Towards automated dynamic scene analysis and augmentation during image-guided radiological and surgical interventions

Contributions in motion-based segmentation and navigation uncertainty

by

Alborz Amir-Khalili

BASc Systems Design Engineering, University of Waterloo, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical and Computer Engineering)

The University of British Columbia
(Vancouver)

December 2017

Abstract

This thesis proposes non-invasive automated scene analysis and augmentation techniques to improve navigation in image-guided therapy (IGT) applications. IGT refers to procedures in which physicians rely on medical images to plan, perform, and monitor an intervention. In IGT, the tomographic images acquired before the intervention may not directly correspond to what the physician sees via the intra-operative imaging. This is due to many factors such as: time-varying changes in the patient's anatomy (e.g., patient positioning or changes in pathology), risk of overexposure to ionizing radiation (restricted use of X-ray imaging), operational costs, and differences in imaging modalities. This inconsistency often results in a navigational problem that demands substantial additional effort from the physician to piece together a mental representation of complex correspondences between the preoperative images and the intraoperative scene.

The first direction explored in this thesis, investigates the application of image-based motion analysis techniques for vessel segmentation. Specifically, we propose novel motion-based segmentation methods to enable safe, fast, and automatic localization of vascular structures from dynamic medical image sequences and demonstrated their efficacy in segmenting vasculature from surgical video and dynamic medical ultrasound sequences.

The second direction investigates ways in which navigation uncertainties can be computed, propagated, and visualized in the context of IGT navigation systems that target deformable soft-tissues. Specifically, we present an uncertainty-encoded scene augmentation method for robot-assisted laparoscopic surgery, in which we propose visualization techniques for presenting probabilistic tumor margins. We further present a computationally efficient framework to estimate the uncertainty in deformable image registration and to subsequently propagate the effects of the computed uncertainties through to the visualizations, organ segmentations, and dosimetric evaluations performed in the context of fractionated image-guided brachytherapy.

Our contributions constitute a step towards automated and real-time IGT navigation and may, in the near future, help to improve interventional outcomes for patients (improved targeting of pathologies) and increase surgical efficiency (less effort required by the physician).

Lay Summary

In this thesis, we present techniques to assist physicians during complex medical procedures that utilize medical images, namely cancer surgery and radiation therapy. Our contributions entail the development and application of advanced computer algorithms that can automatically process and enhance medical images. To assist physicians in finding the location of hard to see blood vessels, we propose fast, safe, and fully automatic techniques that can locate blood vessels based solely on how the vessels move (useful for surgical applications where vessels are covered by a layer of tissue). We also contribute algorithms and systems that can visualize potential ambiguities that occur during medical procedures, which include: uncertain location of tumor boundaries during kidney cancer surgery and effects of erroneous image alignment during radiation therapies.

Preface

This thesis is based on the following papers, resulting from collaboration between multiple researchers.

Studies described in Chapter 2 have been published in:

[5] A. Amir-Khalili, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, G. Hamarneh, and R. Abugharbieh. Auto localization and segmentation of occluded vessels in robot-assisted partial nephrectomy. In *Medical Image Computing and Computer-Assisted Intervention*, pages 407–414. Springer, 2014. Oral presentation - Acceptance rate: $\sim 4\%$, winner of one of five student best paper awards.

[7] A. Amir-Khalili, G. Hamarneh, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, and R. Abugharbieh. Automatic segmentation of occluded vasculature via pulsatile motion analysis in endoscopic robot-assisted partial nephrectomy video. *Medical Image Analysis*, 25(1):103–110, 2015. 2015 Impact factor: 4.565, h5-index: 48.

[126] M. S. Nosrati, A. Amir-Khalili, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, R. Abugharbieh, and G. Hamarneh. Endoscopic scene labelling and augmentation using intraoperative pulsatile motion and colour appearance cues with preoperative anatomical priors. *The International Journal for Computer Assisted Radiology and Surgery (IJCARS)*, 11(8):1409–1418, 2016. 2015 Impact factor: 1.827, h5-index: 25.

Studies described in Chapter 3 have been published in:

[6] A. Amir-Khalili, G. Hamarneh, and R. Abugharbieh. Automatic vessel segmentation from pulsatile radial distension. In *Medical Image Computing and Computer-Assisted Intervention*, pages 403–410. Springer, 2015. Poster presentation - Acceptance rate: $\sim 33\%$.

[8] A. Amir-Khalili, G. Hamarneh, and R. Abugharbieh. Modelling and extraction of pulsatile radial distension and compression motion for automatic vessel segmentation from video. *Medical Image Analysis*, 40:184–198, June 2017. 2015 Impact factor: 4.565, h5-index: 48.

Studies described in Chapter 4 have been published in:

[4] A. Amir-Khalili, M. S. Nosrati, J.-M. Peyrat, G. Hamarneh, and R. Abugharbieh. Uncertainty-encoded augmented reality for robot-assisted partial nephrectomy: a phantom study. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 182–191. Springer, 2013. Oral presentation - Acceptance rate: $\sim 27\%$.

Studies described in Chapter 5 have been published in:

[9] A. Amir-Khalili, G. Hamarneh, R. Zakariaee, I. Spadinger, and R. Abugharbieh. Propagation of registration uncertainty during multi-fraction cervical cancer brachytherapy. *Physics in Medicine and Biology*, 62(20):8116–8135, Oct. 2017. 2015 Impact factor: 2.811, h5-index: 56.

All listed publications were revised and edited by all co-authors.

In Amir-Khalili et al. [4, 5, 6, 7, 8, 9], as the primary author, I was the main contributor to the majority of writing effort, ideation, design, implementation, and testing of proposed methodology under the supervision of Dr. Rafeef Abugharbieh. I also presented the oral presentation for Amir-Khalili et al. [4, 5] and the poster presentation for Amir-Khalili et al. [6] conference papers. In all papers, Dr. Ghassan Hamarneh helped immensely with his valuable input on developing the original idea, improving the methodology, experimental design, and writing the paper. Dr. Masoud S. Nosrati, under the supervision of Dr. Hamarneh, provided input on improving the methodology and implemented the majority of the registration framework used in Amir-Khalili et al. [4].

Dr. Nosrati was the primary author on Nosrati et al. [126], undertaking the majority of the writing effort and developing the majority of the segmentation methodology, implementation, and testing of proposed methods under the supervision of Dr. Hamarneh. As the second author on this paper, under the supervision of Dr. Abugharbieh, I contributed to framing the motivations and the original idea of incorporating motion-based cues into a segmentation framework. I also contributed to writing parts of the introduction, methodology, and results section of the article. Specifically, I was responsible for the explication, development, and implementation of the motion-based vessel segmentation methodologies proposed in this article and producing and the visualization of the results.

In Amir-Khalili et al. [4, 5, 7] and Nosrati et al. [126], Dr. Jean-Marc Peyrat, under the supervision of Dr. Julien Abinahed, provided input on improving the methodology, experimental design, data acquisition, and assisted in establishing contacts with the Hammad General Hospital. Our clinical collaborators, Dr. Osama Al-Alao, under the supervision of Dr. Abudlla Al-Ansari, facilitated the data acquisition process, provided clinical feedback regarding the proposed methodologies, and helped in organizing the clinical user study in Amir-Khalili et al. [7].

