germany) are used with a 256 : 1 reduction gearbox (Faulhaber, Germany). These servomotors have an integrated magnetic encoding system (analogue hall sensors (K2280,
Figure 5.3: Framework overview of the proposed system.

Faulhaber, Germany)) that allows for closed-loop position control using a dedicated motor controller (MCBL 3002 S CF, Faulhaber, Germany). The motor controllers also provide the necessary power to drive the motors. To further reduce the dimensions and weight of the instrument, the motor controllers are placed distally to the actuation pack and enclosed in a box for plug-and-play connectivity. To reduce reflections of signals on the CAN bus due to the multiple devices, a termination resistance of 120\( \Omega \) is used. The embedded software of the motor controllers communicates in a two-fold manner. For the initialisation process, and powering on/off of the motors, the kinematic control software communicates over a CAN bus (Faulhaber CANopen interface) with 1000 \( kbps \) Baudrate. However, the main kinematic control is performed through analog voltage signals using the analog input of the motor controllers. This option was selected over CAN bus due to the faster update rate (200 Hz). The analog voltage input is provided through a data acquisition card (PCIe-6321, National Instruments, USA) that communicates through a Labview interface using a PCI Express interface (1 \( kHz \)). To improve the handheld operation of the scanning device, two multi-purpose buttons are included on the instrument’s casing in an ergonomic position. The two buttons have common ground and are connected through a
pull-up resistor circuit (located inside the controller box) to digital input lines of the data acquisition card. The motor controllers and the pull-up circuit are powered with a 6 V power supply, which is controlled by a switch integrated on the controller box.

The robotic scanning device can be either used in a handheld manner or attached to a robotic arm and be manipulated in a cooperative approach. The latter yields increased stability during scanning and would help avoid fatigue of the surgeon. The robotic positioning arm is an articulated robot with six degrees of freedom, actuated by brushless DC motors and coupled with harmonic drive gears, and can hold a 1.5 kg maximum payload. The arm weights 3.0 kg and reaches 770 mm at extended pose. It can be manipulated by hands-on positioning when a foot pedal is pressed. In case of power loss, each joint can be back-driven by the operator [250].

An overall illustration of the system components is given in Figure 5.3.

### 5.2.2 Mechanical Design of the Robotic Scanning Device

In practice, it is challenging to develop micro-manipulation mechanisms that are physically compact, precise and with sufficient workspace. The proposed robotic scanning device utilises a flexure mechanism to provide precise, controlled motions of the instrument tip with minimal backlash. The steel tube is clamped in a cantilever configuration so that it can deflect in two planes at its free end. Two cam-roller mechanisms\(^1\) are then used in conjunction with micro servomotors to deflect the tube which yields an approximately planar motion of the tube tip. This configuration allows for a slender, low-profile actuation system that fits in to a \(\varnothing 27\ mm\) hand-held device (including the exterior casing). A V-profiled steel cam is welded to the tube shaft; it is engaged by two steel levers with tip mounted bearings that exert lateral forces to the cantilevered tube (see Figure 5.4(b)). The levers are actuated by the two servomotors. The cantilevered tube is mounted in an unloaded position outside of the nominal workspace of the device; this ensures that the

---

\(^1\)The initial concept design of the tubular chassis and V-profiled cam mechanism has been developed and fabricated by Dr. Christopher Payne, while the author of this thesis extended the mechanical design work by optimising the instrument’s workspace, designing all the ergonomic components, extending the functionalities for energy delivery and developing all the software and electronics components.
Figure 5.4: (a) Exploded view of the handheld robotic scanning device and (b) three different scanning positions presenting the steel level positions with the resulting deflection on the tube.

cantilevered tube is always deflected and thus preloaded against the cam-roller mechanisms so as to circumvent backlash errors. The cantilevered tube and micro servomotors are all integrated into a tubular chassis, which is enclosed by an ergonomic casing and can be handheld by the operator. The tubular chassis is fabricated with Veroblack material using a rapid prototyping machine (Objet EDEN350™, Stratasys Ltd., USA) while the casing is fabricated with ABS M30i thermoplastic material (Fortus 400mc, Stratasys Ltd., USA).

5.2.3 Optical Imaging Systems

Different optical imaging systems can be used in conjunction with the proposed robotic scanning device as long as the fibre fits through the Φ2.7 mm diameter tube channel. The main imaging system used in the experiments is the in-house endomicroscopy system that allows high frame rate imaging and hence higher speed scanning (see Appendix A.1). Along with this system, the scanner has been used with an in-house dual wavelength slit scanning system and an in-house endocytoscopy system. Images and mosaics from all systems will be presented in the results section.

As discussed in previous chapters, due to the limited FoV of the microscopic images, it is beneficial to use a mosaicking algorithm to provide wide area histology-like images.
Here, the mosaicking algorithm described in Chapter 3 is preferred due to its computational speed, since it can run in real-time in conjunction with the high-speed optical system with 120 fps acquisition rate. This algorithm is a pairwise image registration approach which provides the shift between each pair of individual images based on the location of the cross-correlation peak. The algorithm is integrated with the Labview environment in the form of a dynamic link library (DLL), written in C++ using the openCV library [233].

5.2.4 Trajectory Generation

For the kinematic software of the instrument, different automatic scanning patterns, including linear, raster and spiral trajectories, can be used. As mentioned before, all the input signals have the form of voltage inputs and, therefore, all trajectories are in the “voltage” space rather than the Cartesian space. However, since it is a linear correspondence, a conversion factor can be determined by measuring and correlating the displacement in both spaces, as will be described in section 5.3.1. An additional requirement is that the generated points should result in a constant linear or tangential velocity. This en-
sures that there is consistent overlap between consecutive endomicroscopy images frames, at least when there are no severe tissue deformations.

To generate a linear scanning trajectory, the operator specifies the scan point frequency $f_s$, the linear velocity $u_s$ and the trajectory length $l_s$. Based on these inputs, every point $(x_i, y_i)$ is generated as:

$$\begin{align*}
(x_i, y_i) &= (x_c + \sum_{i=0}^{n} (\frac{i \cdot u_s}{f_s}), y_c) \\
(5.1)
\end{align*}$$

where $(x_c, y_c)$ is the initial centre point used as offset to the trajectory.

For a spiral trajectory, the Archimedean spiral (also known as arithmetic spiral) model was selected as it has the property that any line originating from the origin intersects successive turnings of the spiral at constant intervals of distance. Every point $(x_i, y_i)$ in the Archimedean spiral is described in Cartesian coordinates as:

$$\begin{align*}
\begin{cases}
x_i &= x_c + b \cdot \theta_i \cdot \cos(\theta_i) \\
y_i &= y_c + b \cdot \theta_i \cdot \sin(\theta_i)
\end{cases}
\end{align*}$$

where $(x_c, y_c)$ is the initial centre point used as offset to the trajectory, $b$ is the angular velocity of the spiral and $\theta_i$ is the corresponding angle in this specific time instance. The initial spiral is considered to start from $\theta_0 = 0$. Since the distance between two successive turnings $\Delta r$ is pre-defined as the desired overlap, then the angular velocity is:

$$b = \frac{\Delta r}{2\pi}$$

(5.3)

The maximum circular area to be scanned defines the final radius of the spiral trajectory $R_{max}$ and therefore the maximum angle at which the motion should end is:

$$\theta_{max} = \frac{R_{max}}{b}$$

(5.4)

The total length of the spiral trajectory can be defined as:

$$L_{max} = \frac{b}{2} \left[ \theta_{max} \sqrt{1 + \theta_{max}^2} + \log(\theta_{max} + \sqrt{1 + \theta_{max}^2}) \right]$$

(5.5)

which can be approximated as:
By defining the scan point frequency \( f_s \) and the linear velocity \( u_s \), the distance \( \Delta_i \) between successive points on the spiral can be defined as:

\[
\Delta_i = \frac{u_s}{f_s}
\]  

(5.7)

Since the sampling should be performed at regular distances across the length of the spiral and given that the distance between successive points on the spiral is pre-defined as well \( (\Delta_i) \), then at each time instance \( i \) the angle \( \theta_i \) is calculated as:

\[
\theta_i = \sqrt{\frac{2}{b} \left[ \frac{L_{\text{max}}/\Delta_i}{\sum_{i=0} \left(i \cdot \Delta_i\right)} \right]}
\]  

(5.8)

Then every point \((x_i, y_i)\) can be derived based on equation 5.2. Due to the fact that the approximated form of the spiral length is used for the derivation, in the beginning of the spiral the points are only approximately uniformly spaced.

### 5.2.5 Robot Closed-Loop Visual Servoing using Microscopic Images

The positional information that microscopic mosaicking provides (in the form of integrated image shifts) can be used for controlling robotic instruments. In particular, a model-free visual servoing control algorithm is used in conjunction with the robotic scanning instrument. As the robot scans a planned trajectory, the microscopic image shifts are used as the input to a closed loop proportional–integral–derivative (PID) control system. This method allows the robotic scanner to compensate for tissue deformation and low frequency external motion effects. For the robotic design considered here, the rotation of the probe with respect to the tissue surface is considered negligible and can be ignored. A similar approach was initially described in the work of Rosa et al. [216] but here is the first time that external motion effects are explored and that such algorithm is implemented in a high-speed spiral handheld scanner.

The concept of this control algorithm is to use the image shifts, in the 2D mosaic image plane, to control the XY positions of the instrument in the “voltage” space. As mentioned
before, all position inputs have the form of voltage inputs and, therefore, all trajectories are in the “voltage” space rather than the Cartesian space. The Z-component is defined manually by the operator when the instrument is placed normal to the tissue surface and does not change during the scanning task as the area covered is very small (see more in section 5.3.1). To align, therefore, the x-axis of the probe’s space with the x-axis of the image space (and hence also align the y-axes), an angle $\varphi$ should be calculated in order to find the matrix $R$ and rotate the XY position of the mosaic frame of reference to the XY position in the voltage frame of reference:

$$p_V = \begin{bmatrix} x_V \\ y_V \end{bmatrix} = L \cdot \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix} \cdot \begin{bmatrix} x_I \\ y_I \end{bmatrix} = L \cdot R \cdot \begin{bmatrix} x_I \\ y_I \end{bmatrix}$$

(5.9)

where $L$ is the factor that transforms image coordinates to voltages.

The angle $\varphi$ is calculated at the beginning of each experiment. The calculation of $\varphi$ is manually performed through an optimisation process where the scanner is commanded to follow a straight line and, using the real-time mosaicking output, the error between the commanded and the real coordinate points is minimised.

At the beginning of each scanning task, the probe position in the mosaic image is initialised and set as: $p_I(t = 0) = (0, 0)$. This position corresponds to the centre $p_{Ic} = (x_{Ic}, y_{Ic})$ of the mosaic image. As the scanning continues, the integral over time of the relative image shifts corresponds to the current probe position $p_V(t)$ at the time instance $t$. Consequently, at every time instance, the measured probe position $p_V(t)$ can be compared with $p_V^*(t')$, the desired probe position at time $t'$. Due to the different time sampling rate between the generated trajectory and the endomicroscopy images, the closest generated time instance $t'$ to the actual measurement $t$ is used as the desired probe position. To compensate for position differences, a PID based controller is implemented; here only the PI terms are used as the derivative term is prone to measurement and timing errors. In practise, the PI controller is sufficient for the proposed visual servoing algorithm as the P term eliminates the steady state error while the I term accelerates the movement towards the set-point faster and further eliminates the steady state error. Additionally, the scanning task is mainly quasi-static as the variables change slowly due to the velocity being
approximately constant [215].

The required correction to the probe trajectory in voltage space can be calculated as:

\[
\begin{bmatrix}
\Delta p_{Vx} \\
\Delta p_{Vy}
\end{bmatrix} = K_p \cdot \begin{bmatrix}
p_{Vx}^*(t') - p_{Vx}(t) \\
p_{Vy}^*(t') - p_{Vy}(t)
\end{bmatrix} + K_I \cdot \int_0^t \left( \begin{bmatrix}
p_{Vx}^*(\tau') - p_{Vx}(\tau) \\
p_{Vy}^*(\tau') - p_{Vy}(\tau)
\end{bmatrix} \right) d\tau
\] (5.10)

where the gains \(K_p\) and \(K_I\) are the proportional and integral terms respectively. The tuning of the gains is performed manually, using the standard method of eliminating the steady state error and minimising the overshoots. The selected PI terms are: \(K_p = 10\, \text{min}^{-1}\) and \(K_I = 0.4\, \text{min}^{-2}\).

A schematic of the control algorithm is depicted in Figure 5.6.

![Control schematic of the proposed system.](image)

Figure 5.6: Control schematic of the proposed system.

### 5.3 Experimental Results

#### 5.3.1 Mechanical Performance Evaluation and Characterisation

A potential issue with the cantilevered tube is that the motion of the tube tip is approximately planar which means that if the axial component (Z-component) difference is larger than the working distance of the optical imaging system, then optimal imaging results are not guaranteed. For this reason, a CAD simulation was performed in order to calculate the axial difference for various extreme positions of the workspace and the results are
Figure 5.7: Shaft length analysis: (left) demonstration of change in axial position at the extremes of the instrument’s workspace and (right) axial change versus shaft length.

depicted in Figure 5.7. As can be seen, the selected length of 58 mm provides less than 50 microns difference. In practice this is sufficiently small to be compensated for by the elasticity of the tissue.

Furthermore, the position of the motors and consequently the configuration of the cam-roller mechanism is optimised to achieve the largest possible workspace in the smallest possible overall dimensions for the actuation system. A CAD simulation of four different configurations has been performed, as can be seen in Figure 5.8. As a result, the configuration of 10 mm positional difference between the cam-roller mechanisms is chosen as it provides the largest simulated workspace (17.95 mm²).

5.3.2 Custom Tracking Rig

The novelty of the proposed mechanism is that it provides high accuracy and repeatability since no tendons or gears are involved. To validate this, a measuring system is needed that provides accurate tracking measurements in the range of several μm, approximately less than 10 times smaller than the effective FoV of the microscope probe. Commercial measuring systems such as electromagnetic and optical trackers provide no more than 100 μm accuracy which is only 2 times smaller than the FoV. Therefore, a novel custom
Figure 5.8: Analysis of the cam-roller mechanism positions: (left) four different position configurations explored during the analysis and (right) simulated workspace results for the four different configurations.

A tracking rig has been fabricated with tracking accuracy of several $\mu m$ that can effectively measure the instrument’s tip without imposing any additional weight on the instrument.

The proposed rig consists of a camera fixed rigidly along the axis of the scanning mechanism of the instrument, see Figure 5.9 and a lens assembly between the two in order to focus the tip of the probe onto the camera. An imaging probe inserted into the tube is used to relay light to the tip of the instrument, and the camera records $1280 \times 1024$ pixels video while the instrument scans. The output images have a dark background with a bright circle where the instrument’s tip is. Further processing is then needed to measure the circle’s centre and consequently the instrument’s relative tip position at any time.

After converting the output into a binary image, the tip of the instrument appears as a circle of 15 pixels diameter. Circle tracking, based on the two-stage Circular Hough Transform method [251], is used and the centres of the circles are recorded online (see Figure 5.9 for zoomed result). The overall resolution of the tracking system is measured to be 7.5 microns which is 32 times smaller than the microscopic image FoV (240 microns) and, therefore, adequate for use as a tracking method for the system.
Figure 5.9: Workspace analysis (a) Custom camera rig, (b) zoomed region of the tracked tip after binary conversion showing the tracked circle and its centre, (c) zoomed image presenting an exemplar tracked spiral trajectory and the instrument tip before binary conversion, and (d) overall workspace region with linear and non-linear parts.

5.3.3 Mechanical Performance Evaluation

By using the optimal mechanical design parameters and the previously mentioned tracking device, the overall workspace of the instrument can now be evaluated. By commanding the instrument to scan across the whole workspace, the linear and non-linear parts of the workspace can be found and the optimal workspace can be identified. As presented in Figure 5.9, in the upper and right edges of the workspace the levers of the mechanism lose contact from the cam, whereas in the lower and left edges there is a mechanical stop in the cam mechanism. The resulting middle linear workspace is approximately 14 $mm^2$ ($3.7 \times 3.7 \ mm$ area size). The resulting limits of the linear workspace are used to constrain the motion of the instrument so as to stay within this linear area.

By commanding the instrument to move in a straight line within the linear workspace, the scaling factor between voltage space and Cartesian space is calculated using the mosaic image that is generated. This is $664 \text{microns}/V$ for each direction. In order to find the correspondence between the motor rotations (in degrees) and the analog input voltages, the following motor settings are used: (a) voltage range = 20 V, (b) motor increments
Figure 5.10: Repeatability evaluation of two different scanning patterns along with the corresponding trajectory errors relative to the commanded trajectory, (a) three repeats of spiral trajectories, (b) four repeats of raster trajectories.

range (without gear transmission) = 100000, (c) motor increments range (without gear transmission) per revolution = 3000 and (d) gear ratio = 256. As a result, the following equation is followed to find the value of $R$ (degrees per volt):

$$R = 360 \cdot \frac{1}{256} \cdot \frac{1}{3000} \cdot \frac{100000}{20} = 2.3438 \frac{Degrees}{Volt} \quad (5.11)$$

Using a trial-error approach in a non-deforming lens paper target, the approximate maximum linear scanning velocity is found at $8V/s$ (or $5.312mm/s$) for the open loop approach and $2.5V/S$ or $(1.660mm/s)$ for the visual servoing approach. However, due to mosaicking inconsistencies in the extreme velocities (due to smaller overlaps between images) and the limited communication bandwidth of the data acquisition card, the velocity is chosen to be approximately to the values $2.6mm/s$ for the open loop control and $1mm/s$ for the closed loop control. These velocity values provided consistent good quality mosaics.
The circle tracking method is also used to evaluate the repeatability and accuracy of the instrument. Two different scanning approaches (spiral and raster) have been evaluated. In Figure 5.10 the different trajectory patterns are demonstrated; three repeats of spiral trajectories and four repeats of raster trajectories, using the same parameters across the repeats, are performed in sequence. The position errors relative to the original plan are presented below the trajectories, with the maximum position error not exceeding 30 microns and the mean error across all the trajectories being 10.7 microns. The latter implies that the instrument has similar repeatability and accuracy to the tracking system while providing a FoV that is approximately 22 times larger than that of the microscopy probe. This is the first time that a custom large area robotic scanning device demonstrated such performance in terms of accuracy and repeatability.

5.3.4 Microscopic Visual Servoing Evaluation

The validation and evaluation of the visual servoing approach for microscopic images was performed, initially, using a custom phantom that exhibits a pre-defined grid pattern (see Figure 5.11(a)). The phantom was printed by a laser printer on a sheet of paper and coated by a fluorescent marker, making it visible by the endomicroscopy probe. Every square in the grid pattern has a line thickness of 73 µm and width of 237 µm. The accuracy of the printed phantom was confirmed using a bench-top microscope, showing that a maximum error of ±6 µm and ±5 µm respectively could be introduced to the ground truth by the printer. The mosaic results are presented in Figure 5.11(b). The lines and width of the grid pattern are measured using the mosaic images to validate the visual servoing algorithm, as can be seen in table 5.1. A total of 22 sample measurements are made on each image, recording the line thicknesses and the square widths. It should be mentioned that the measurements are performed manually, leading to minor segmentation errors for the distance calculation. However, these errors are not comparable to the overall FoV of the scanning device and therefore do not affect the final result. Despite the high accuracy of the instrument, the open loop performance is not perfect in terms of the grid visualisation, even in the situation that the surface is non-deforming. This can be explained by various mechanical imperfections and especially by the fact that the orientation of the instrument
Figure 5.11: Visual servoing results on a grid pattern. (a) Grid along with a zoomed picture of the printed pattern, (b) a mosaic of approximately 1.4 mm in diameter without and with visual servoing control. A dashed line is shown for visual assessment of the faithful reconstruction of straight lines from the phantom. (c) The two spiral mosaics (orange tint) are overlaid manually on the printed pattern (grey) for visual comparison.
Table 5.1: Comparison between the bench-top microscope’s measurements and the generated mosaics with and without visual servoing.

cannot be exactly perpendicular to the scanning surface, leading to a minor drift in one direction.

![Open Loop Mosaic](image1)

![Visual Servoing Mosaic](image2)

200 µm

Figure 5.12: Visual servoing mosaic results when motion is introduced using the motorised translation stage, along with the graphs showing the current position (real position from the mosaic) and the desired (commanded) position. In open loop, the real position deviates from the desired trajectory whereas, with the visual servoing the real position follows the desired trajectory. The control trajectory shows the commanded output to the instrument. Pseudo-colour is applied to the mosaics for improved visualisation.

However, a significant difference in terms of open loop and closed loop control is observed when unexpected motions and deformations appear. In order to demonstrate and evaluate this, the rigid scanner was clamped and commanded to perform spiral scans on a moving surface. Initially, lens tissue paper stained with the topical fluorescent contrast agent acriflavine was placed on top of a motorised translation stage that was commanded to
move in a random motion along both the x-axis and the y-axis within a range of ±100 \( \mu m \) with velocity of 1 \( mm/s \).

As can be seen in Figure 5.12, the spiral pattern cannot be formed in the open loop approach. This is evident in the position graph on the right hand side which shows that the mosaic trajectory, which is approximately the real trajectory followed by the probe (allowing for any small deformations), deviates significantly from the commanded trajectory (whereas it would ideally coincide). Conversely, with the visual servoing algorithm, the effects of the motion are almost eliminated as the mosaic trajectory follows the desired spiral trajectory. Note that the control trajectory, which consists of the real output signals provided to the instrument, in the visual servoing result does not follow any specific model, demonstrating the importance of the model-free PID approach selected in this work. A random motion as an external disturbance when the probe is scanning a spiral over the grid pattern, described before in the validation, is presented in 5.13. Further \textit{ex vivo} results on lens paper and colonic tissue (also stained with acriflavine) are depicted in Figure 5.14. It is important to mention here that the colonic tissue results show not
Figure 5.14: Mosaic results without and with visual servoing on lens paper and colonic tissue while an external disturbance is applied. (a) Lens paper with disturbance along x-axis of $0.1 \, \text{mm/s}$, (b) colonic tissue with random motion disturbance of $0.1 \, \text{mm/s}$. Pseudo-colour is applied on the lens paper mosaics.

only the suppression of the external motion but also compensation for tissue deformation. This, as a result, tends to increase the effective FoV for deformable surfaces, as can be seen in Figure 5.15 where the diameter of the area covered with visual servoing is $1.1 \, \text{mm}$, whereas without it is $0.94 \, \text{mm}$.

Figure 5.15: Mosaic results for breast tissue scanning without and with visual servoing showing the difference in the scanning area covered.
5.3.5 Mosaicking Results

The performance of the robotic scanning instrument was further assessed with *ex vivo* tissue experiments (see Figure 5.16). Since BCS surgery is the main application of this instrument, human breast tissues (normal and tumour) were scanned with the instrument while clamped. Clamping was used as handheld operation introduced motion artefacts due to hand motions, creating issues with large area mosaicking. The primary goal of these experiments was to scan, with high-speed, specific areas of interest to assess their histological features. Therefore, tissue areas of approximately 1 mm (Figure 5.16(a,e-f)) and 1.7 mm (Figure 5.16(b-d)) outer diameter were covered in 13 and 7 seconds respectively. The time difference here is because a slower speed was selected in the scan in Figure 5.16(a-c). Additional experiments on porcine colonic tissue were performed to assess the performance of the instrument on other types of tissue (Figure 5.16(e-f)).

The mosaics presented illustrate the performance of the instrument while rigidly clamped in order to demonstrate the enhanced scanning capabilities of the proposed work. This is the main reason that a robotic arm with cooperative-control function is included in this system. The operator can manipulate the arm in a hands-on manner, place it in the desired position and then press the button to scan the area under investigation. On the other hand, the operator has the option to unmount the instrument from the robotic arm and operate it in a handheld manner. Exemplar mosaicking results comparing use of the instrument handheld and clamped in the robotic arm are demonstrated in Figure 5.17. High resolution microscopic scanning of lens paper areas of 2 mm and 3.7 mm diameter is shown.

As mentioned before, the instrument features a tubular channel that different fibre-based optical imaging probes can be inserted into. To demonstrate use of the instrument with other endomicroscopes, an in-house dual wavelength slit scanning confocal system and an endocytoscopy system were tested. Exemplar *ex vivo* mosaic results for each system are presented in Figure 5.18. The human breast tissue scanned with the endocytoscopy system is stained topically with the contrast agent methylene blue. The lens paper mosaic using the dual-wavelength system was acquired by staining locally one part of the paper with the contrast agent acriflavine and the remaining part with methylene blue. Scanning
Figure 5.16: Example of *ex vivo* tissue mosaics. (a) Human normal breast tissue (fibrous connective tissue), (b) human normal breast tissue (adipose tissue admixed with fibrous connective tissue), (c-d) human breast invasive carcinoma, (e) porcine large bowel and (f) porcine large bowel fat.
Handheld Operation  
Clamped Operation

Figure 5.17: Example mosaics of lens paper while manipulating the instrument in a handheld manner or clamped on the robotic arm. Pseudo-colour is applied for improved contrast.

Endocytoscopy  
Dual-Wavelength Confocal Microscopy

Figure 5.18: Example ex vivo mosaic results using the in-house endocytoscopy with human normal breast tissue and the dual-wavelength slit scanning confocal system with lens paper.
Figure 5.19: (a) Tip of the instrument presenting the $CO_2$ laser and microscopy fibres, (b) instance from the laser ablation procedure on a paper card when the laser fired, (c) image from the custom tracking camera presenting the mechanical offset between the $CO_2$ fibre (red) and the microscopy fibre (blue), (d) two instances overlaid, from different timestamps, presenting the movement by the offset value in order the red dot can coincide with the initial position of the blue dot, and (e) real-time mosaic, from the same region of a paper card, before and after the ablation in the centre of the spiral pattern.

was performed over the edge between the two parts in order to visualise both contrast agents and hence mosaicking of two wavelengths.

### 5.3.6 Energy Delivery Results

Finally, for the first time in robotic-assisted endomicroscopy, preliminary results showing how a fibre-delivered $CO_2$ laser ablation system can be integrated into the scanning device are presented. The concept for the use of a laser ablation system is that the operator can assess the tissue microscopically and, if cancer cells are identified or tumour margins, then, with the laser, can mark or cauterise the examined area in real-time. The laser (OmniGuide Inc., USA) was used in “Super Pulse” mode (single pulse, power 3 $W$ and duration 40 $ms$) in order to deliver a small precise mark on the target. The laser fibre is inserted through the same channel as the microscopy fibre bundle and is separated by
Figure 5.20: Example of a real-time mosaic image from the same region of the grid pattern before and after the laser ablation. The grid pattern was then imaged through a bench-top microscope, presenting an ablation of 104 µm in diameter and thermal damage of approximately 210 µm.

Before Ablation  After Ablation  Bench-top Microscope

a small horizontal mechanical offset at the tip of the instrument (see Figure 5.19(a-c)). Also, at the tip of the instrument, the laser fibre is fixed in an vertical offset of 1 – 2 mm from the tip of the microscopy probe due to its working range (see Figure 5.19(b)). Since the two fibres are laterally offset, in order to be able to ablate at the centre of the microscopy mosaic, an offset is applied to the probe’s position during CO₂ laser firing. Determining this offset requires that the mechanical offset should be accurately measured (see Figure 5.19(c)). For this, the custom tracking rig is used and the two fibres are tracked in real-time. Since using the CO₂ laser would damage the camera, this is replaced with a multi-mode fibre coupled to a red LED for the purpose of the calibration. The horizontal offset is found in the voltage space by commanding the tip to move so that the red circle (CO₂ fibre) coincides with the initial position of the blue circle (microscopy fibre) (see Figure 5.19(d))

To validate the previous approach, a laser pulse is shot at the centre of an area that has been mosaicked using a spiral pattern and then it is mosaicked again. Figure 5.19(e) shows the area before and after the ablation. The laser mark, as expected, is at the centre of the mosaic. The same process is followed using the grid pattern, previously described, in Figure 5.20. The laser delivered a very short pulse, ablating an area of 104 µm in diameter with minimal thermal damage (< 50 µm radially), as can be seen in the bench-top microscope’s image. Here, also, the real-time mosaicking does not provide as accurate
pairwise registration as previous mosaicking examples due to the very small correlation
of the images in the ablated area. However, the results demonstrate the potential for the
combination of microscopy with a laser ablation fibre in the same reference frame to assist
with intraoperative tumour marking and ablation.

5.4 Discussions and Conclusions

In this chapter, a novel, robotic, high-speed scanning device for intraoperative tissue iden-
tification and margin assessment is demonstrated. It is first evaluated for use in BCS where
intraoperative assessment of tumour margins is critical, but the clinical applications can
be extended to other open surgical procedures (e.g. skin cancer or neurosurgery), where
a direct line of sight may be available to the tissue of interest. Since flexibility is not
required in the proposed system, high resolution endomicroscopy probes can be utilised,
while generating fast large area mosaics through its highly accurate and repeatable mech-
anism. This is an advancement over current histology techniques, such as frozen section,
and other point-based optical imaging modalities as it both significantly reduces the time
taken to acquire results, and provides high resolution, histology-like images over an area
of 3 mm$^2$ in less than 10 s without the removal or destruction of any tissue structure.
Also, the small footprint of the overall system in the operating theatre, either as a hand-
held instrument or combined with the small robotic arm, allows it to be easily combined
with other imaging modalities such as indocyanine green (ICG) fluorescence and hyper-
spectral imaging, therefore creating a platform for intraoperative image guidance and
therapy based on real-time histological information.