In Amir-Khalili et al. [9], Dr. Roja Zakariaee, under the supervision of Dr. Ingrid Spadinger, provided the clinical motivations behind our contributions, facilitated access to the data, provided the segmentations and annotations of the data, and insights on the clinical significance of our findings.

Table of Contents

List of Tables

List of Figures

Glossary

da Vinci Surgical System Surgical laparoscopic robot developed by Intuitive Surgical, Inc. in Sunnyvale, California, USA.

DC gain Transfer function value at the frequency $s = 0$ for continuous-time transfer functions and at $z = 1$ for discrete-time functions.

elastix A free, open-source, and multi-platform toolbox for automatic registration of volumetric medical images developed originally by University Medical Center Utrecht in The Netherlands.

ITK-SNAP A free, open-source, and multi-platform software for semi-automatic segmentation of volumetric medical images.

MATLAB Matrix Laboratory, a multi-paradigm numerical computing environment developed by The MathWorks, Inc. in Natick, Massachusetts, USA.

RENAL Nephrometry Score A method for quantifying the salient characteristics of renal mass anatomy in an objective and reproducible manner.

SOMATOM Scanner X-ray computer tomography scanner developed by Siemens Medical Solutions in Erlangen, Germany.

List of Acronyms

AA abdominal aorta.

AM appearance model.

AUC area under the receiver operating characteristics curve.

BM brightfield microscopy.

CCA common carotid artery.

CNN convolutional neural network.

CT computed tomography.

DIR deformable image registration.

DOF degrees of freedom.

DSA digital subtraction angiography.

DSC Dice similarity coefficient.

DUS dynamic ultrasound.

DVH dose-volume histogram.

EKF extended Kalman filter.

EV endoscopic video.

EVM Eulerian video magnification.

FS frequency smoothing.

GBS graph-based segmentation.

GPU graphical processing unit.

GT ground truth.

GWIR group-wise image registration.

HT Hough transform.

HU Houndsfield unit.

IGT image-guided therapy.

IVC inferior vena cava.

KMVS kinematic model-based vessel segmentation.

MF monogenic flow.

MFCCB multi-fraction cervical cancer brachytherapy.

MIS minimally invasive surgery.

MO morphological operations.

MON monogenic.

MR magnetic resonance.

MRA magnetic resonance angiography.

MTD magnitude of total vertical and horizontal displacements.

NCC normalized cross-correlation.

NURBS non-uniform rational B-spline.

OAR organs at risk.

OCT optical coherence tomography.

OF optical flow.

OOF optimally oriented flux.

PBMS phase-based motion segmentation.

PRMM pulsatile radial motion model.

RA renal artery.

RAPN robot-assisted partial nephrectomy.

RF radio frequency.

ROC receiver operating characteristics.

RU registration uncertainty.

RV renal vein.

SQF spherical quadrature filters.

TRE target registration error.

TSMT tubular spring-mass tracker.

US ultrasound.

USCM ultrasound confidence maps.

Acknowledgments

As I reflect back on the long and tumultuous journey that led me to this point in my life, I feel fortunate and grateful for all of the wonderful people who have helped me find my way.

First, I would like to thank my research supervisor Dr. Rafeef Abugharbieh for being so patient with me and for giving me the freedom and means to pursue research topics that I was deeply passionate about. Thank you for your advice, for sending me to so many conferences, and for having faith in me. Not many people are afforded such luxuries. Thank you!

My sincere gratitude to all of my co-authors, without whom this thesis would not have been possible. Special thanks to Dr. Ghassan Hamarneh, co-author on all of the works presented in this thesis and a constant source of amazing and novel ideas. It was a privilege to work with you. Everything I know about scientific writing, I learned from Dr. Abugharbieh and Dr. Hamarneh. Thank you both for being so generous with your time! Thank you to Dr. Masoud Nosrati and Dr. Jean-Marc Peyrat for entertaining the zany ideas of an overenthusiastic fledgling researcher throughout the day, in the middle of the night, during holidays, and across continents. Thank you all for your help with the late conference deadlines and for teaching me how to persevere and strive for the best. With your help, our 2014 MICCAI paper became one of my proudest achievements.

I am deeply indebted to my supervisory committee: Dr. Antony Hodgson and Dr. Purang Abolmaesumi. Thank you for your support and for sparking my interest in the field of image-guided therapy. This thesis would not have been possible without your help.

Thank you to the QRSC staff: Dr. Abulla Al-Ansari, Dr. Julien Abinahed, Dr. Jean-Marc Peyrat, Dr. Osama Al-Alao, Dr. Nikhil Navkar, Winnifred Gonsalves, Georges Younes, and Dr. Sarada Dakua for being such amazing friends and hosts during my stay in Qatar. It was an unforgettable experience and a pleasure to work with you and to learn from you.

Thanks to all of my fellow BiSICL members, you guys are the best group of people I have ever had the pleasure of working with. Thank you for exposing me to your research; I have learned so much from all of you. To all of my friends from EiS and UBC, thank you for putting up with my shenanigans.

Belated thank you to my University of Waterloo professors: Dr. Alexander Wong, Dr. Paul Fieguth, and Dr. David Clausi for nurturing my interest and love for signal processing and machine learning. And a long overdue thank you to my secondary school science and mathematics teachers: Josef Osif, Mike Langille, and Bob Walter for setting me on this exciting path fourteen years ago.

Last, but not least, thank you to my family for their unconditional love and support; especially to my parents. Thank you to my better half, Maria-Jose Ramirez, the paragon of patience!

Dedicated to my grandmother, whose stoic and valiant battles with cancer inspired my work.

Chapter 1

Introduction

Computers have become ubiquitous appliances in the modern world and our dependency on them is incontrovertible. The impact that computers and computer algorithms have had on medicine and medical research are equally profound. Computing plays an integral role in modern medicine and it has been used to address problems spanning a wide gamut of applications that include health informatics, medical imaging, patient monitoring systems, clinical decision support systems, medical databases, surgical robotics, biomedical research, and many more. Breakthroughs in computer processing power and novel algorithms have allowed researchers to acquire, store, and analyze detailed volumes of medical data; which has led to ground breaking medical discoveries and the development of more advanced and accurate diagnostic and interventional tools. Although research on applications of computing in medicine has been ongoing for more than 40 years, recent accelerating advancements in medical image analysis and computer-assisted interventional systems are a testament to the potential of computing in further revolutionizing current medical practices and research methods [114].

Over the past three decades [85], a growing body of research contributions, promoting the usage of intraoperative imaging technologies and computer algorithms as guidance tools during both surgical interventions and interventional radiotherapy, have coalesced to form the exciting and burgeoning field of image-guided therapy (IGT) [82, 83, 114]. Research into new and improved IGT systems demands expertise from diverse disciplines that range from medicine to engineering and computer sciences. Such multidisciplinary collaborations have been shown to be fraught with challenging system design problems [28, 115]; more so when the system is being developed in a clinical context, where human lives are at stake. Within the solution to these challenges, however, lies the key to unlocking new paradigms of interventional techniques and therefore benefits for patients.

1.1 Thesis Motivation

This thesis is motivated by unresolved engineering problems in computer-assisted IGT navigation and investigates novel automated scene analysis and augmentation methods that can be leveraged to overcome them. This introductory chapter entails the definition of IGT, explication of associated challenges, and the rationale behind our scientific contributions proposed to address emerging challenges in IGT.

1.1.1 Image-Guided Therapy

Image-guided procedures are broadly defined as interventional techniques in which the physician relies heavily on medical imaging technologies to plan, perform, and monitor an intervention. IGT systems can therefore be thought of as the integration of three major contributing technologies: imaging, guidance (navigation), and therapy delivery devices [82]. Indeed, the unifying objectives and associated benefits of IGT within this definition are difficult to encapsulate in just a few sentences due to the broadness of its scope, which encompasses both surgical and radiotherapy applications, and an overlap with the field of minimally invasive surgery (MIS) research. There is, however, an important distinction to be made between IGT and MIS techniques. In a recently published book on IGT, Jolesz [82] defines IGT as systems that leverage medical imaging to both, improve targeting and control over therapy delivery, and to decrease the invasiveness of a procedure; the latter of which is a shared objective with that of MIS methods such as laparoscopy, whereas the former is not.

IGT systems that are characterized by the aforementioned objectives may be described using the time-line based view proposed by Yaniv and Cleary [203], and categorized into three phases: preoperative planning, intraoperative plan execution, and postoperative assessment. Within this time-line based view, IGT systems generally follow a sequence of constituting steps (Figure 1.1) that was later described by Cleary and Peters [28].

Naturally, the interest of this thesis lies in the applications of computer-assisted technologies in IGT, the main involvements of which are in the preoperative and intraoperative phases. Thus, the key challenges that stand out in this systemic view of IGT are the medical imaging technologies and the apparent need for reconciliation between preoperative and intraoperative imaging modalities. Preoperative imaging involves the acquisition of tomographic image volumes, typically X-ray computed tomography (CT) and magnetic resonance (MR) images. These tomographic images enable the physician to peer through the skin of the patient and view the anatomical, and sometimes functional, state of different systems within the patient's body. Tomographic images are therefore essential to the generation of the

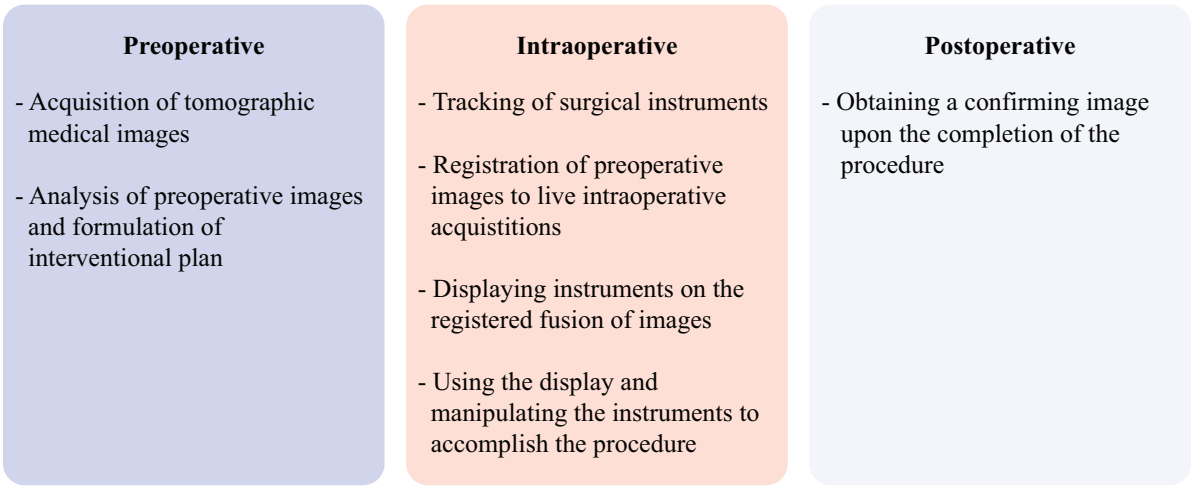


Figure 1.1: Typical steps in image-guided therapy systems organized in a time-line based view.

interventional plan, which further implies that the quality of preoperative imaging may have a direct impact on the effectiveness of the interventional plan and, by extension, the outcome of the therapy.

In addition to the quality of imaging and the interventional plan, it stands to reason that the execution of interventional plan during the operation also has a notable impact on IGT outcomes. Due to time-varying changes in the patient's anatomy, risk of overexposure to ionizing radiation, operational costs, and incompatibility with therapy delivery devices, the interventional plan may not necessarily represent the intraoperative state of the patient or, more importantly, the interventional plan may not directly correspond to what the physician is able to see via the intraoperative imaging modalities.

This inconsistency may complicate the execution of the interventional plan as it often results in a navigational problem that demands substantial additional cognitive effort from the physician to piece together a mental representation of complex correspondences between the preoperatively conceived plan and the intraoperative scene [183]. In lay terms, the navigation problem in IGT is akin to—albeit a more complicated version of—the common problem of finding one's way, in first person, from the current street address to a destination using only a memorized representation of the route from a map presented in bird's-eye view. Anyone who has ever been lost can easily appreciate the implications of this problem in the context of IGT wherein, by extending the metaphor, the path being traversed lies in 3D space, the map is outdated and is of questionable quality, the streets are not marked, landmarks move in relation to each other, and a wrong turn may jeopardize the outcome of the intervention.

This intraoperative navigation problem, from a technical perspective, is a decisive issue for IGT systems and is a suitable candidate for computer-assisted guidance solutions. Hence, we look into the domain of computer-assisted intraoperative navigation to identify the emerging engineering research problems in IGT.

1.1.2 Computer-Assisted Intraoperative Navigation

The solution to the computer-assisted navigation problem in IGT is by no means trivial. A one-size-fits-all computer-assisted solution is impractical because of drastic discrepancies between the imaging modalities or therapy delivery devices used during the preoperative and intraoperative phases of different IGT applications [16]. Nonetheless, existing computer-assisted guidance methods that are currently employed to address IGT navigation problems may be considered as specialized amalgams of the following established signal processing, control system, and computer vision methodologies [28, 82, 203]:

- Low-level image processing: denoising, enhancement, and augmentation of image information
- Image segmentation: delineation or partitioning of regions/structures of interest
- Registration: alignment of multiple data sets into a single coordinate system
- Tracking: determining location and position of tools and anatomical structures
- Visualization: displaying the fusion of complimentary information from registered images

Among these methodologies, registration and tracking are often considered to be the most integral and challenging components of the IGT navigation problem [28, 82, 106, 170, 181].