In this work, a different approach was pursued compared to previous works, including
that of Zuo et al. [4, 218] and Rosa et al. [212, 216], because the focus of this device is to
provide the surgeon with a tool for rapid assessment of tissue microstructure. In previous
works, the duration of the scanning task of similar sized areas is close to one minute,
a fact which can create hurdles in the adoption of this technology as the surgeon is re-
quired to stay steady during the task. The combination of the high frame rate microscopy
system with the accurate/repeatable high-speed scanning mechanism and the supporting
robotic arm can significantly alleviate these issues. Combined with interventional capabilities, shown for the first time in this work, this tool could not only provide real-time histological information but also mark and ablate suspicious areas and tumour margins intraoperatively, without the need for a separate tool.

However, there are still some limitations that need to be tackled in future versions of the tool. As mentioned before, the rigidity of the instrument allows high resolution fibres to be inserted, but it prevents its use at operating sites that require flexible access. This is particularly problematic due to the nature of endomicroscopy probes that require the probe to be as normal to the surface as possible in order to get good quality microscopic images. Additionally, handheld operation of the instrument is still challenging as the contact forces during the scanning task are not compensated. This is significantly improved with the use of the supporting robotic arm’s hands-on approach to place and stabilise the instrument, but this adds significantly to the cost and complexity of the system. Handheld or mounted use could be improved by incorporating an active or passive control mechanism to control the tip of the instrument on irregular surfaces. This force control mechanism can be either passive (such as the pneumatic approach discussed in Chapter 4) or active (using a force sensor on the instrument or the robotic arm) or rigid (using an outer stabilisation tube).
Chapter 6

Intraoperative Robotically-Assisted Transanal Microscopic Imaging with Large Area Scanning

6.1 Introduction

In this chapter, a novel articulated robotic assisted instrument is proposed to facilitate image mosaicking for endoluminal endoscopic microsurgeries. Although radical surgery for cancer is generally oncologically effective, endoluminal excision offers a much less invasive alternative. In particular, for colorectal cancer, there is a considerable clinical demand for local excision of rectal lesions as it is the fourth most commonly diagnosed malignancy in UK and the second most common in Europe and third worldwide, as of 2012 [10]. One third of these are staged as ‘T1’ on diagnosis, meaning that they are potentially suitable for local excision.

Local excision by endoscopic microsurgery is increasingly used in minimally invasive endoluminal procedures. Transanal Endoscopic Microsurgery (TEM) is an advanced local excision procedure which has gained more attention recently due to the associated positive clinical outcomes. This operation is established for early stage rectal cancers (pT1), with
Figure 6.1: (left) Transanal Endoscopic Microsurgery procedure, (centre) endoscopic images provide no pathological information that can assist in tumour margin identification and (right) the Micro-IGES robotic system.

a low recurrence rate of $4 - 10\%$ [252]. The TEM procedure currently relies on preoperative endoscopic biopsies for tissue characterisation and MRI or transanal US for staging malignancies [253]. There is no intraoperative image guidance used during surgery and the resection margins are assessed post-operatively (see Figure 6.1(centre)). This leads to high recurrence rates ($> 20\%$) associated with inadequate staging, unclear resection margins and incompletely excised adenomas [252, 254]. Additionally, the procedure suffers from poor ergonomics and lack of predictable control of standard, straight laparoscopic instruments due to collisions between the instruments within the confined luminal space. This prevents the reliable excision of lesions and limits the use of the procedure to a small number of surgeons in specialist centres (see Figure 6.1(left)). All these limitations suggest that a real-time, \textit{in situ} and robotically assisted approach for identifying the tumour margins and confirming the total resection of the cancer intraoperatively could improve patient outcomes.

As mentioned previously, endomicroscopy can provide a potential solution to tissue assessment as it is a real-time, high resolution imaging technique for \textit{in situ} visualisation of tissue micro-structure. Confocal laser endomicroscopy, for example, has demonstrated high accuracy in distinguishing between neoplastic and non-neoplastic colorectal tissue
[255], and so has potential applications in the real-time guidance of endoscopic microsurgery. Also, recently, a study has shown that slightly lower resolution HRME can distinguish between neoplastic and non-neoplastic colorectal polyps with an accuracy of 94% [256].

However, due to the small-size of optical probes, imaging inside luminal structures can be greatly facilitated by the stability and accuracy of robotic instruments. For this reason, a 6-DoF robotically actuated scanning instrument (“Micro-IGES endomicroscopy instrument”) has been developed that can perform large area mosaics (over 30 mm$^2$) in a master-slave configuration. The instrument is developed in parallel with a robotic system, called Micro-IGES system, that is intended to be used for robotic assisted TEM procedures (see Figure 6.1(right)).

Compared with previously reported devices, the instrument allows larger areas of tissue to be scanned in a short period of time, offering improved flexibility and a master-slave control scheme. The key features of the tool are its simple mechatronics and the ability to scan over large curved endoluminal surfaces. It has a small diameter shaft which means that it can be inserted through a standard surgical port, facilitating a wide range of applications in cancer surgery. It can, also, be used in conjunction with standard laparoscopic instruments or other robotic instruments. Results from ex vivo colorectal tissues demonstrate efficient large area scanning with ease of user control.

6.2 Materials and Methods

6.2.1 System Overview

6.2.1.1 Mechanical Design

An overview of the key characteristics of the Micro-IGES endomicroscopy instrument is shown in Figure 6.2. It is a 6-DoF robotic instrument, comprising a prismatic joint and five rotational joints. The prismatic joint and a rotational joint (roll) are located at the proximal end of the instrument, while distally there are four joints (three active and one
Figure 6.2: Key characteristics of the Micro-IGES robotic assisted microscopy instrument.

passive). The design of the instrument is driven by the environment constraints imposed by endoluminal structures and the imaging constraints, as the microscopy probe’s orientation should be normal to the scanning surface. For this reason, the motion of the second and third joints is coupled in order to constrain the second and fourth links always to be parallel (see more information in the kinematics section 6.2.2). In addition, the prismatic joint is formed by a rack and pinion mechanism, where the rack is integrated in the rail casing of the overall Micro-IGES system and the pinion is part of the endomicroscopy instrument (see Figure 6.3(c)).

The actuation of the instrument is tendon-based with the exception of the prismatic joint that is gear-based (0.4mod). The tendons are routed through the steel shaft of the instrument through a low-friction, pulley system to capstans which are rotated by the motors (see Figure 6.3(a)). All the tendon-routing and fixations of the front articulated end-effector of the instrument are presented in Figure 6.4. The last distal joint is driven by two tendons that are routed through two low-friction bowden cables for backlash reduction (see Figure 6.3(b)). This means that the distal joint is independent of the rest of the instrument’s configuration and therefore independent of backlash issues originating from mechanical imperfections in the rest of the links. To reduce backlash to the rest of the distal links a pulley mechanism is integrated in the fabricated links in order to eliminate tendon length changes even in acute angle configurations. Regarding the middle coupled
mechanism (4-5 DoF), as can be seen in Figure 6.4, a tendon is routed from a fixed point on the 4th link to the 2nd, where it is also fixed, and then back to 4th where the tendon is fixed again. This is the same as having two tendons with equal lengths fixed in the 2nd and 4th link through the 3rd link to create a parallelogram mechanism (see Figure 6.4(2)). The actuation of this mechanism is performed through a pair of tendons that is fixed at the 3rd link (see Figure 6.4(1)).

All actuation is driven by five brushless DC servomotors (0620C-006-B, Faulhaber, Germany) with integrated 256 : 1 ratio gearboxes for three of them, and 64 : 1 for the remainder, custom-made to have the gear head and the encoder head in a parallel structure to reduce size. Each DC servomotor exhibits 0.36 $mNm$ torque without the gearbox.
Figure 6.4: Tendon-routing of the articulated end-effector. Dashed lines depict that the tendons are in different planes and circles in the end of the tendons are the fixation points.

The instrument features a hollow channel of 2.1 mm diameter in the smallest region where different imaging probes can be interchanged. The channel is connected by a hollow, flexible tube from the back of the actuation pack to the proximal end of the instrument for easier fibre probe interchangeability (see Figure 6.3(d)). Each imaging probe is routed through the channel and fixed distally by a screw-based mechanism. Here a leached fibre bundle (SCHOTT, UK) is used. The fibre bundle contains approximately 17,000 fibre cores. The spacing between the fibre cores is 8 µm. The total diameter of the bundle is 1.3 mm and the length is 2 m. This results in a 16 µm resolution and a 800 µm FoV.

6.2.1.2 Instrument Dimensions

The shaft and the joint mechanism have an 8 mm outer diameter which can fit into one of the instrument channels of conventional transanal endoscopic surgical ports (typically instrument insertion ports have 10 mm diameter openings). The overall length of the instrument is 283 mm which enables the surgeon to image tumours that are deep inside the rectum, enlarging the effective imaging workspace. All actuation mechanisms and motors are located within the same compartment with overall dimensions of 90 mm length, 54 mm height and 44 mm width. The dimensions follow the overall design constraints of the Micro-IGES system so that multiple instruments can operate simultaneously for improved surgical workflow (see Figure 6.2 and 6.3(d)).
6.2.1.3 Materials

The distal articulated links and motor couplings are fabricated with stainless steel through the Direct Metal Laser Sintering process (Concept Laser Mlab Selective Laser Sintering (SLS), ES-Technology, UK) for improved durability and accuracy, while the instrument casing and several of the body parts such as the pulley holders and wire fixations are fabricated with a fused deposition modelling prototyping process with ABS M30i thermoplastic material (Fortus 400mc, Stratasys Ltd., USA).

6.2.1.4 Electronics

The motor encoders use analogue hall sensors (K2280, Faulhaber, Germany) to sense the tendon positions and, as a result, the position and orientation of the tool. Every motor is connected to a motor controller (MCBL 3002 S CF, Faulhaber, Germany), with the latter providing power to drive the motor. To further reduce the dimensions of the instruments and maintain the compatibility with the Micro-IGES system, the motor controllers are placed distally to the actuation pack and enclosed in a custom casing for plug-and-play connectivity (see Figure 6.5). The motor controllers are connected to a 6 V power source with a current limit of 1 A in order to eliminate the possibility to damage the controllers, the motors and the gearboxes. The embedded software of the motor controller communicates with the kinematic control software over a CAN bus (Faulhaber CANopen interface).
with an update rate of 100 $Hz$ and a 1000 $kbps$ Baudrate. To reduce reflections of the signal on the CAN bus due to the use of multiple devices, a termination resistance of 120 $\Omega$ is used. To interface the CAN bus with the control system, a USB-to-CAN interface is used (IXXAT, HMS Industrial Networks, Germany) that can be connected to any USB 2.0 Hi-Speed port (480 $MBit/s$) of the control system.

### 6.2.1.5 Software

The overall motor and control software was developed in the Labview Environment (National Instruments, US). The software features two control modes for the instrument, so that it can either perform autonomous scanning with different patterns (line and raster) for large area surveillance or be controlled in a master-slave manner with the use of an input device (here a Phantom® Omni (Sensible Technologies, USA) is used). All the system’s interconnected components described in this section are depicted in Figure 6.6.
The control software can be separated in four components: the interface code for the CAN communication, the GUI with the live 3D visualisation of the instrument, the Master-Slave Interface and control, and the autonomous scanning interface.

The program for the CAN communication is based on the Faulhaber CANopen protocol, already integrated within the motor controllers, while dynamic-link libraries (DLLs) are used for interfacing with the protocol for fast and reliable communication. Each of the motor controllers corresponds to a CAN node and is assigned a particular node ID. To improve the message bandwidth, a “Trace” function has been implemented that, after a “sync” command, answers in one message with the position and current of each motor. Depending on the node ID, the position value is multiplied with the gear ratio value and the Current value (Amps) is checked to confirm that it is in a safe range. If not, then the motion is stopped. The same result occurs when position limits are enabled and the commanded position exceeds this limit.

The GUI (see Figure 6.7) includes all the interfaces with the motor controllers as well as a 3D graphical representation of the distal joints of the instrument and the proximal
rail position. Specifically, on the left of the interface, all high level controls of the instrument can be found, such as homing, enable, disable, stop motion, position limits enabler, velocity control and relative/absolute position control. As well as the 3D visual control scene on the right, in the lower part of the GUI all the values of the joint positions and the currents are presented in the form of graphs for troubleshooting purposes. The GUI could be made more minimal by adding a “clinical” mode where the surgeon sees only the controls and indications critical to the operation.

The other two components will be described in detail in the following sections.

### 6.2.2 Kinematic Model

The instrument, as already mentioned, includes a prismatic joint and a roll joint proximally, and four joints in the distal part. All of the instrument’s DoFs are depicted in Figure 6.8(a). Due to the coupled motion of the 4th and 5th joint (see Figure 6.8(b)), angles \( p_1 \) and \( p_2 \) are equal (\( p_1 = p_2 \)).

Based on the aforementioned, the instrument kinematics can be derived. In general, the kinematics of a robotic manipulator describe the relationship between the joint space \( (q \in \mathbb{R}^{n \times 1}) \) and the task space \( (x \in \mathbb{R}^{6 \times 1}) \), where \( n \) is the number of DoF. The Denavit – Hartenberg (DH) are traditionally used to described the kinematics of serial chain robotic manipulators.

The DH parameters are presented in Appendix D, whereas the corresponding joint limits are presented in Table 6.1. Regarding the latter, it should be noted that the translation range is defined as the motion range inside the rail of the Micro-IGES system. Also, the roll joint can rotate more than 360 degrees but for safety reasons the range is limited to 360 degrees.
The kinematic equations of the developed instrument of 6 links with joint parameters $q_i$ are given by:

$$T_0^6 = \prod_{i=1}^{6} T_{i-1}^i(q_i)$$  \hspace{1cm} (6.1)$$

where $T_{i-1}^i(q_i)$ is the homogeneous transformation matrix of the link frame $i$ with respect to the link frame $i - 1$. Therefore, the matrix $T_0^6$ specifies the location of the $i$th coordinate frame with respect to the base coordinate system, which is located at the intersection between the $q_3$ axis and the roll axis $q_2$ (see Figure 6.8(a)).