Challenges in Medical Image Registration

Registration and tracking are difficult in many IGT scenarios, e.g., resection of soft-tissues [12, 185, 206], considering the different reference frames of acquisitions, the heterogeneity of the imaging modalities (in terms of their dimensionality, appearance, and the meaning of the underlying physical measurements), and the likely motion and deformation of anatomical structures. For example, in the context of robot-assisted laparoscopic surgery, raw intensity information from a preoperative CT scan bears little visual and structural resemblance to 2.5D color video data acquired intraoperatively by a stereoscopic endoscope. In such extreme cases, where the intraoperative imaging modalities cannot wholly capture the deformation of organs, the problem of image-based registration is often considered to be ill-posed.

Consequently, in IGT applications that target deformable soft-tissue structures, real-time registration is difficult to achieve without the help of complex hardware solutions involving electromagnetic or optical tracking systems and the use of fiducial markers. During complex abdominal interventions, e.g., kidney cancer surgery [12, 185], some IGT guidance systems also necessitate the use of additional, potentially harmful, intraoperative imaging systems. Aside from the often considerable operational costs associated to these hardware tools, the use of such tools in some situations, e.g., aforementioned abdominal interventions, is arguably antithetical to one of the principles of IGT as they may increase the invasiveness of the procedure [36, 106]. This trade-off between guidance accuracy and invasiveness is difficult to quantify and justify in many interventional applications. There is therefore a trend, in the field of IGT research, towards scene analysis and augmentation solutions that strive to overcome navigation problems without resorting to the use of invasive tools or harmful imaging techniques.

Intraoperative Image Segmentation

A possible non-invasive solution to facilitate the task of preoperative to intraoperative registration is to identify corresponding landmarks independently in both preoperative and intraoperative images, i.e., segmenting the intraoperative scene independently of preoperative images. Among computer-assisted approaches to the IGT navigation problem, some automated image analysis and augmentation methods [64, 113, 124–126, 130, 136, 205] adhere to a more mathematically elegant and holistic perspective of the navigational problem. Though it is often useful to compartmentalize navigation into different sub-problems of image segmentation, registration, tracking, and visualization, it is also important to note and appreciate the fundamental relationships that exist between them. These sub-problems may be mathematically modelled in such a way that the computational solution to the navigation problem would jointly produce an optimal solution for the requisite segmentation, registration, and tracking sub-problems. Such mathematical frameworks cast the problem of navigation as a numerical optimization problem, the objective of which is deftly formulated in such a way that it corresponds to the true objective of navigation—and associated sub-problems—as closely as possible. It stands to reason that, within such mathematical formulations, an insight into improving one sub-problem would likely result in an improvement to another [64]. The detection of corresponding landmarks, such as location of blood vessels, between the two intraoperative and preoperative imaging modalities may therefore simplify and expedite the registration step.

Automatic localization or segmentation of structures from preoperative data has been thoroughly researched for many years [155]. Analysis of preoperative image data is also typically not subject to hard performance constraints as preoperative data can be analyzed and processed prior to the actual operation. While on the other hand, segmentation of structures from intraoperative imaging modalities are arguably more challenging; this is due partly to stringent demands on real-time performance [82] and partly to the limitations of the intraoperative imaging modalities, e.g., poor signal quality or inability to see through the surface of tissues. But, intraoperative modalities are not as limited as they seem. In many IGT application, e.g., image-guided surgery, intraoperative imaging modalities are used to provide real-time information from the interventional scene to the physician and, as a result, are able to capture dynamic or temporal behavior of the patient's anatomy. Automatic localization of structures may be difficult to do in real-time from a single static image, but the dynamic image information acquired by the intraoperative imaging modality can also be used to improve localization [5–8, 53, 111, 142].

Navigation Uncertainty

In addition to the challenges associated to registration and intraoperative segmentation, a perhaps more fundamental problem in computer-assisted IGT navigation systems is regarding the inevitable sources of uncertainty. A perfect navigation system is unattainable in practice due to inherent errors or uncertainties in imaging, segmentation, localization, registration and tracking; which also is precisely why IGT navigation systems are never fully automated and instead rely on the expertise of a physician in the final decision making process. Even if a mathematical navigation framework has been established to align the preoperative plan with the intraoperative frame of reference and associated location of therapy delivery instruments, it is of critical importance to consider the computational errors and uncertainties that may rise during this numerical optimization process. It is important—and arguably ethical—to estimate and display imprecision of such IGT navigation systems to the physicians such that they are not forced to make decisions based on false determinations. Though some IGT systems can visualize the imprecision of the tools using a circle or an ellipse to represent the likely registration and localization uncertainties [159, 160, 184], the majority of IGT systems are limited by processes or results that do not encode uncertainty information in one of the navigational tasks (segmentation, tracking, and registration), none of which is guaranteed to be accurate, especially the paramount registration stage. In this thesis, we use the adjective 'crisp' to refer to such deterministic or non-probabilistic variables, results, and processes that do not encode uncertainty information. Moreover, without an uncertainty-encoding visualization or augmentation, physicians are rendered oblivious to the levels of trust that should be bestowed on the navigational results presented to them. This observation furthermore raises another important systemic problem in regards to the ergonomic factors of the system, i.e., the level of trust a physician can bestow on the navigation system in light of the innate uncertainties that exist in the imaging and computer-assisted navigation solutions [176].

1.2 Problem Statement

1.2.1 Thesis Objectives

The objective of this thesis is to develop automated scene analysis and augmentation techniques that can improve intraoperative navigation in IGT procedures without increasing the level of invasiveness to the patient. In this section, we translate this objective into concise research questions that will guide us to better understand the limitations of existing solutions and to frame our contributions within the related prior art associated to computer-assisted navigation methods.

1.2.2 Research Questions Addressed

In Section 1.1.2, we motivated that the independent segmentation of structures in preoperative and intraoperative images may improve the difficult task of image registration. We also motivated that the kinematic behavior of anatomical structures, which is encoded in the dynamic information of medical image sequences, may be used as an independent source of information for the purpose of segmentation. In this thesis, we choose to investigate motion-based segmentation of blood vessels because vascular structures exhibit unique periodic pulsatile radial motion characteristics (they are pulsating tubes) that may be captured with intraoperative imaging modalities, e.g., dynamic ultrasound (DUS) and endoscopic video. Furthermore, as we elaborate later in Section 1.4, automatic vessel localization methods are applicable to many image-guided medical diagnosis and interventional procedures. Segmentation of vessels in IGT is however challenging due to (i) low spatial and temporal imaging resolution (motions may be faint), (ii) presence of occlusions (vessels hidden under tissues), and (iii) spatial and temporal noise artefacts. The first research question that we seek to address in this thesis is hence:

Research Question #1: How can motion information be used to automatically segment blood vessels from dynamic medical images during IGT where: there are multiple sources of motion, the data is noisy, vessels may be occluded by layers of tissues, and observable vascular motions are faint?