Using the standard DH notation and parameters, the previous transformation matrix can
be calculated as:

\[
T_{i-1}^i = \begin{bmatrix}
    c\theta_i & -s\theta_i & c\alpha_i & a_i s\theta_i \\
    s\theta_i & c\theta_i & c\alpha_i & a_i c\theta_i \\
    0 & 0 & 0 & 1 \\
\end{bmatrix}
\]  

(6.2)

Here, for simplicity, \(\cos(t) = ct\) and \(\sin(t) = st\).

Therefore, to find the end-effector’s position and orientation, a chain product of successive coordinate transformation matrices of \(T_{i-1}^i\) should be calculated as:

\[
T_6^0 = T_1^0 T_2^1 T_3^2 T_4^3 T_5^4 T_6^5 = \begin{bmatrix}
    R_0^6 & P_0^6 \\
    0 & 1 \\
\end{bmatrix}
\]  

(6.3)

where \(R_0^6\) is the corresponding rotation matrix and \(P_0^6\) is the corresponding position matrix, calculated as:

\[
R_0^6 = \begin{bmatrix}
    s\theta_2 s\theta_6 + c\theta_2 c\theta_3 c\theta_6 & c\theta_2 s\theta_3 & c\theta_2 c\theta_3 s\theta_6 - s\theta_2 c\theta_6 \\
    -c\theta_2 s\theta_6 + s\theta_2 c\theta_3 c\theta_6 & s\theta_2 s\theta_3 & s\theta_2 c\theta_3 s\theta_6 + c\theta_2 c\theta_6 \\
    s\theta_3 c\theta_6 & -c\theta_3 & s\theta_3 s\theta_6 \\
\end{bmatrix}
\]  

(6.4)

\[
P_0^6 = \begin{bmatrix}
    12.55 c \theta_2 c \theta_3 + 7s \theta_2 s \theta_4 + 11.3s \theta_2 s \theta_6 + 7c \theta_2 c \theta_3 c \theta_4 + 11.3c \theta_2 c \theta_3 c \theta_6 \\
    12.55s \theta_2 c \theta_3 - 7c \theta_2 s \theta_4 - 11.3c \theta_2 s \theta_6 + 7s \theta_2 c \theta_3 c \theta_4 + 11.3s \theta_2 c \theta_3 c \theta_6 \\
    d_1 + 12.55 s \theta_3 + 7s \theta_3 c \theta_4 + 11.3s \theta_3 c \theta_6 \\
\end{bmatrix}
\]  

(6.5)

In these equations, \(\theta_5 = -\theta_4\) due to the parallelogram mechanism. As a result, as can be seen in equations 6.4 and 6.5, the parallelogram’s angle \(\theta_4\) affects only the final position and not the orientation. This explains how this mechanism can be used to control the distance between the end-effector and the tissue.

Regarding the inverse kinematics of the instrument, the Jacobian matrix can be calculated. The Jacobian represents the relationship between the task space and the joint space, as it
relates the end-effector with the joint velocities \((\dot{x}_i, \dot{q}_i)\):

\[
\dot{x}_i = J \dot{q}_i \iff \begin{bmatrix} \dot{p} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} J_p(q_i) \\ J_\omega(q_i) \end{bmatrix} \dot{q}_i \tag{6.6}
\]

where the Jacobian components are calculated as:

\[
J_p(q_i) = \frac{\partial P^6_0}{\partial q_i} \tag{6.7}
\]

\[
J_\omega(q_i) = \begin{bmatrix} S_{3,2} \\ S_{1,3} \\ S_{2,1} \end{bmatrix}, \quad \text{where } S^{3x3} = \dot{R}^6_0 (R^6_0)^T \tag{6.8}
\]

### 6.2.3 Master-Slave Control

To improve the ergonomics of the tool and add the ability to be remotely controlled, a master-slave control is implemented. This allows the addition of a computerized kinematic medium between the instrument and the operator and therefore the addition of motion scaling for precision scanning. In this work, the master device is a Phantom® Omni (Sensable Technologies, USA). It provides six DoF manipulation and positional sensing using joint proprioceptive sensors (optical encoders). It also includes a gripping stylus in the end-effector with two buttons that will be used to activate different modes in the proposed instrument. The communication is performed through a IEEE-1394 FireWire port.

<table>
<thead>
<tr>
<th>Button 1 Enabled</th>
<th>Button 2 Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instrument</strong></td>
<td><strong>Omni</strong></td>
</tr>
<tr>
<td>Joint 1 (translation)</td>
<td>Z Position</td>
</tr>
<tr>
<td>Joint 2 (roll)</td>
<td>Joint 6</td>
</tr>
<tr>
<td>Joint 3</td>
<td>Joint 1</td>
</tr>
</tbody>
</table>

Table 6.2: Master-slave control, presenting how the joints and positions of the master device are mapped to drive each of the robotic instrument’s joints.
Based on the kinematics of the master device and its mechanical joint configuration, a master-slave control presented in Table 6.2 is used to operate the instrument. In particular, the configurations shown in Table 6.2 present the mapping between individual joints of the tool and the kinematic parameters of the Omni device (either position or joint parameters). The control of the instrument is performed by clutching / un-clutching using one of the two buttons included in the master device depending on the joint that needs to be activated. The mapping is a linear function of each of the components presented in Table 6.2, e.g. the translation of the instrument is derived by multiplying a scaling factor (found experimentally to fit the instrument’s ergonomics and tasks) with Z position of the master device. Similar process is followed for the rest of the instrument’s joints.

Communication with the master device is implemented using a DLL in C++ based on the manufacture-provided API and has a 200 Hz update rate.

### 6.3 Experimental Results

#### 6.3.1 Workspace Results

The functionality of the robot is first assessed by calculating its workspace. Using the generated forward kinematics equations (eq. 6.5) and the DH parameters (see Appendix D), the workspace of the instrument can be evaluated. In Figure 6.9, the instrument’s workspace is depicted. The origin of the world coordinate frame \((0,0,0)\) is equal to the insertion point of the instrument from the Micro-IGES surgical port; the insertion point is at +10.5 mm offset towards the +X axis. The instrument is inserted in the lower port of the surgical port because the operational site in TEM is in the lower quarter of the port; if the tumour is identified in the side or upper quarters of the rectal wall then the overall robotic system can be rotated in order to place the operational site in the lower quarter of the port. In Figure 6.9, the environment’s workspace requirements are also overlaid onto the tool’s workspace in the form of two cylinders (minimum workspace with 4 cm diameter and maximum workspace with 5 cm diameter). As is demonstrated, the tool can reach more than the whole quarter of the rectal wall.
Figure 6.9: Illustrations of the instrument’s workspace in comparison to the available endoluminal workspace for a TEM operation. The latter is depicted in the left figure, from a side view, as a cylinder of 50 mm diameter and in the right figure, from the front view, as two cylinders of 40 mm and 50 mm diameter. In both figures the instrument is displaced in an offset position equal to its insertion point from the surgical port.

In order to represent the dexterity of the instrument, a manipulability measure is calculated in the overall workspace. The manipulability index was introduced by Yoshikawa [257] and is defined as:

$$\omega = \sqrt{\det(JJ^T)}$$

where $J$ is the Jacobian matrix calculated in equations 6.7 and 6.8.

The measure is affected by the singular values of the workspace and increases away from singularities. For our calculations, this measure was enhanced with the penalisation term $P(q_i)$ introduced by Tsai et al. [258] that additionally considers the behaviour of the tool at the joint limits (being near the joint limits is penalised). If the distance to the lower joint limit is $q_i^-$ and to the upper limit is $q_i^+$, then the penalisation term $P(q_i)$ is defined as:

$$P(q_i) = 1 - \exp\left(-k \prod_{i=1}^{6} \frac{(q_i - q_i^-)(q_i^+ - q_i)}{(q_i^+ - q_i^-)^2}\right)$$

where $k$ is a scaling factor (here $k = 10$) that is used to adjust the behaviour of the tool in the joint limits. The final value of the dexterity measure is calculated by multiplying the
Figure 6.10: Workspace dexterity evaluation; red: high dexterity, blue: low dexterity.

manipulability index \( \omega \) with the penalisation term \( P(q_i) \).

### 6.3.2 Tendon Backlash

In tendon-actuated robotic instruments, the performance is often limited by non-linear friction and the backlash hysteresis phenomena (due to tendon slack), posing challenges in precise control of the system. Since no sensors are integrated on the tip, this phenomena should be modelled and quantified.

Using an electromagnetic (EM) tracking system (Trackstar 2, Ascension Tech. Inc.), the motion of each individual joint of the instrument is measured. Two mini 2 – \( \text{mm} \) outer diameter EM sensors with 6 DoF tracking (model 180, Ascension Tech. Inc.) are placed inside the placement slots of two rapid prototyped extensions. This design of the extensions was preferred as it amplifies the distances between the sensors and therefore more accurate measurements can be performed. The experimental setup is presented in Figure 6.11(a). The calibration data is recorded while each joint was commanded to move...
in two different encoder values, while the commanding encoder values and the respective EM values are recorded.

The calibration data recorded for each joint are grouped into forward $g^+ = \{e^+, q^+\}$ and backward direction $g^- = \{e^-, q^-\}$ based on the joint velocity $\dot{q}$ (positive and negative respectively) and points with low velocity are removed from the grouping selection. As can be seen in Figure 6.11 and 6.12, after the data selection, two linear lines are fitted. The lines are determined based on an optimisation process that considers that two parallel lines can be fitted simultaneously to the two data sets:

$$\{\hat{s}, \hat{b}^-, \hat{b}^+\} = \arg\min_{\{s, b^-, b^+\}} \left\{ \sum_{i=1}^{n} (q_i^--s(m_i^- - b^-))^2 + \sum_{i=1}^{m} (q_i^++s(m_i^+ - b^+))^2 \right\}$$

(6.11)

where $n, m$ are the number of $g^-$ and $g^+$ points respectively. Here $s$ represents the slope of the parallel lines and is constrained to be identical for both lines. The backlash can then be determined by the distance between two points as: $b = |b^+ - b^-|$. 

As can be seen from the graphs, joint 1 (translation) has a very repeatable behaviour with linear slopes but large backlash. This is mainly due to the gear transmission’s backlash.
which is linear and therefore the difference between the two slopes can be considered
directly to eliminate the backlash. Also, the roll (joint 2) presents very low backlash
which is mainly due to the placement of the mechanism that is located proximally, at the
actuation pack. Joint 6, the most distal joint, presents also very linear behaviour and this
is significantly affected by the incorporation of bowden cables that eliminate any changes
in the length of the tendon and therefore backlash. Joints 3 and 4 present more non-
linear behaviour due to the friction between the joints and, as a result, they can be better
modelled using more advanced backlash elimination techniques, such as in [259, 260].
6.3.3 Exemplar Mosaic Results

The performance of the 6-DoF robotic scanning instrument was further assessed with *ex vivo* experiments. Initially, lens paper was stained with topical fluorescent contrast agent acriflavine, and the instrument performed an open loop automatic raster scan. This operation was selected to demonstrate the ability of the scanner to automatically scan a very wide area requiring investigation. Figure 6.13 presents the preliminary mosaic results of lens paper, showing the fibrous structure of the paper. The area covered in the left image is approximately $88 \text{ mm}^2$ and in the right image approximately $28 \text{ mm}^2$.

In addition, raster scans over a grid pattern phantom, the same as described in the previous chapter, on a flat surface are presented. As described, the phantom was printed by a laser printer on a sheet of paper and coated by a fluorescent marker. Every square in the grid pattern has a line thickness of $73 \mu m$ and width of $237 \mu m$. In Figure 6.14, the mosaicking results over the overall grid’s surface are presented and can be compared to the mosaicking result of the high-speed scanner in Figure 5.11. The comparison depicts the large area scanning capabilities of the presented robotic instrument as well as its accuracy and precision during automatic raster scanning.
Since the target clinical application for this instrument is the TEM operation, it was also tested using the WOLF TEM simulator and colonic tissue. This TEM simulator was used as it is used for surgical training for this specific operation. Similarly to lens paper, topical contrast agent acriflavine was applied before the scanning task. The instrument was then commanded to perform a raster scan trajectory, demonstrating that is capable of large area mosaicking with \textit{ex vivo} porcine colonic tissue.

Further \textit{ex vivo} experiments were performed after the incorporation of the master-slave control, presenting the ability of the instrument to scan large tissue areas under user control. For these experiments, a curved artificial phantom was used (see Figure 6.16(b)) in order to simulate the endoluminal structures encountered during TEM operations. On top of the phantom, porcine colonic tissue was placed, stained also with topical contrast agent acriflavine. The overall setup, with the user's console, the robotic instrument and the microscopy system, is presented in Figure 6.16(a). Due to the lack of the endoscopic view, the user was controlling the instrument under direct visual control and also using the image feedback from the real-time endomicroscopy GUI.
Figure 6.15: (a) Experimental setup for ex vivo experiments using the TEM simulator and porcine colon tissue; the instrument inside the Micro-IGES rail is depicted and in the background the clinical graphical interface shows the real-time microscopic images and mosaics, (b) images from the experiments demonstrating insertion of the instrument using the master-slave control, assuming perpendicular orientation to the tissue and then scanning automatically once the tip is in place, and (c) images from an external camera showing the real-time mosaics and images from the clinical interface.

Figure 6.16: (a) Experimental setup using a curved artificial phantom and the master-slave control and (b) zoomed view of the curved phantom with the colonic tissue placed on top.
Two different modes were explored in these *ex vivo* experiments. In the first mode, the user had to initially position the instrument in the starting point of the scanning task and then automatically the instrument would scan the surface using the joint pairs either 1-2 (translation and roll) or 1-3 (translation and sweep joint). In the second mode, the user would scan an area under direct visual feedback from the microscopy GUI and the real-time microscopy image. The mosaicking results are presented in Figure 6.17.

![Figure 6.17: (a) Automatic raster scanning mosaicking results (first mode) and (b) real-time mosaicking results under completely manual control using the master device.](image)

### 6.4 Discussions and Conclusions

In this chapter, an intraoperative robotically-assisted transanal microscopic imaging device for large area scanning is provided. For TEM procedures, complete excisions of cancerous tissue can not always be confirmed until histological examination of the excised tissue is completed. The proposed instrument clearly demonstrates a potential solution to the clinical requirement of intraoperative tissue identification during endoscopic micro-surgical procedures where margin assessment is critical for the outcome of the procedure.
Compared to the devices presented in the previous chapters, and also other robotically-assisted endomicroscopy works, this device presents a high degree of flexibility even for tubular structures as encountered in endoluminal surgeries, while providing very large area cellular level images through mosaicking. The results demonstrate 2D mosaics covering areas more than 80 $mm^2$. The overall scanning duration is sufficiently short to avoid the interruption of normal surgical workflow and it offers real-time tissue characterisation. Also, the overall dimensions of the instrument do not hinder the operator from inserting other instruments while scanning. The proposed master-slave control operation allows for the surgeon to be located remotely from the patient, improving the overall ergonomics of advanced procedures, such as TEM.