Our second research question investigates the systemic problem of navigation uncertainty. In the context of IGT, the added value of the end-to-end computation of navigation uncertainties has been previously demonstrated in the context of orthopedic interventions [159, 160]. In image-guided orthopedic interventions, which involve the alignment of rigid organs (i.e., bones) and rigid surgical tools (e.g., plates, screws, and surgical drills), the effects of navigation uncertainties (specifically from tracking, calibration, and registration) can be gleaned from the imprecision of rigid transformations. In the context of image-guided interventions that target deformable tissues, on the other hand, computation and propagation of navigation uncertainties are more challenging for two reasons: (i) there are additional sources of uncertainties (e.g., segmentation and surface reconstruction), and (ii) deformable image registration uncertainties are more challenging to compute in an efficient manner. As such, the second research question is:

Research Question #2: How can different sources of navigation uncertainties be computed, propagated, and visualized for image-guided medical interventions that target deformable tissues?

To address both of these two research questions, it is crucial to first understand the capabilities and limitations of the imaging modalities that are currently used during IGT. In the following sections we first provide an overview of three popular medical imaging modalities used in IGT before surveying the prior art relating to the proposed research questions.

1.3 Medical Imaging Technologies in Image-Guided Therapy

Medical imaging technologies play a central role in IGT. Existing computer-assisted navigational solutions employed during IGT are designed around the capabilities and limitations of these medical imaging technologies. This section thus provides a general overview of three very different imaging modalities that are commonly used during IGT: CT, ultrasound, and stereo endoscopy. Among these modalities, CT is the oldest and most widely used imaging modality, often employed during all three (pre, intra, and postoperative) stages of IGT interventions. Ultrasound (US) is another well established, although safer and more portable, imaging modality favored primarily for diagnostic and intraoperative applications. Stereo endoscopy, commonly used during robot-assisted MIS, is a newer extension of traditional optical endoscopy that has garnered considerable attention in recent years.

Although this thesis focuses primarily on the three aforementioned modalities, these modalities and the applications for which we present our contributions are, in essence, effective examples for the exposition of our proposed mathematical methodologies. By design, the methodologies presented in this thesis are not explicitly restricted to these modalities and may be extended to other prevalent imaging modalities used in IGT including MR imaging, single-photon emission computed tomography, positron emission tomography, X-ray fluoroscopy, diffusion-weighted MR imaging, and various contrast enhanced specializations of these modalities.

1.3.1 X-ray Computed Tomography Imaging

X-ray imaging has historically played a central role in IGT systems and may be considered as the original inspiration behind this paradigm of interventional techniques and technologies [114]. Among existing X-ray-based medical imaging modalities, which essentially are techniques that utilize X-ray sources and X-ray detectors in different configurations to image a patient, CT is considered to be the workhorse of all interventional procedures that require cross-sectional imaging [61]. CT imaging is often preferred over other tomographic imaging techniques, such as MR, for imaging anatomical structures because of its relatively faster acquisition speeds, finer imaging resolution, superior visualization, and cheaper operational costs.

As a preoperative imaging modality, X-ray imaging modalities are primarily used during image-guided orthopedic interventions, in the context of which CT imaging is used to plan for the treatment of complex fractures and implant placements [60]. CT is also often considered to be the gold standard for diagnosing and staging of different types of cancers including lung carcinomas [121], renal cell carcinoma [146], pancreatic carcinoma [50], and colorectal cancers [129]. By extension, CT has proven to be a suitable modality to use during the planning stages of many different image-guided surgical and radiotherapy interventions.

X-ray imaging technologies, however, cannot differentiate between certain soft-tissues as effectively as MR imaging techniques. Therefore, the utility of CT in diagnosing certain types of cancers such as early detection of prostate cancer [32] is limited. Another critical downside of X-ray imaging—and by extension CT imaging—is in regards to the use of harmful ionizing radiation. As a consequence of the potential for adverse health effects to patients and clinicians, the use of CT imaging during the intraoperative stages of IGT is reserved strictly for interventions where the benefits of CT imaging outweigh its adverse effects.

Despite associated health risks, CT imaging is currently used as an intraoperative imaging modality during many sensitive IGT procedures including: image-guided external beam radiotherapy, brachytherapy, CT-guided needle biopsy or radio frequency ablation, vascular procedures, orthopedics and neurosurgery; with many more applications to emerge with future developments in CT-guided surgical robotic technologies [173, 189].

1.3.2 Dynamic Ultrasound

US imaging, or ultrasonography, is an imaging technology that is also ubiquitous and well established in clinical practice and medical research. The prevalence of US in the medical imaging community is primarily due to its ability to acquire images of soft-tissue structures located beneath the surface of the skin (e.g., tendons, muscles, joints, vessels and internal organs) in real-time without exposing the patient or the physician to harmful ionizing radiation. Medical US is predominately used during clinical diagnostic procedures but many advocate for its use within image-guided surgical interventions in part due to the aforementioned benefits and also because medical US machines are inexpensive and portable compared to some of the other existing medical imaging technologies.

Though the affordances associated to US in an interventional context are fairly obvious, the drawbacks of US have in past limited its use to rather rudimentary procedures, e.g., catheterization and needle biopsies. US imaging is impeded by various limits on its field of view [45] and a low signal-to-noise ratio. Due to the physics of US, image quality is highly dependent on the size, orientation, and acoustic properties of structures being imaged. As a result, US image quality depends on the expertise of the US operator; and even with a highly skilled operator, it is often impossible to image structures behind bone and air pockets. Moreover, US image quality is also degraded by the presence of signal dependent speckle noise [139], which further necessitates the need for an expert to interpret the acquired images. This need for expertise, or specialization, in US image acquisition and interpretation has been an insidious barrier to the adoption of US as the primary intraoperative imaging modality during surgical interventions.

The capabilities of US imaging technologies have been improving steadily since its introduction, overcoming the associated drawbacks with the help of breakthroughs in hardware and software. There have been marked advancements in US hardware in terms of transducer sensitivity, beam-forming, and image processing speed, which have improved the trade-off between image quality versus acquisition speed. Other hardware-based solutions such as tracking systems and 3D volumetric US probes [45] have also been developed to overcome the field-of-view limitations. Concurrent with the advancements in

hardware, countless automated algorithms have been proposed by the research community to reduce the reliance on an expert for analysis and interpretation of the images. Additionally, advanced technologies such as Doppler US, US elastography [208], high-intensity focused US, and contrast-enhanced US have also been developed to extend the capabilities of US devices beyond imaging of soft-tissues.

The steady and incremental improvements in US capabilities have paved the way for the proliferation of US imaging into formerly unconventional surgical applications such as orthopedic surgery [62] and cancer treatment [116]; with more applications emerging on the horizon. Coupled with the heightened level of projected growth and competitiveness among manufacturers of US equipment, as well as the trend towards open-software US platforms [96], medical US imaging is poised to remain as a dominant modality of intraoperative imaging during image-guided interventions.

1.3.3 Stereo Endoscopic Video

Recently, robot-assisted MIS techniques have emerged as viable alternatives to traditional abdominal laparoscopic surgery techniques; providing surgeons with the added benefits of superior ergonomics, advanced vision systems, intuitive control over wristed instruments, scaling or miniaturization of movements, and filtering of hand tremor; which have led to increased dexterity and higher precision [95]. The purpose of such robots is not to replace the surgeon, but to effectively augment and enhance the dexterity and perceptual capabilities of the surgeon instead by overcoming the ergonomic ambiguities that exist in laparoscopic surgery. Surgical laparoscopic robots such as the da Vinci Surgical System are thus designed in a unidirectional master-slave configuration. In such configuration, the surgeon interacts with the patient by directly controlling the robot via a surgical console that in turn provides visual feedback from the surgical site to the surgeon via the high-definition stereo endoscopic camera.

The stereo endoscope of the da Vinci Surgical System, is regarded as one of its greatest advantages over traditional laparoscopic methods. The stereo cameras provide the surgeon with added depth perception, which has been shown to improve surgical performance for both novice and experiences operators [21]. Furthermore, the video processing pipelines used within the da Vinci Surgical System and other similar robot-assisted systems create a natural and practical platform for the deployment of advanced computer-assisted algorithms for automated scene analysis and augmentation [135]. As a direct result, research into endoscopic video analysis has gained considerable momentum in recent years and many novel imaged-based computer-assisted systems have been proposed to leverage endoscopic video as a primary intraoperative imaging modality for automated guidance. A timely survey by Bernhardt et al. [16] identified 279 academic publications relating to surgical endoscopic scene augmentation published between 2000 to June 2016, more than half of which were published after 2011.

Most of image-based laparoscopic guidance systems proposed to date are motivated by a common objective: to compensate for the loss of tactile or haptic feedback incurred with the transition from traditional open style to laparoscopic methods of surgery. The importance of tactile sensations in context of surgery is patently evident as, in open surgeries, surgeons rely on manual palpation of tissues to identify important structures hidden beneath the surface. The location of these important structures are often directly tied to surgical outcomes as they include: blood vessels, nerves, and pathologies such

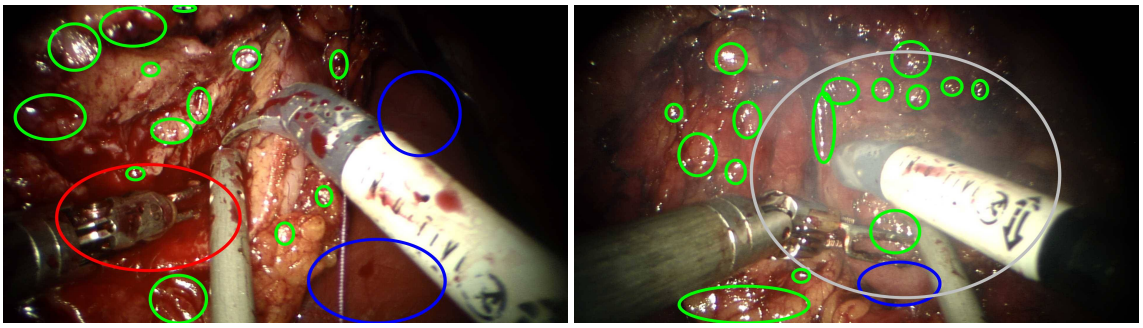


Figure 1.2: Examples of challenging noise artefacts present in video sequences acquired by one of the stereo-endoscopic cameras of the da Vinci Surgical System during a partial nephrectomy procedure. These artefacts, caused by the presence of blood (red), white specular highlights (green), smooth organ surfaces (blue), and smoke (gray), pose unique challenges to automatic localization of anatomical landmarks.

as tumors and cysts. The dominant trend in laparoscopic guidance has therefore been to compensate for the loss of the ability to perceive hidden or occluded structures by augmenting the intraoperative endoscopic video with patient-specific models extracted from preoperative CT and MR image volumes.

The trend of image-based surgical guidance is however not free controversy and criticism by advocates of hardware based solutions, which instead aim to integrate haptic feedback directly into the next generation of surgical robots [37]. Admittedly, the field of image-based laparoscopic guidance has yet to reach maturity and automated augmentation of endoscopic video still remains an open problem partly due to the presence of challenging conditions such as deformable tissue motion, limited field-of-view, presence of non-Lambertian specular highlights, smoke, blood, and tissue surfaces that are smooth (Figure 1.2). The haptics-based approach to this problem, on the other hand, is perhaps more challenging as the development of a true-to-life haptics feedback system is contingent on: (i) the development of complex sensors to measure instrument interactions and (ii) interfaces to convey force/tactile sensations back to the surgeon; both of which are far from being justifiable in terms of cost despite the push towards improved haptic displays and force/tactile sensors that are high-performance, low-cost, biocompatible, and sterilizable.

Indeed, advantages can be achieved in robotic surgery by employing the underutilized haptic sensory channel but such advantages will arguably not surpass that of traditional open surgery; whereas image-based guidance methods that leverage advanced computer vision techniques possess the potential to improve the surgical practice beyond what is currently possible [16].

1.4 Segmentation of Vasculature from Dynamic Medical Image Sequences

The three medical imaging modalities presented in the previous section provide an exemplary representation of just how different medical images can be; CT images measure the radiodensity of tissues, B-mode US measures the differences in acoustic impedance, and stereo endoscopy captures sequences

of color images of the surfaces of organs inside of the patient at two different viewpoints. Automatic localization of corresponding landmarks is challenging as the static appearance of tissues and organs does indeed vary across these imaging modalities. However, the underlying temporal behavior and apparent motions of some anatomical structures remains unchanged despite the difference in appearance. We propose that these dynamic characteristics are a missing piece to the large puzzle that is the IGT navigation problem.

Within the scope of this thesis, we aim to improve IGT guidance by investigating non-invasive automated scene analysis and augmentation techniques that can localize anatomical landmarks, i.e., blood vessels, from dynamic intraoperative imaging modalities. Specifically, we aim to establish that an understanding of the temporal anatomical behavior of blood vessels can be mathematically modelled and leveraged to facilitate automatic localization of these vessels in a computationally efficient manner. Fortunately, blood vessels in particular have unique shape and motion characteristics, i.e., they are essentially pulsating tubes that periodically pulsate with a frequency that is predominantly governed by the heart-rate. These pulsatile motion characteristics, once computed, can be utilized as novel features for vessel segmentation and may be incorporated alongside other complimentary static visual features such as color, intensity, shape, and texture.

Due to the sheer breadth of associated applications, which extend beyond IGT into clinical diagnostic assessment, automatic vessel segmentation from medical images has been a popular subject of research for many years and has resulted in many notable contributions. Common applications of vascular imaging range from routine non-invasive diagnostic procedures to complex surgical interventions. Vascular imaging is routinely used to assess the risk for cardiovascular morbidity by (i) directly imaging and analyzing the coronary arteries with intravascular US, MR, or CT imaging; (ii) quantifying arteriosclerosis from color images of the retina [128]; (iii) segmenting atherosclerotic plaque from US [17], MR [34], or CT images [105] of the common carotid artery (CCA); or (iv) monitoring changes in vascular distensibility from MR images of the aorta [23] and CT angiography images of the CCA [65]—all of which have been identified as independent predictors of stroke [24, 66, 84, 97, 201]. In the domain of IGT, the real-time acquisition speed and noninvasive nature of US imaging have popularized its utility for guidance during commonly performed, yet laborious, vascular access (cannulation) procedures such as insertion of central venous and arterial pressure catheters [104]. Moreover, vascular imaging is used regularly during preoperative planning and screening of surgical interventions like kidney and liver transplants [63, 86]. Finally, in addition to the traditional applications of X-ray fluoroscopy and CT angiography during image-guided cardiac catheterization [59] and aneurysm surgery [138], vascular imaging is finding new applications in intraoperative guidance during robot-assisted prostate and kidney cancer surgeries [5, 7, 110, 187].