The proposed parallel-constrained articulated mechanism in combination with the last joint allows for the tip of the instrument to bend up to 90 degrees and in the same time the contact height can be regulated. This is essential as the operator can adjust the contact force by mainly moving only one joint. Force could also be controlled by an alternative approach in future versions of the instrument. For example, a pneumatic approach, similar to Chapter 4, could be incorporated using the channel inside the instrument. As an alternative to a mechanical approach, image-based control can be incorporated for ensuring optimal contact with the tissue and good quality microscopy images.
Chapter 7

Robotic-Controlled Intraoperative Multi-Modality Imaging Fusion ¹, ²

7.1 Introduction

In this chapter, the integration of endomicroscopy with ultrasound imaging through a robotically actuated instrument for TEM operations will be presented. As discussed in previous chapters, a single imaging modality does not always provide sufficient information to identify the tumour margins. Multiple modalities can be used to provide a more detailed assessment of the tissue. One such example is described by Yang et al., where US, OCT and photoacoustic tomography are integrated into a single instrument, providing information on a similar plane at different depths and resolutions [262]. Alternatively, modalities that provide contrasting information may be combined, such as in the work of Higgins et al. [263] that proposed the combination of fluorescence microscopy and OCT to acquire

¹Content from this chapter was published as:

²This work has been conducted in collaboration with Mr George Dwyer, during his master thesis project, with equal contributions from the authors. The author of this thesis created the initial prototype for combining US with endomicroscopy, the endomicroscopy mosaicking software, the calibration model and algorithm for the image fusion and processed all the data from the section “mechanical design evaluation”. George Dwyer designed and fabricated the articulated mechanism, the actuation pack and performed the final experiments. The ultrasound 3D reconstruction software is written by Dr Pratt [261].
images on planes perpendicular to one another, allowing a three-dimensional assessment of the tissue. A similar concept has been proposed in Chapter 4 with the multi-modality pick-up probe using OCT and endomicroscopy. The disadvantage of OCT, however, is the maximum imaging depth within scattering tissue is only $l - 2 \text{ mm}$.

An alternative imaging modality, currently used in the operating theatre for intraoperative imaging, is ultrasound. Ultrasound provides a tomographic image, similarly to OCT, but with higher penetration depth (more than 1 \text{ cm}) at the expense of lower resolution.

Endomicroscopy and US provide diagnostically complementary information - endomicroscopy provides cellular level information at a superficial level whereas US provides depth resolved information at a macroscopic level. Combination of these modalities could assist the assessment of surgical margins, with endomicroscopy used for high resolution lateral margin determination, and the overall three-dimensional (3D) structure of the cancerous tissue inferred from the US images. In transanal surgery, which is the clinical focus of this device, the generated 3D structure can also provide the thickness of the rectum under the operating site, assisting in preventing perforation of the lumen. For intraoperative use this information should be represented in a concise and intuitive way, with images from each modality co-registered over the target area of interest. The main factor impeding this is the difference in the FoV of the two modalities.

The fusion of multiple imaging modalities can be achieved only by registering both image sets in a common frame of reference. Two options provide this capability: using an external tracking system (optical or EM tracking) or using a mechanical registration mechanism. The first option requires additional hardware which can potentially prove expensive and complex. Mechanical registration, which is the preferred solution in this work, is a simpler approach as both modalities are physically connected and the registration model can be derived through the CAD model of the mechanism.

The remainder of this chapter introduces a novel articulated robotic instrument for the fusion of endomicroscopy and ultrasound to provide intraoperative assessment of surgical margins. Robotics provide the ability to increase the dexterity and stability of the instrumentation, while decreasing the complexity of instrument manipulation. The device
incorporates an articulated 1-DoF robotic segment to increase the dexterous workspace and a scanning 1-DoF mechanism, incorporating passive adaptation, that ensures consistent large area endomicroscopy mosaics. In addition, the use of endoscopic tracking is demonstrated, allowing three-dimensional reconstruction of the ultrasound data displayed onto the endoscopic view. An *ex vivo* study on porcine colon tissue has been performed, demonstrating the clinical applicability of the instrument.

## 7.2 Materials and Methods

### 7.2.1 System Architecture

#### 7.2.1.1 Mechanical Design

The robotically actuated instrument is based on a two DoF mechanism, one for articulation and one for endomicroscopy scanning. Here, a commercial US probe (UST-533,
Hitachi-Aloka Medical Ltd, Japan) has been incorporated as it features a linear array and is small in size. This US probe has a transducer length of approximately 10 mm, whereas the endomicroscopy probe used has a FoV of approximately 1 mm (leached fibre bundle, SCHOTT, UK). In order to produce comparable images, the endomicroscopy probe must be translated across the length of the US transducer. To accommodate this, an external frame has been designed in order to mount the US probe (Figure 7.1(a)). This frame features two metal rails mounted above the US transducer that provides the necessary stabilisation structure to support and constrain the translation mechanism (Figure 7.1(c)). The mechanism is then driven using a pair of tendons positioned in alignment with each rail. A miniature pulley is positioned at the head of the probe to redirect the tendon in the direction of the shaft. The tendon pair is then aligned through a small aperture, leading to the instrument shaft.

High quality endomicroscopy images depend on maintaining a low contact force between the endomicroscopy probe and the tissue. This is difficult to achieve manually, causing the tissue to be out of focus or deformed. Here a simple solution is adopted due to size constraints. In particular, a passive force adaptive mechanism (Figure 7.1(c)) using two stainless steel compression springs is developed. This mechanism varies the offset between the endomicroscopy probe tip and the US array surface passively, accommodating the force exerted on the probe’s tip. As only one DoF is required, a brass rod runs concentrically through each of the springs. This constraint assists with stabilising the movement of the probe, but it also introduces additional friction to the mechanism. In addition, the endomicroscopy probe is fixed to the lower component of the mechanism using two grub screws.

To obtain clinically useful information, the instrument needs to be manipulated accurately through the surgical workspace. Current instruments are mainly manipulated manually (e.g. rigid laparoscopic instruments or endoscopes) and the operation’s outcome is mainly dependant on the surgeon’s dexterity. By taking this into account, an active DoF articulation module, featuring two passive joints, has been fabricated to provide lateral movement to the instrument. This capability allows large surface areas to be scanned; by synchronising the motion of the linear endomicroscopy scanning and the lateral motion
of the articulated segment, large raster microscopic mosaics can be generated along with their respective cross-sectional US images.

Due to the large variation in the diameter of the US probe, a segment with a channel would be very large as it would need to accommodate the US transducer. Therefore, in order to constrain just the US probe wire, a two part segment system was designed. This allows the flexible probe to be assembled directly onto the US probe wire, circumventing the need of using large end diameter (Figure 7.1(d)). This lateral DoF provides the necessary channels for both imaging modalities and the tendons for actuation. Due to its rigidity, the US probe wire is positioned centrally in order to minimise its effect on the motion of the robotic probe. Also, a minimum bending radius of 10 mm is provisioned in the design requirements due to the mechanical constraints of the US probe wire. The designed instrument, due to the dimensional constraints of the US transducer, has an overall diameter of less than 15 mm, allowing the instrument to be introduced through a standard 15 mm laparoscopic port.

The final prototype was fabricated using additive manufacturing techniques. In particular, the actuation pack is made using fused deposition modelling prototyping process with ABS M30i thermoplastic material (Fortus 400mc, Stratasys Ltd., USA) and the in-
instrument using Veroblock digital material (Objet EDEN350TM, Objet Geometries Ltd., USA).

### 7.2.1.2 Electronics

The tendon pairs allow the actuation system to be placed proximally (see Figure 7.3). In the actuation side, the tendons are routed through a pulley system and anchored in two capstan mechanisms accordingly for each pair. The capstans are then used to tension the tendons so that backlash is eliminated. The capstans are clamped directly to metal shafts which rotate in axial alignment with the motors; the shafts and the motors are joined using custom motor coupling components. For this instrument, due to there being no size restrictions, the servo motors chosen for actuation are the Dynamixel RX-24F (Robotis Inc., USA), featuring an integrated potentiometer position sensor which provides a resolution of 0.29° and allows loads up to 2.6 Nm. The maximum load allows the US probe to be manipulated even at extreme angles. Communication with the motors is performed through a RS-485 multi-drop bus which allows communication speeds of up to 1 Mbps. The communication from each controller is performed through a Universal Asynchronous Receiver/Transmitter (UART) that converts the signal from the main controller to RS-485. Each packet is defined by the main controller, which is in this case a computer running the Labview development environment (National Instruments) with the Dynamixel Application Programming Interface. The API provides functions to control the communications, define and interpret each packet, transmit instruction packets and receive status packets. Using the integrated position sensors, the motor controllers are able to perform compliant control. With the compliance control method, the output torque of the motor can be adjusted in relation to the difference in the current position and the goal position. The overall framework described in this section with all the interconnected components is depicted in Figure 7.2.

In order to calculate the kinematic model of the instrument, the Denavit-Hartenberg (DH) parameters of the probe were calculated (see Appendix E). Each joint, including passive joints, are assigned a coordinate frame. This resulted in a five coordinate frame transform-
Figure 7.3: Actuation pack assembly (left) CAD rendering along with individual mechanical components and (right) fabricated actuation pack.

...there are two passive and one active revolute joints at the lateral segment and one prismatic at the scanning mechanism. Workspace analysis was conducted to address the operational requirements.

### 7.2.2 Endomicroscopy Imaging

The optical imaging system is a widefield fluorescence endomicroscope (see Appendix A.2). Here a leached fibre bundle (SCHOTT, UK) is used to relay the images from the tip of the instrument to the system due to its flexibility as it can bend 90 degrees with less than 1 mm bending radius. This is preferable for the scanning component of the instrument as it does not affect its motion and keeps the overall size to a minimum. The total diameter of the bundle is 1.3 mm and the length is 0.8 m. The effective FoV of the endomicroscopy image is 800 µm.

In this study, all raw images were processed to remove the honeycomb pattern and broken fibres, as in Chapter 3. All the mosaicking results were generated offline but, similarly to all the previously described instruments, the system currently includes real-time mosaicking capabilities. Image blending based on distance weighting was also incorporated in
order to reduce intensity variations between the individual images in the mosaic.

### 7.2.3 Ultrasound Imaging

The UST-533 US probe incorporates a linear array transducer, approximately 10\textit{mm} in length, with FoV of approximately 60\textit{mm}. The array contains 128 elements, producing a lateral resolution of approximately 50\textit{µm}. The probe is connected to a Prosound Alpha10 cart (Hitachi-Aloka Medical Ltd, Japan) which powers the probe and interprets the received signal. The US imaging cart has a DVI output which is connected to a DVI to SDI converter (Matrox Electronic Systems Ltd, Canada). The SDI image is captured by a Quadro SDI Card (NVIDIA Corporation, USA).

In order to extend the capabilities of the instrument, a marker-based tracking approach was also implemented based on the work by Pratt \textit{et al.} [261]. Tracking an US probe allows the captured image data to be coupled to spatial information. Therefore by scanning the US probe over an area, a large dataset can be acquired over a spatially known volume. The instrument is tracked by attaching a custom marker (Figure 7.2 - tracking marker) using either monocular or stereoscopic based algorithms; the instrument is then tracked in a coordinate system fixed with respect to the endoscope position, without introducing additional equipment to the operating theatre. The custom marker is a KeyDot (Key-Surgical Ltd, USA), which features non-symmetrical chessboard patterns. The marker is firstly calibrated to the US image, which provides the spatial transformation between the marker on the US transducer and the US image. This transformation is calculated using a calibration phantom of known dimensions. The spatial features of the phantom are then registered to the image features from the transducer. The spatial transformation calculated is comprised of image scaling followed by translation and rotation over six DoF. For accurate tracking of the US probe, the endoscope must also be calibrated. This was achieved by following a method described by Zhang [231]. From this, the probe can be tracked in real time; a video sync generator (Blackmagic Design Ltd) allows the US image data and video from which the probe tracking data is derived to be synchronised.

As described in the work of Pratt \textit{et al.} [261], the 3D reconstructed volume can be gen-
erated if sequential US images are associated with corresponding spatial information. Using the marker tracking, the four coordinate positions for each corner of the US image together with the image size are stored. A bounding volume is then made using the maxima and minima of each coordinate axis. For each voxel in the bounding volume, the perpendicular distance to each of the US image planes is calculated. The closest is used to define the voxel intensity, as long as the distance is below a certain predefined threshold. There is no interpolation between slices, and the algorithm relies on the US images being relatively close together. The reconstructed volume can then be rendered volumetrically, as the spatial location of the volume is known in relation to the endoscope and can be overlaid onto the endoscope display.

### 7.2.4 Fusion

The scanning mechanism integrated at the instrument’s end-effector incorporates the ultrasound transducer and in a defined mechanical offset (Figure 7.4(a)), which is outside of the ultrasound’s FoV but parallel to the ultrasound imaging plane, the endomicroscope probe is scanned to create linear mosaics. To achieve the fusion of the two modalities, the images must first be registered spatially, which requires both modalities to be calibrated relatively to each other. From the mechanical design (CAD) of the tool, the relative parallel (horizontal) offset between the microscopy probe and the US image can be calculated (Figure 7.4(a) - mechanical offset). However, the CAD does not provide accurate information about the vertical offset between the two images, *i.e.* there is no mechanical information about the accurate position inside the US probe of the US linear array that produces the image (see Figure 7.4(a) - spatial alignment offsets).

Therefore, a calibration phantom was designed and fabricated, featuring two aligned cubic volumes (Figure 7.4(b-c)). The first contains a hollow cubic volume, filled with US gel as a conductive medium; whereas, the second volume contains stained lens tissue paper supported by a fixed rigid structure. The US probe is fixed in place above the hollow volume, while the endomicroscopy module is unconstrained. As a result, the endomicroscopy probe can be translated across the entire scanning mechanism limits, imaging the lens pa-
per section and the known volume edges. Simultaneously the spatially identical volume edges are presented in the US images. Using these known volume dimensions, the resultant endomicroscopy mosaic and US image can be registered. As it can be understood, in order for the ultrasound images to appear in the same plane as the endomicroscopy, the probe needs to be moved up to the mechanical offset position, and vice versa for the endomicroscopy mosaic image.

7.3 Experimental Results

7.3.1 Mechanical Design Evaluation

The instrument configuration was initially analysed in Matlab, using the calculated DH parameters and a maximum rotation from the origin of ±22 degrees. Through simulation, the lateral range of the instrument was calculated as 33.66 mm with each segment reaching the mechanical limit. The fabricated instrument was then analysed using the endoscopic tracking; it was positioned with the marker perpendicular to the endoscope. The instrument was then manipulated to both tracking limits, −38 degrees and 55 degrees, proximally and distally to the endoscope. The endoscopic tracking then allowed assessment of the movement repeatability. The motion used for the assessment was from
Correlation between encoder angles and argument of measured tool-tip points ($R=0.989$)

Figure 7.5: Performance evaluation by endoscopic tracking; (a) Tracking information for repeated rotational trajectories (blue points), with fitted circle (red) for correlating the angle of the shaft to the encoder angle values; (b) the correlation between the position encoder sensing (blue) and the measured rotation angle (green) by the endoscope.

the proximal tracking limits to the distal limits and then returning to the initial position. This was repeated ten times while recording the joint positions from the motor encoders. The measured angle of rotation (green line) correlates well with the rotational motor encoder values (blue line), with $R = 0.989$ (Figure 7.5(b)). This shows that it is possible to utilise the encoder values to determine the tip-position of the probe, despite mechanical imperfections and the long shaft.

7.3.2 Endomicroscopy Scanning of a Large Surface Area

Linear scans were performed using the instrument with the scanning velocity set to $0.7 \text{ mm/s}$ which was the highest velocity possible while producing the appropriate image quality. Lateral steps were also included in each linear scan, producing a large 2D microscopy data set. Each lateral step was 1.35 degrees, over a range of 27 degrees and 16 $\text{mm}$ lateral range; the angular velocity of the lateral joints was 2.85 $\text{degrees/s}$ for each step. This experiment was then repeated using the passive force adaptive mechanism over the same parameters. Both sets of data were then mosaicked with blending, resulting in mosaic images of approximately $10 \text{ mm} \times 13 \text{ mm}$ (fixed probe clamp) and $9 \text{ mm} \times 13 \text{ mm}$ (passive force adaptive mechanism)(Figure 7.6). In this study, lens cleaning paper was used as
a surface and topical contrast agent acriflavine was applied to provide the fluorescence signal.

### 7.3.3 Calibration of Endomicroscopy and Ultrasound

By fixing the instrument to the calibration phantom, as shown previously in Figure 7.4, the spatial transformation between the microscopy and the US image can be found. The endomicroscope is scanned across the calibration phantom and the images are mosaicked, producing a mosaic with a length of 12.6 mm. Each image has clearly defined areas with spatially known dimensions. With this knowledge, the images can be scaled to a corresponding size. The mosaic can then be superimposed above the US image and aligned to show the relative position of the two datasets in 3D space (see Figure 7.7). The process of superimposing the two images is facilitated by the fact that the ultrasound image presents the volume limits of the calibration phantom. The same limits appear in the microscopy mosaic (limit where the lens paper appears - see Figure 7.7(b)). These limits allow the spatial alignment offsets between the two image planes to be found (see Figure 7.4(a) –
Figure 7.7: Calibration image showing the alignment between the mosaic and US image, and the spatial offset of 4.6 mm between the centre of the images.

spatial alignment offsets). Following this process, fusion of the two imaging modalities in the same reference frame is feasible.

7.3.4 Ultrasound Imaging

To assess the ability of the instrument to perform 3D reconstruction, a kidney phantom with an enclosed tumour was scanned (see Figure 7.8(a)). The US scans for 3D reconstruction were captured and the lateral DoF joint angles were assigned to keep the marker within the endoscope’s FoV at all times. This resulted in a scan over 20 degrees, before returning to the origin, at an angular velocity of 2.85 degrees/s. This produced a data set of 89 slices, each with four sets of tracked corner coordinates. The data set was then reconstructed into a basic volume (see Figure 7.8(b)). The software then provides different transfer functions that can be applied to improve the visualisation before it is overlaid onto the endoscopic view, according to the operator’s preference (see Figure 7.8(c)). Each transfer function defines the opacity and colour of the voxels based on image pixel values, allowing different areas of tissue to be identified visually.
Figure 7.8: 3D reconstruction of kidney phantom, (a) experimental setup along with the 3D reconstructed overlay of the artificial tumour, (b) basic 3D reconstruction of the US dataset and (c) 3D reconstruction using advanced transfer function for visualisation.

Figure 7.9: Ex vivo imaging test with porcine colon tissue experimental setup, (a) overview of the instrument during scanning, (b) details of the endoscopic tracking using a stereo endoscope, and (c) artificial phantom to simulate a potential tumorous structure to assess the US 3D reconstruction.

7.3.5 Ex Vivo Test with Porcine Colon Tissue

In order to demonstrate that the instrument is clinically viable, an ex vivo imaging study was undertaken on porcine colon tissue using the lateral instrument with the passive force adaptive mechanism (see Figure 7.9). The main difference from previous experiments was the use of a semicircular phantom, featuring a diameter of 50 mm (Figure 7.9(c)), the approximate diameter of a human rectum. The phantom also provided a point of reference for the US data. The endomicroscopy data and US images were captured simultaneously using the endoscopic tracking. The instrument was placed in contact with the tissue enclosed in the phantom with the endoscope positioned perpendicular to the instrument for optimal viewing of the KeyDot marker (Figure 7.10(b)).
Figure 7.10: *Ex vivo* imaging of porcine colon tissue imaging study, (a) instance from the endoscopic view (operator’s view) presenting in real-time the microscopic images on the right and an overlay of the real-time US image on the US probe, (b) 3D reconstruction result of the US images presenting the details of the tissue but also the artificial tumour structure below (c) 4mm × 12mm area mosaic showing surface detail of the colon including the crypts.

The scans were taken over 27 degrees with 1.35 degree-steps. In each step, an US image was captured. The scanning mechanism velocity was 0.7mm/s and the angular velocity of the lateral joints was 2.85degrees/s, resulting in a 4mm × 12mm area mosaic and an accompanying 3D reconstruction of the US data (Figure 7.10(b-c))

### 7.4 Discussions and Conclusions

In this chapter, the proposed instrument presents an alternative approach to assess the tumour margins intraoperatively. Here, as well as surface information from high resolution microscopic images, depth information from US images would allow assessment of the tumour’s 3D structure. This framework permits, therefore, the margins to be confirmed *in situ* and *in vivo*. Finally, endoscopic tracking assists the operator to track the biopsy site during the assessment process, and is used to reconstruct the US images into a 3D volume, providing an overall 3D map of the area under investigation.

The workspace analysis of the proposed instrument shows that the two passive joints provide a workspace greater than that required for surgery within the rectum. However, this allows the instrument to be operated within its mechanical limits, reducing the torque required for actuation and reducing the stress on US probe. The comparison between the
fixed probe and the passive force adaptive mechanism demonstrated that the fixed probe produced larger spaces between linear scans and less consistent mosaics (Figure 7.6(a and b)).

The *ex vivo* results presented in this chapter demonstrated the potential clinical use of the instrument intraoperatively. The US images obtained were consistent throughout the scanning process (Figure 7.10(a)), however as the tissue imaged was healthy tissue with no abnormal pathology, the functionality of the US and 3D reconstruction was not fully exhibited. The endomicroscopy data illustrate the ability to scan and produce mosaics, \(4\,mm \times 7\,mm\) in size, with almost no visible deformation and clearly visible crypts. The average size of the malignancies resected in TEM procedures is \(31\,mm\), with a recommended margin of \(10\,mm\) [253]. The current size of raster scans precludes the entire malignancy from being imaged; however the region of real interest is the tumour margin, which can be scanned. Increasing the scan area further can lead to inconsistencies in the produced mosaic. These inconsistencies are possibly due to the constant bidirectional loading and unloading of the tissue while scanning. This causes little visible deformation until the strain of the tissue exceeds the force applied by the endomicroscopy probe, resulting in visible deformation of the tissue. Alternatively, errors in the mosaicking algorithm may be predicting an incorrect shift between sequential frames.

The proposed image fusion framework demonstrates the feasibility of how two complementary imaging modalities can be used effectively intraoperatively. The proposed platform can be extended also to include larger mosaics with the 3D tissue structures reconstructed using the US.
Chapter 8

Conclusions and Future Research Directions

8.1 Contributions of this Thesis

This thesis explores robotically-assisted manipulation of endomicroscopy probes for real-time optical biopsy during surgery. Improved iterations of prototypes have been demonstrated, coupled with real-time high-resolution image mosaicking and advanced scanning, allowing intraoperative pathological analysis and surgical planning. The development and evaluation of these devices led to a number of technical, clinical and scientific contributions:

- Development of multi-purpose pick-up probes to be used with existing surgical robotic instruments. Demonstration of improved handling of flexible miniaturised microscopy probes while providing integrated tracking and multi-modality imaging for biopsy and re-targeting.

- Development of a novel low-friction air bearing to be used as an axial force adaptive mechanism, maintaining constant contact between the tissue and the imaging probe. The principle is presented in the form of a pick-up probe to be used with existing
surgical robotic platforms that lack haptic capabilities. Demonstration of improved mosaics and consistent tissue contact across multiple users.

- Development of a high-speed handheld or dockable robotic scanning device for breast conserving surgery that allows precise scanning over wide, deformable, moving tissue areas with image-guided control. Demonstration of high resolution, high-speed mosaicking over normal and cancerous breast tissue.

- Implementation of an energy delivery fibre for targeting microscopic regions of tissue for marking and ablation. Demonstration of mosaicking and targeting of a specific position in the instrument’s workspace.

- Development of a dexterous 5-DoF robotic instrument for use in endoluminal microsurgeries. Demonstration of large area mosaicking over complex curved endoluminal structures in a short period of time, offering improved flexibility and a master-slave control scheme.

- Development of a robotic articulated instrument for multi-modality image fusion (ultrasound and endomicroscopy). Demonstration of the ability to calibrate both modalities in the same frame of reference and to generate large 3D maps of the tissue using surface, cellular-level information and cross-sectional, structural information.

The technical chapters of the thesis address some of the unmet clinical challenges outlined in the introduction of this thesis regarding the applications of endomicroscopy for intraoperative tissue characterisation and tumour margin delineation. Transforming conventional microscopy modalities, which are generally bench-top systems, into probe-based microscopes is hindered by the difficulties of manual control. This is amplified by the fact that the FoV of optical biopsy techniques, as seen in Chapter 2, is limited and not comparable to histology images. To obtain, therefore, meaningful information, the surgeon would need to perform tedious movements while performing large area surveillance. The difficulty increases further when scanning inside cavities (e.g. breast cancer) or over complex 3D surfaces. These challenges are addressed in this thesis from both software and hardware perspectives.
From the software perspective, as described in Chapter 3, the images generated from the scanning motions are used by an image processing framework to generate real-time, *in vivo* mosaic images and provide histology-like characteristics over a wide area of tissue. The algorithms are integrated with a high frame-rate line scanning endomicroscopy system, that can achieve 120 fps, without compromising image qualities. It has also been demonstrated how these mosaics can be fused intraoperatively with the macroscopic view of the surgeon and provide multi-scale integration for improved re-targeting of the obtained biopsy. All these imaging approaches can be integrated with the developed robotic and non-robotic instruments and improve the prospects for intraoperative margin evaluation by providing the surgeon with much more comprehensive information.

Following software integration, the rest of the thesis focuses on the hardware perspective. Currently, probes are almost exclusively controlled manually by the clinician, leading to inconsistent velocities and poorly controlled trajectories. Throughout this thesis, the introduction of surgical robots and robotic-assisted probe manipulation has been presented to improve the probe manipulation compared to manual instruments. In Chapter 4, a first approach towards robotic-assisted probe manipulation using existing surgical robotic platforms is demonstrated in the form of multi-purpose, pick-up probes. Three different robotic pick-up probes are presented with varying functionalities and capabilities; (1) a single modality pick-up probe with integrated tracking for position estimation and multi-scale fusion; (2) a multi-modality pick-up probe for OCT and endomicroscopy imaging simultaneously, which also integrates endoscopic tracking and can regulate the tissue distance based on the OCT image; and (3) a force adaptive pick-up probe based on an ultra low-friction air bearing mechanism that provides adaptive axial force control. A user study, using the force adaptive pick-up probe combined with the high-frame rate endomicroscopy system, demonstrates the effectiveness of this concept for improved large area mosaicking and tissue surveillance. The simplicity and versatility of the pick-up probes allow them to be easily deployed in the surgical workflow and be used either with existing robotic systems or manual laparoscopic surgical instruments.

In addition to existing surgical systems, the development of bespoke, mechatronically-enhanced robotic devices for endomicroscopy is demonstrated. In Chapter 5, a high-speed
handheld robotic scanning device is presented with the aim to provide rapid assessment of tissue microstructure during breast conserving surgery. The instrument includes a highly accurate and repeatable scanning mechanism, providing high resolution, histology-like images over an area of 10 mm$^2$ in less than 10 s. The instrument is combined with a high frame rate line scanning system and can be easily adapted to other imaging modalities. A visual servoing approach based on microscopic images is integrated to improve the accuracy of mosaics, especially on deformable, moving tissue surfaces. The interchangeability of fibres bundles, also, allows the integration of fibre-based CO$_2$ laser ablation. The combination of microscopy with a laser ablation fibre is presented for intraoperative marking of tumour margins.

The lack of flexibility of the previous device is addressed with the dexterous 5-degree-of-freedom robotic instrument presented in Chapter 6. The device was developed in parallel with a robotic system that is intended to be used for robotic assisted TEM procedures, a challenging procedure that lacks intraoperative image guidance. The instrument offers increased flexibility even for complex endoluminal structures and the small dimensions of the device allow it to be inserted through a standard surgical port and be used in conjunction with standard laparoscopic instruments or robotic instruments. The instrument features, distally, a novel parallel mechanism that facilitates the operator to approach complex tubular structures. By using a master-slave control scheme and automatic scanning capabilities, large area mosaics are generated over curved endoluminal surfaces efficiently.

The concept of multi-modality imaging, that was introduced with the pick-up probe in Chapter 4, was further explored with a bespoke articulated robotic instrument in Chapter 7, that integrates endomicroscopy and ultrasound imaging. The latter provides cross sectional depth information at a macroscopic level and endomicroscopy cellular level surface information. A novel calibration method is presented in order to fuse, in the same frame of reference, diagnostically complementary information and create 3D maps of the area under investigation. Robotic actuation facilitates the acquisition of consistent microscopic images as well as providing the necessary articulation to scan large area surfaces. Also, the use of endoscopic tracking is demonstrated, allowing 3D reconstructions of the ultra-
sound data to be displayed onto the endoscopic view, enhancing the intraoperative image guidance. Results on *ex vivo* tissue and an artificial curved phantom presented the clinical applicability of the instrument.