Extraction of vascular structures are of such importance that many acquisition techniques and imaging modalities have been specifically developed to enhance the appearance of vasculature in medical images. Such techniques include contrast enhanced CT or MR angiography, laser speckle imaging [118], near infrared fluorescence imaging [187], color Doppler US and optical coherence tomography (OCT) [80]. Although these modalities and techniques enhance the appearance of the imaged

Table 1.1: Categorization and comparison between state-of-the-art methods for automatic vessel segmentation.

	Modality	Application	Method	Multi-Scale	Motion Model
Frangi et al. [49]	2D DSA & 3D MRA	Vessel enhancement	Vesselness	Y	N
Lorigo et al. [101]	3D MRA & 3D CT	Vessel segmentation	Active contours	N	N
Staal et al. [171]	RGB image	Segment retinal vessels	Vesselness	Y	N
Vermeer et al. [193]	2D image	Segment retinal vessels	AM	N	N
McIntosh and Hamarneh [107]	3D MRA & 3D CT	Vessel segmentation	Vesselness+ TSMT	Y	N
Law and Chung [98]	3D MRA	Detect curvilinear structures	OOF	Y	N
Řříha and Beneš [142]	DUS	Segment carotid artery	OF+MO+HT	N	Y
Schaap et al. [152]	3D CT	Segment coronary artery	Supervised AM	Y	N
Rigamonti and Lepetit [141]	2D RGB	Segment retinal vessels	Random forests	Y	N
Becker et al. [13]	2D RGB & 3D BM	Vessel segmentation	CNN	Y	N
Amir-Khalili et al. [5, 7]	EV	Segment renal vessels	PBMS	Y	Y
Hennersperger et al. [71]	2D US	Segment carotid artery	Vesselness+ AM+USCM	Y	N
McLeod et al. [111]	DUS	Detect dural pulsation	EKF+FS	N	Y
Gastounioti et al. [53]	DUS	Segment plaque	GWIR+GBS	N	Y
Gao et al. [52]	3D MR	Segment carotid artery and abdominal aorta	HT + NURBS	Y	N
Amir-Khalili et al. [6, 8]	DUS	Segment carotid artery	MF+PRMM	Y	Y

AM: Appearance Model
BM: Brightfield Microscopy
CNN: Convolutional Neural Network
CT: Computed tomography
DSA: Digital Subtraction Angiography
DUS: Dynamic Ultrasound
EKF: Extended Kalman Filter
EV: Endoscopic Video
FS: Frequency Smoothing
GBS: Graph-Based Segmentation
GWIR: Group-Wise Image Registration
HT: Hough Transform
MF: Monogenic Flow
MO: Morphological Operations
MRA: Magnetic resonance angiography
NURBS: Non-uniform rational B-spline
OF: Optical Flow
OOF: Optimally Oriented Flux
PBMS: Phase-Based Motion Segmentation
PRMM: Pulsatile Radial Motion Model
TSMT: Tubular Spring-Mass Tracker
US: Ultrasound
USCM: Ultrasound Confidence Maps

vasculature, many of the aforementioned clinical applications stand to benefit from a fully automatic vessel localization algorithm. The need for automated vessel localization, or segmentation, has inspired novel contributions in the field of medical image analysis. In the next subsection, we provide an overview of important contributions made in this field and survey the emerging trend of incorporating temporal information (kinematics) into automatic vessel segmentation algorithms.

1.4.1 Related Works

In Table 1.1, we have summarized notable contributions made towards fully automatic segmentation of vascular structures with an emphasis on seminal techniques and recent approaches that incorporate temporal motion models. The reader is referred to comprehensive surveys of vessel segmentation tech-

niques [89, 90, 99, 179] for more information on other existing methods.

The early attempts at automatic vessel segmentation focus on applying advanced low-level pixel based image analysis techniques to static intensity information acquired from the aforementioned imaging modalities. Such attempts include the exploitation of ridge-like features in the image [171], Hessian-based vesselness features [49, 71, 98], and model/physics based approaches [71, 193]. Other high-level techniques have also been proposed by embedding these low-level features in broader frameworks, which include: vessel trackers [107], deformable 3D cylindrical non-uniform rational B-spline (NURBS) surfaces [52], a combination of wavelet-based features and machine learning [168], active contours [101], and supervised machine learning techniques [13, 141, 152]. With the exception of Doppler US and OCT, the techniques listed above and in cited survey papers [89, 90, 99, 179] focus on extracting low- and high-level features from static information alone, ignoring the most characteristic feature of a pulsating vessel, i.e., its kinematics or temporal behavior.

On the other hand, US and OCT can exploit the pulsatile flow kinematics of blood inside the vessels to facilitate localization. Such modalities are capable of measuring the directionality and relative velocity of structures (usually blood) by leveraging the Doppler effect. The flow of blood, however, is not the only temporal characteristic of vascular structures. The pulsatile radial distension and compression of the vascular walls (from the lumen to tunica externa) is another characteristic that can be observed and measured using almost any imaging modality so long as the temporal and spatial resolutions are adequate.

The first use of temporal features for the purpose of vessel segmentation did not explicitly model the kinematics [142]. In their paper, the authors simply assumed that the only meaningful movement in DUS scan of the CCA imaged along the transverse axis is the pulsatile movement of a circular pattern. Based on this assumption the authors propose to use an optical flow (OF) sequence and simply average the absolute value (magnitude) of motion across the entire sequence to generate features. These features are then processed with median filtering and morphological operations (MO) to generate a binary mask. High-level features are finally extracted from the Hough transform (HT) of the binary mask and the resulting features, along with the last frame of the sequence, are fed into a Bayesian classifier to compute the center and radius of the CCA.

We initially proposed to exploit the kinematics of pulsating vessels in the context of kidney cancer surgery to identify major vessels that are hidden under layers of connective tissues [5, 7]. Rather than a simple computation of the average magnitude of motion using OF, we proposed the use of a temporal bandpass filter to isolate features that are in sync with the heart-rate. In our approach, we reformulated the Eulerian video magnification (EVM) [195] pipeline into a multi-scale phase-based motion segmentation (PBMS) algorithm to detect the motion of renal vessels by analyzing the magnitude of temporal change in the local phase information of an endoscopic video (EV) sequence. Our PBMS method, although novel in application, only operates on the magnitude of local pulsatile motion and is consequently prone to false positives when tested on other applications and imaging modalities, failing to differentiate between the motions specific to vasculature versus neighboring structures that happen to move at the same frequency as blood vessels. In more recent publications [6, 8], to reduce the number

of false positives and to extend the application of our method to more challenging imaging modalities, such as DUS sequences of the CCA, we proposed a novel kinematic model-based vessel segmentation (KMVS) pipeline that couples a pulsatile radial motion model (PRMM) with a more detailed computation of motion characteristics that entails an estimation of the local magnitude and orientation of motion. We showed that this updated pipeline increases the accuracy of kinematics-based vessel segmentation and that by reconstructing the monogenic signal [43] and computing the motion vectors using a monogenic flow (MF) technique, the local orientation of motion may be estimated in a more computationally efficient manner compared to the previous PBMS method.