## 8.2 Future Research Directions

The main achievement of the work presented in this thesis is the development of robotic-assisted endomicroscopy platforms that can be used to investigate further software and hardware improvements to ensure the usability and efficacy of endomicroscopy in surgery.

Regarding online image processing, in Chapter 3, a pairwise registration approach is used in order to provide high frame rate mosaicking capabilities in the proposed robotic devices. However, this approach can introduce cumulative mosaicking errors, especially in continuous large area surveillance and manual probe manipulation, that hinders accurate tissue assessment by the clinician. Therefore, the work presented can be improved with post-processing mosaicking algorithms by taking into account non-rigid deformations and performing global registration [105] or incorporating the position information provided by robotic devices or tracking systems [115, 118, 211].

Improvements of individual microscopic images is also possible. With the development of miniaturised piezoelectric actuators or other micro-positioning techniques, microscopy probes with super-resolution can be developed that use very small displacements (in the order of several microns) to acquire multiple images and fuse them to create a high resolution microscopy image. High frame rate microscopy systems coupled with super-resolution probes could significantly improve the diagnostic accuracy of microscopy.

A limitation of the current probe-based confocal endomicroscopy systems is that, due to the inherent optical sectioning, only a single, fixed focal plane is provided. By enabling rapid axial scanning, *in vivo* 3D cellular level imaging can be performed. Methods such as liquid lenses and MEMS are potential solutions to enable this capability [97, 98, 96, 99, 100]. The integration of this technology with robotic actuation could
significantly improve the clinical value of endomicroscopy for *in vivo* tissue surveillance. Improvements in the diagnostic yield of endomicroscopy may also be possible with multi-wavelength systems, allowing use of multiple contrast agents.

Due to the direct probe-tissue contact that is required in cellular level imaging modalities such as confocal microscopy, consistent microscopic images can be acquired only if the instrument moves smoothly across highly deformable tissues. The advantages of force sensing in endomicroscopy scanning are demonstrated in Chapter 4 with the force adaptive pick-up probe and in Chapter 7 with the passive spring-based mechanism. However, the other robotic devices do not incorporate any sensing capabilities and are limited to the operator’s visual feedback in order to assess the forces that are applied to the tissue surface. This can be problematic for complex, curved tissue surfaces. In these cases, force sensing is needed, either by active or passive means, for maintaining consistent contact with the tissue. This forges mechanical stabilising mechanisms, which can deform the tissue and cause difficulties in re-targeting. The miniaturisation of force sensors, such as fibre Bragg gratings (FBG) and strain gauges, could benefit the development of multi-DoF force sensing for miniaturised pick-up probes and flexible robotic instruments. An alternative approach, partially described in this Chapter 4, is the use of OCT or other similar imaging modalities, to maintain a consistent distance from the tissue surface. All these force sensing approaches, however, add further components to the instruments that increase their dimensions and complexity. A recent approach, from the Hamlyn Centre, towards “sensor-less” probe-tissue contact during endomicroscopy scanning has great potential. The approach processes the incoming video stream of endomicroscopy images and quantifies the image quality based on various image criteria; the outcome of the processing is fed into a reinforcement learning algorithm that controls the distance from the tissue.

A first approach towards flexible robotically-assisted surgical devices for endomicroscopy is presented in Chapters 6 and 7 and clearly demonstrates the advantages of robotic-assisted scanning over manual scanning. The development of single-port, highly articulated or flexible, endoscope-based robotic platforms with robotically actuated segments, as discussed in Chapter 2, demonstrate improved versatility and reachability over manual
endoscopic tools. This is essential to demonstrate the advantages of optical imaging technologies and eliminate all the ergonomic challenges. Also, the small field-of-view of microscopy images makes them susceptible to tremor and involuntary movements by the surgeons, suggesting there would be a benefit from a tremor suppression mechanism. This should be accompanied by a haptic feedback control loop that enables the surgeon to feel or palpate the tissues under investigation.

Another area that is proposed in this thesis and can be investigated further is multimodality and multi-scale imaging. High resolution imaging probes provide micron-scale details in a local area, while techniques, such as hyper-spectral, FLIM and ultrasound imaging, cover macroscopic details on a much coarser scale. Cross-scale imaging is therefore a good candidate for precision guidance and targeting intraoperatively. This can be greatly enhanced by robotic systems, allowing stable operation and local-global registration (re-targeting). Additionally, the integration of all these intraoperative imaging probes can be enhanced by the incorporation of laser ablation and micro-diathermy techniques, allowing precision surgery with point-by-point targeting capabilities on the fly. This is introduced for the first time in endomicroscopy in Chapter 5 and further investigation should be performed.

Further user trials and usability studies would be beneficial for the robotic toolbox presented in this thesis as their clinical uptake will then be evaluated. Even though most of the proposed devices provide automatic endomicroscopy scanning, and have been designed based on their easy integration in the normal surgical workflow, further ergonomic studies will provide a better insight on how these instruments can be better interfaced by the surgeons. For example, the high-speed scanning instrument, presented in Chapter 5, is designed to be used in a handheld fashion or attached to a robotic arm, providing versatility in the way it is used by the surgeon. However, the preferred mode is strongly dependant on the operator. Also, further user studies will provide a better understanding on the clinical usefulness and efficacy of the visualisation of multi-modality and multi-scale imaging, to determine how they are interpreted by the surgeon and whether they can be easily understood. Visualisation and information on-demand are aspects that can be explored in future usability studies in order not to interfere with the normal surgical workflow.
In addition to technical improvements, endomicroscopy as a clinical imaging modality for surgical margin detection and non-invasive monitoring of treatment should be further explored. More extensive clinical trials will be required in order to fully prove its clinical utility and allow it to be adopted in the normal clinical workflow. Additionally, since contrast agents must be applied on the tissue to achieve a good imaging contrast, further investigation should be performed in order to understand the utility of existing and new contrast agents with different tissues and at different concentrations. Also, the way that contrast agents are administered (locally or intravenously) should be further explored in order to provide optimal imaging results. In addition, an inherent issue with endomicroscopy, that should be addressed, is that \textit{in vivo} optical imaging requires the surgeon to have histopathology training, which is not an ideal solution, possibly distracting the surgeon from focusing on the treatment of the patient and making adoption less likely.

On-going research investigates the use of automatic image classification algorithms to create an atlas of images corresponding to specific diseases that will allow the system to identify diseased tissue and hence assist the surgeon [264]. In the meantime, parallel research is exploring online, real-time data streaming and collaboration between a remote pathologist and the surgical theatre that would alleviate the issue of histology training.

Finally, it is clear that the challenges of robotics for surgical microscopy is not only limited to mechatronics, but also optics and computer vision techniques to generate a single, unified imaging platform. New and emerging technologies are bringing together advanced optics, imaging, artificial intelligence, mechatronics and robotics to provide disruptive technologies that will benefit patients and surgeons. The new era of smart surgical robotic devices will offer a reduced footprint in the operating theatre, being less invasive and intuitive to use, while the integration of high resolution optical imaging will provide surgeons with an additional source of information to help them improve the precision of interventions, and ultimately improve outcomes for patients.
Appendix A

Optical Imaging Systems

A.1 Virtual Slit-Scanning Endomicroscopy System

The line-scanning system operates under a well-established principle of slit-scanning [47]. The system is developed by Dr Michael Hughes, so only a brief description is given here. Specific details of this implementation, which allows for high speed imaging, are presented by Hughes et al. [43].

This endomicroscopy system is essentially a fluorescence confocal microscope coupled to a fibre bundle probe to relay images from the tissue. A 488 nm laser is shaped to a horizontal line and scanned vertically over the proximal face of the fibre bundle probe by a scanning mirror. The bundle transfers the laser line to the tissue (with approximately 4 mW of power), where it excites a line of fluorescent stained tissue. The line of fluorescence returns along the bundle, where it is de-scanned, wavelength-filtered to remove reflected illumination light and imaged onto a linear CCD (2048 $\times$ 1 pixels). Fluorescence from out-of-focus regions of the tissue is defocused onto the camera and so largely undetected, resulting in an optical sectioning effect which reduces background blur and allows imaging of thick tissue. 2D fluorescence microscopy images, with a frame rate of up to 120 fps, are assembled from 500 lines acquired for consecutive positions of the scanning mirror. Common contrast agents, such as fluorescein and acriflavine, which are applied intravenously and topically respectively, are compatible with the system. Also, in
principle the system is compatible with any fibre bundle imaging probes and are mounted to the system using a custom adaptor.

The system is shown in Figure A.1.

Figure A.1: (a) Photograph and (b) schematic of the endomicroscopy imaging system, showing only the key components. Intermediate relay optics, and optomechanics are not shown for clarity. (c) Enclosed endomicroscopy system with custom fibre probe connected.

They key functional differences between the line-scanning endomicroscope and conventional point-scanning confocal endomicroscopes, such as Cellvizio [235], are the higher
frame rate (120 \textit{fps} compared to typically 12 \textit{fps}) and slightly reduced optical sectioning strength (\textit{i.e.} a greater fraction of background blur survives to impact on the image). The higher frame rate is important for improving mosaicking as greater overlap can be achieved between adjacent images with the same probe velocity [43].

\section*{A.2 Widefield Fluorescence Endomicroscopy}

Here the optical system used for Chapter 7 will be presented. This system is based on widefield fluorescence endomicroscopy [265]. It is illuminated by a royal blue LED (M455L3, Thorlabs, NJ, USA) directed through a low pass (450\textit{nm}) excitation filter (FES0450, Thorlabs) as the excitation source. The emission is captured using an emission filter (\geq 500\textit{nm}) (FEL0500), followed by a 5\textit{cm} focal length tube lens. Finally, the image is captured by a monochrome 8 bit camera (DCU223, Thorlabs). For increased flexibility, a leached fibre bundle is used to transmit the emission light and relay the emitted light to be captured. The proximal end of the bundle is fitted with a $\times10$ objective lens, while a dichroic mirror is placed proximally in order to bifurcate the path between the excitation source and the emission capture components. Raw images are captured at 30 frames per second, through a Labview interface.

![Enclosed widefield endomicroscopy system with a Fujikura fibre bundle attached.](image)

Figure A.2: Enclosed widefield endomicroscopy system with a Fujikura fibre bundle attached.
Appendix B

Autonomous Scanning for Endomicroscopic Mosaicking and 3D Fusion

During the PhD, I also collaborated extensively with Mr Lin Zhang and Mr Menglong Ye for performing intraoperative real time fusion of stereoscopic images with microscopic images. This work uses the Da Vinci robot and the DVRK framework with a 6-DoF stereoscopic visual servoing approach to perform autonomous large area endomicroscopy scans over a user-defined area and to fuse the 3D tissue surface with microscopic mosaics on-the-fly.

B.1 System Overview

The hardware components of the proposed framework consist of a patient side manipulator (PSM) of a da Vinci® robot with dVRK controllers, a stereo laparoscope, and a custom endomicroscopy system. The dVRK controllers are connected to a host PC via a

---

1Content from this chapter was published as:
IEEE 1394 firewire interface in a daisy chain topology. The stereoscopic system provides SD (720x576) video streaming for both left and right channels at 25 Hz. In the host PC, the stereo stream is captured by a Kona 4 PCIe frame grabber (AJA Video System). The endomicroscope system is the in-house laser line scanning system (see Appendix A.1). The microscopic images are captured in 300x300 pixels, which are served to the the controller at 80 Hz via a TCP/IP connection. In addition to the host PC, a second PC with modest configurations has been used dedicated to perform 2D mosaicking and 3D visualisation. The stereo images captured from the camera are used for 3D tissue surface reconstruction via a stereo matching method. The reconstructed 3D surface is used for two purposes: (1) to plan a scanning trajectory based on its position and normal information; (2) to provide 3D fusion with the 2D image mosaics. The video stream from the laparoscope system is also used for pose estimation of the endomicroscopy probe.

As shown in Figure B.1, the end-effector of the robot grasps an adapter that holds the endomicroscopy probe. A marker (KeyDot®, Key Surgical, Minnesota) is attached to the adapter for accurate pose estimation of the probe. The stereo system are calibrated using the method in [231] to obtain the intrinsic and extrinsic camera parameters, such that the 3D depths of the tissue can be recovered via stereo matching. The trajectory for an endomicroscopy scan can be planned with a simple 2D position input from the user who identifies the area-of-interest for tissue diagnosis. A visual control component closes the loop by comparing the current and desired probe poses and commanding the robot to minimise their difference, which assists the probe to follow the planned trajectory. The robot’s end-effector pose in Cartesian space is read and set via a dVRK-ROS component which is connected to a low level PID controller implemented by the SAW package using the cisst library [266]. Microscopic images captured from the endomicroscopy system are stitched together in real-time based on normalised cross correlation. The image mosaic is fused with the reconstructed surface on-the-fly to provide both macro- and micro-views of the scanned region. With accurate and robust pose estimation, the developed visual servoing approach enables the probe to conduct smooth and accurate scanning, so as to achieve continuous image mosaicking over a large and complex 3D surface intraoperatively.

206
With the proposed 6-DoF visual servoing approach, the robot end-effector can adapt the pose of the mini-probe to the tissue surface. This allows 3D fusion across macro- and micro-views, which are obtained from the stereo camera and the endomicroscope, respectively. It should be noted that the microscopic image mosaics are generated in a 2D image space with much higher resolution than the 3D tissue surface obtained from a standard stereoscope. To deal with the difference in resolutions, the 3D tissue surface has been interpolated to achieve the same resolution as the microscopic images. The interpolation is based on Radial Basis Functions (RBF), and can be performed immediately following stereo reconstruction. Efficient 2D-3D mapping can therefore be established for 3D fusion. The main steps of the 3D fusion process are: (1) 3D tissue surface reconstruction from stereo images; (2) interpolation of the 3D surface to achieve microscopic resolution;
Figure B.2: 3D fusion. The surface is densely interpolated on the area (colored as gray) for pathology analysis, and 2D-3D mapping is created so that visualisation of mosaics on 3D can be updated efficiently.

(3) creation of a 2D-3D mapping from the 2D pixels on the microscopic image mosaics to the interpolated 3D surface; and (4) update of the 3D surface on-the-fly when the microscopic images are being mosaicked. With a pre-interpolated 3D surface, the depth fusion task is efficient and can be performed at 25 Hz. To further improve the efficiency of the surface interpolation, the fusion process is only performed over the planned scan area. An illustration of the 3D fusion has been included in Figure B.2. Three ways of visualising the scanning and mosaicking on-the-fly are provided: (1) a view from the camera that observes the robot and target tissue; (2) a real-time updated 2D mosaicking map; and (3) an online updating 3D fusion of microscopic images onto the reconstructed surface. To assist the users to observe not only the scanning process but also the endomicroscopy image feed, the endomicroscopy images are overlaid over the marker on the probe, in a form of augmented reality.

B.2 Results

The assessment of the accuracy of the ELAS stereo reconstruction was performed using the Hamlyn cardiac datasets (http: //hamlyn.doc.ic.ac.uk/vision/) and presented compet-
Figure B.3: Exemplar 2D mosaicing results. (a) Mosaics of a 3x3 mm region on a kidney phantom; (b) Mosaics of a 2x2 mm region on an ex vivo porcine colon fat tissue sample.

itive accuracies over the state-of-the-art COCV [267] (ELAS: RMS 1.31 ± 0.98 mm - COCV: RMS 1.85 ± 0.82 mm). On the other side, the algorithm achieves real-time performance at 25Hz without any GPU programming. The use of the macroscopic visual servoing allowed to achieve sub-millimetre accuracy (3D translation) as, during the local autonomous scanning, the motion with visual servoing has achieved 0.211 mm, which is smaller than the FoV (0.24 mm) of the endomicroscope.

In order to assess the visualisation, autonomous scans were performed on a custom-made PVA cryogel kidney phantom and, to produce high quality microscopic images, lens tissue paper stained with the topical fluorescent contrast agent acriflavine, was placed on top of the surface. The scanning tasks were performed on the phantom surface using raster trajectories. An example of the obtained 2D mosaics is shown in Figure B.3(a), where a 3x3 mm region is covered. The actual covered region is slightly larger than 3x3 mm, due to the extra loading-unloading phase. The 3D fusion results are presented in Figure B.4(a) and (b), where Figure B.4(b) also demonstrates that the approach is able to perform multi-region scanning. These results have shown that the proposed framework is able to provide continuous image mosaicking, benefiting from the smooth and accurate motion generated by the visual servoing approach.

Experiments were also performed using ex vivo animal tissue. A porcine colon tissue sample was stained with the topical acriflavine contrast agent. Examples of the 2D mosaicking and 3D fusion results are provided in Figure B.3(b), and Figure B.4(c) and (d),
respectively. Both the phantom and *ex vivo* experiments have demonstrated that the framework is capable of performing autonomous large area 3D scans while generating high quality image mosaics. It is also worth mentioning that due to the advantage of global surface reconstruction and visual servo control, the system is able to scan multiple regions on the same surface. To this end, the robot moves the probe to an original standby pose after each scan is completed, and a new scanning region can be selected on the same surface.

![Figure B.4: Exemplar 3D fusion results. (a) Online fusion results and (b) multi-region fusion results of the phantom experiments, (c) and (d) online fusion results of the *ex vivo* colon experiment.](image)

This work was also extended, recently, to include a similar visual servoing approach, as described in Chapter 5, based on microscopic images. The robot moves the pick-up probe along the global trajectory based on endoscopic visual servoing and, when the probe arrives in the starting position of the scanning task, the robot drives the probe along the local scanning trajectory using both microscopic and macroscopic images. This technique presented better repeatability and accuracy compared to using only the macroscopic visual servoing.
Figure B.5: Example of 3D fusion results using microscopic and macroscopic visual servoing, (a) lens paper and (b) *ex vivo* porcine stomach tissue.

An example of the results obtained with the last approach from *ex vivo* experiments on lens paper and porcine stomach tissue samples are presented in Figure B.5.
Appendix C

3-DoF Robotic Instrument for Large-Area Transanal Microscopic Imaging

In this section, a robotic probe which provides large area mosaics (over 15 $mm^2$) is proposed. The device incorporates 3 active DoF, one for translating the microscopic probe and performing linear scans, one for rotating the instrument’s shaft, assisting with accessing areas around the lumen and providing the second axis for 2D scanning, and one joint mechanism to allow insertion through a surgical port and then allow perpendicular access to the tissue on the luminal walls. The described system includes image enhancement and mosaicking algorithms. The key features of the probe are its simple mechatronics and the ability to scan large 2D curved surfaces. It has a small diameter shaft which can be inserted through a standard surgical port, facilitating a wide range of applications for endoluminal surgery.

C.1 Mechatronic Design for Large Area Scanning

The overall design of the instrument is depicted in Figure C.1. The scanning part consists of a linear three phase DC-servomotor with analogue Hall sensors (LM 1247-020-
Figure C.1: 3D CAD renderings and photographs of the developed scanning instrument depicting: (1) the joint mechanism, (2) tip joint at various angles, (3) the rotational actuator for the shaft’s roll, (4) the microscopic probe’s channel, (5) the rotational actuator connected with a screw-based mechanism for driving the distal joint, (6) the tensioning mechanism for the tendons driving the joint, (7) the voice coil actuator for the shaft’s translation, (8) the workspace of the instrument while inside the patient, (9) the fabricated driver unit and (10) the fabricated instrument with casing.

01, Faulhaber, Germany) and two brushless DC-servomotors with analogue Hall sensors (0620C-006-B, Faulhaber, Germany). This linear DC-servomotor is chosen because of its compactness, precision, rapid response and adequate stroke length and force. Two dimensional scanning motion is achieved by axial translation of the shaft by the linear DC-servomotor, and rotation of the shaft axis, provided by one of the brushless DC-servomotors through a gear-transmission mechanism with a ratio of 1:3. The other brushless rotational DC-servomotor is connected to a screw-based shaft around which the tendons of the joint are wrapped. A tension assembly mechanism, presented in Figure C.1(6), comprising a wave spring and a screw-nut combination, keeps the tendons always under tension. The joint in the distal part has an internal channel of 2.5 mm where the microscopic imaging probe passes, and the probe is fixed through a stress-less spring-leaf-like assembly. The same channel can also be readily used for other types of imaging probes.
Porcine colon tissue
Endoscope
Endomicroscope

Figure C.2: Relative positions of endoscope, checkerboard pattern, and endomicroscopy probe.

The instrument body and the connecting parts such as the gear holders and the pulley mechanisms are fabricated with Veroblock material using a rapid prototyping machine (Objet EDEN350TM, Objet Geometries Ltd., USA). The linear DC-servomotor employed can provide a total stroke length of 100 mm and continuous force of 3.6N. The joint can rotate up to ±90 degrees (constrained using a mechanical stop on the joint), additional to the 360 degrees that the rotating shaft provides. The aforementioned technical features provide a significant workspace that the instrument can scan along the tumour site. The dimensions of the actuation assembly are 35 mm height, 118 mm length and 28 mm width and the overall weight is 200 g. The shaft and the joint mechanism have an 8 mm diameter which can fit into one of the instrument channels of conventional transanal endoscopic surgical ports (generally 10 mm diameter). Also, the shaft has a length of 250 mm which assists the surgeon to image tumours that are far from the port, enlarging the effective imaging workspace. Furthermore, all the motors provide accurate position sensing through their integrated Hall sensors; information that can be used for accurately tracking the probe tip.

For this instrument, the widefield endomicroscopy system presented in Appendix A.2 is used.
C.2 Optimal Scanning Velocity

Since the time taken to generate large area mosaics is crucial for the clinical translation of the instrument, the velocity selection is an important aspect of the proposed implementation. Given the core spacing $S_c$ of the fibre bundle, the camera exposure time $t_e$ and the fact that two cores are required per resolution element, then the maximum velocity without blurring the image is given by $v_{(\text{max},b)} = 2S_c/t_e$. Velocity selection should also account for the minimum overlap between consecutive images needed for effective mosaic generation. Therefore, given the FOV, the image frames acquired per second (fps) and assuming a required 75% overlap between consecutive images, the maximum velocity based on the mosaic requirement is given by $v_{(\text{max},m)} = 0.25 \cdot \text{FOV} \cdot \text{fps}/100$.

For the experiments reported below, a camera exposure of 10 ms and a frame rate of 35 Hz were used, resulting in the maximum velocities for the developed system of (a) $v_{(\text{max},b)} = 1.6 \text{mm/s}$ and (b) $v_{(\text{max},m)} = 8.5 \text{mm/s}$ . As a result, the desired linear velocity for the system is the lower value, $v_{(\text{max},b)} = 1.6 \text{mm/s}$.

C.3 Results and Validation

Figure C.3: Performance evaluation by probe tracking. (a) Repeated measured linear trajectories (black, blue red) compared with expected position (green line), (b) tracking information for repeated rotation-al trajectories (blue points), with fitted circle (red) for correlating the angle of the shaft to the encoder positions, (c) the correlation between the position encoder sensing (blue) and the measured rotation angle (green).

215
Validation of the device has been conducted using an NDI Aurora Electromagnetic Tracking System (NDI Corp, CA). In Figure C.3(a), the repeatability of the linear scan is assessed, with the black, blue and red lines showing the tracked tip position over three runs, compared to the expected position from the motor controller (green). In Figure C.3(b) the trajectory of the tip during rotational scanning is shown. Projecting these points onto a best-fit plane and fitting a circle, allows the centre of rotation, and hence the rotational angle to be calculated. The measured angle of rotation (green line in Figure C.3(c)) correlates well with the rotational motor encoder values (blue line), with $R = 0.9957$. This shows that it is possible to utilize the encoder value to determine the tip-position of the probe, despite mechanical imperfections and a long shaft.

In Figure C.4(a) a large mosaic of the USAF resolution target is presented, showing that the instrument can provide consistent linear mosaics. The plot shows the calculated position as a function of time (blue points) with a best fit line (red), showing the linearity of the scan. 2D Mosaics from lens cleaning paper and colorectal porcine tissue are showed in Figure C.4(b) and Figure C.4(c) respectively. Alternatively, Figure C.4(d) shows that very large linear mosaics could be reconstructed using the large range of motion of the linear
motor, proving an essential tool for the surgeon for scanning over early stage colorectal tumours larger than 1 cm. The time taken to scan the Figure C.4(b) area mosaic of 21.6 mm$^2$ is 54 sec, the Figure C.4(c) area mosaic of 15 mm$^2$ is 68 sec, and the Figure C.4(d) linear mosaic of more than 15 mm is 38 sec.
## Appendix D

### DH Parameters of the 6-DoF Robotic Instrument

The DH parameters for the developed 6-DoF Micro-IGES endomicroscopy instrument can be defined as presented in Table D.1

<table>
<thead>
<tr>
<th>Link</th>
<th>Theta $\theta_i$ (rad)</th>
<th>Link Offset $d_i$ (mm)</th>
<th>Link Length $a_i$ (mm)</th>
<th>Link Twist $\alpha_i$ (rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>$d_1$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>$\phi_2$</td>
<td>0</td>
<td>0</td>
<td>$\pi/2$</td>
</tr>
<tr>
<td>3</td>
<td>$\phi_3$</td>
<td>0</td>
<td>6</td>
<td>$\pi/2$</td>
</tr>
<tr>
<td>4</td>
<td>$\phi_4$</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$\phi_5$</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>$\phi_6$</td>
<td>0</td>
<td>11.3</td>
<td>$\pi/2$</td>
</tr>
</tbody>
</table>

Table D.1: Denavit–Hartenberg parameters of the 5-Dof robotic instrument.
Appendix E

DH Parameters of the Articulated Instrument for US and Microscopy Fusion

The DH parameters for the developed instrument of Chapter 7 can be defined as presented in the following table:

<table>
<thead>
<tr>
<th>Link</th>
<th>Theta $\theta_i$ (rad)</th>
<th>Link Offset $d_i$ (mm)</th>
<th>Link Length $a_i$ (mm)</th>
<th>Link Twist $\alpha_i$ (rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>$-\pi/2$</td>
</tr>
<tr>
<td>2</td>
<td>$\theta_i/3$</td>
<td>0</td>
<td>4.775</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>$\theta_i/3$</td>
<td>0</td>
<td>4.775</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>$\theta_i/3$</td>
<td>0</td>
<td>7.075</td>
<td>$-\pi/2$</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>$a_5$</td>
<td>$\pi/2$</td>
</tr>
</tbody>
</table>

Table E.1: Denavit–Hartenberg parameters of the robotic instrument for US and pCLE fusion.
Appendix F

Mechanical Details of the Presented Instruments

In this Appendix, the most important operation details of each instrument presented in the previous chapter are demonstrated.

Figure F.1: Pick-up mounts presented in Chapter 4; (a) Single modality pick-up probe dimensions and bottom view; (b) Multi-modality pick-up probe dimensions and bottom view; and (c) passive adaptive pick-up probe dimensions.
Figure F.2: Handheld robotic scanning device presented in Chapter 5 overview; (a) full and cross-sectional view of the instrument presenting its main design features; (b) visualisation of the approximate workspace of the cantilevered tube; and (c) the scanning mechanism in different instances demonstrating how the two cam rollers push the triangular cam and as a result the cantilevered tube, that is rigidly attached to the cam, moves.

Figure F.3: Micro-IGES robotic assisted microscopy instrument presented in Chapter 6 overview; (a) dimensions of the articulated instrument’s end-effector; (b) main components of the proximal actuation and motor pack; and (c) overall dimensions of the presented instrument.
Figure F.4: Bottom view of the multi-modality robotic instrument presented in Chapter 7, showing the linear scanning range (13.6 mm) and the angular range (±66°); (left) different instances of the linear scanning mechanism presenting the microscopy probe scanned in parallel to the ultrasound image plane and (right) different instances of the US probe sweeping motion. The combined sweeping motion of the US with the linear scanning of the microscopy probe allow for multi-modality 3D image fusion.
Appendix G

Image Reuse Permissions

IEEE Images

All images replicated in this thesis from IEEE publications have been referenced according to the following statement from the IEEE:

“The IEEE does not require individuals working on a thesis to obtain a formal reuse license, the permission format is as following:

In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.”

http://ieeexplore.ieee.org/

Springer Images

All images replicated in this thesis from Springer publications have been referenced according to the following format: (©[year of publication] reproduced from [authors] with permission from Springer). Individual licenses for each reused image have been obtained from Springer:
Licence No: 3934321213573 for Figure 2.14(d)

Licence No: 3934721114505 for Figure 2.15(b)

Licence No: 3934710419613 for Figure 2.15(c)

Licence No: 3934801109146 for Figure 2.17(c)

Licence No: 3934810078979 for Figure 2.18(b)

Licence No: 3934871482350 for Figure 2.18(c)

Licence No: 3931951168902 for Figures and context in Chapter 4 for the force adaptive probe

**Images From Companies And Research Institutes**

Figure 2.8(d) was obtained with permission from Farm PD

Figure 2.13(d) was obtained from personal written communication with Dr Cheol Song
Bibliography


[167] A. Bajo, R. E. R. Goldman, L. Wang, D. Fowler, and N. Simaan, “Integration and preliminary evaluation of an insertable robotic effectors platform for single port ac-