Concurrent with our efforts, other novel methods have been proposed to address similar challenges with the help of pulsatile kinematics models. In a recent publication, it was demonstrated that a kinematic model of periodic low velocity out-of-plane motion of structures in DUS using extended Kalman filter (EKF) and frequency smoothing (FS) can localize dural pulsation for spine needle interventions [111]. The proposed method operates in real-time and is capable of detecting subtle motions that are imperceptible in Doppler US. Furthermore, the proposed visualizations were shown to reduce the normalizing path length and number of attempts required to perform a mock epidural procedure on a spinal phantom model. Although this method was shown to be effective in the novel application presented, similar to our PBMS method, it will likely not be able to distinguish between vascular structures and others that happen to translate at the same frequency as vessels. The FS aspect of the proposed method may allow the EKF approach to perform better than PBMS, but this method cannot benefit from an advanced kinematic model of vasculature due to the lack of a mechanism to account for the spatial orientation of motion.

In the context of CCA atherosclerosis assessment, another method was proposed to learn the kinematic dependencies between atherosclerotic and healthy vascular tissue in DUS by combining groupwise image registration (GWIR) with a graph-based segmentation (GBS) scheme [53]. Rather than implementing a physics-based kinematic model, the authors proposed a data-driven approach to learn a complex discriminative model. To do this, the magnitude of total vertical and horizontal displacements (MTD) are first computed for every pixel throughout the sequence using GWIR. Then, independent component analysis is used to identify the dominant and independent motion classes, which are used as a basis to which the MTD of each pixel is mapped using mutual information. A final mutual information value is assigned to a given pixel through majority voting. The likelihood of a pixel belonging to a binary class (healthy or atherosclerotic) given the final map is first learned and then used as the data term to perform GBS and generate contiguous contours around the atherosclerotic regions. Segmenting atherosclerotic plaque from DUS is challenging and the proposed pipeline performs well. It can be argued that the pipeline may be modified to segment vascular structures in addition to the plaque regions. Even though real-time performance is not a strict requirement for diagnostic clinical applications, the speed of the algorithm is of clinical value. The authors do not mention the runtime of their pipeline and the GWIR method used in the paper was projected to take minutes to complete, at best, if optimized and implemented in C++ [169]. It is thus unlikely that the proposed method would be able to perform in real-time.

We introduce our contributions in intraoperative segmentation of blood vessels from dynamic medical image data starting with Chapter 2, in which we present our automatic phase-based motion segmentation technique. Our proposed PBMS method can localize blood vessels that are hidden under layers of tissues (visually occluded) by estimating the magnitude of apparent motion via the change in local (spatial) phase information and measuring this change over multiple spatial scales and orientations to encode the motion information from neighboring pixels. In Chapter 3, we follow with the evolution of our methodologies that result in improved accuracy, increased computational performance, and broadened applicability to other medical imaging modalities. Our next-generation kinematic model-based vessel segmentation methods extend the computation of motion to include the orientation of motion, in addition to the magnitude of motion, and take advantage of a mathematical pulsatile radial motion model to localize vasculature.

1.5 Navigation Uncertainty in Image-Guided Therapy

In comparison to vessel segmentation, the computation and visualization of intraoperative guidance uncertainty during IGT is a relatively new topic of research that has recently found applications in orthopedic surgery [159, 160], image-guided neurosurgery [147], longitudinal studies of Alzheimer's disease [162], as well as pharyngeal [148] and prostate [119] radiotherapy. A method for visualizing the influence of propagated non-rigid registration uncertainties onto probabilistic segmentation has also been presented [163]. Similar visualizations are currently being explored in radiotherapy for brain tumors [148] and also for probabilistic extrapolation of glioma invasion with variable margins [92].

1.5.1 Sources of Navigation Uncertainty

Sources of uncertainties in computer-assisted navigation emanate from both preoperative and intraoperative stages of IGT. Among the contributing sources to navigation uncertainty, the uncertainties associated with the image segmentation stage of IGT are well known. Probabilistic segmentation of image data is not novel and many automated segmentation techniques are capable of producing fuzzy or probabilistic labels that represent the underlying uncertainties in the resulting segmentation output; more so now that statistical atlas-based [78] and machine learning [156] techniques for medical image segmentation are gaining in popularity. It is also possible to estimate segmentation uncertainties within semi-automated image segmentation frameworks that are intended for accurate segmentation of preoperative data [58]. Despite the availability of probabilistic segmentation methodologies, the segmented models produced for the purpose of IGT navigation are often converted to determinate or crisp labels during the preoperative planning and intraoperative scene augmentation.

Another important, and often ignored, source of uncertainty in many of the emerging image-guided surgical navigation frameworks stems from the extraction of 3D surface geometry from stereo endoscopic video. In the past, attempts have been made to model the encoding of 3D geometry in a pair of 2D images using Bayesian frameworks [14], probabilistic stereo reconstruction methods have been proposed [74, 87], and probabilistic scene analysis has also been used to detect smooth problem areas prior to matching [172]. These approaches merely leverage a probabilistic model to arrive at a globally

optimal depth reconstruction and, surprisingly, none of these approaches attempted to propagate the computed uncertainties to create a probabilistic representation of the extracted geometry.

We explore methods to estimate and visualize sources of uncertainties in Chapter 4, in which we present our preliminary phantom experimentation using an uncertainty-encoded navigation framework designed for image-guided minimally invasive robot-assisted partial nephrectomy (RAPN). The presented framework encodes uncertainties through the computation and visualization of uncertainties that may occur during preoperative CT segmentation and computational stereopsis steps of navigation.

Although the computation of segmentation and geometric uncertainties are important, the combined effects from these separate sources of error cannot be quantified in many IGT applications without their integration within a probabilistic registration framework. In fact, the lack of research into the propagation of uncertainty in IGT can be attributed in part to the need for an effective mathematical framework for computing and propagating the effects of uncertainty during the registration stage of IGT [159].

1.5.2 Uncertainty in Deformable Image Registration

Deformable image registration (DIR) algorithms are arguably the most important source of uncertainty in IGT navigation systems and, in such contexts, registration uncertainty (RU) is challenging to compute and visualize. Unlike simpler rigid and affine transformation models, DIR is fraught with challenges stemming from the complexity of parameter optimization, choice of similarity metric, and evaluation of accuracy and precision of resulting transformations [81, 170, 181]. Such challenges are exacerbated as transformation models become more complex to accommodate for more realistic deformations.

Despite the challenges associated with DIR, deformable models generally outperform simpler registration models since most organs inside the human body undergo complex elastic deformations. Inevitably, however, errors that occur during DIR would propagate through to all subsequent analyses performed post-registration. Examples include errors in augmented reality visualization during computer-assisted surgery [72], biased estimation of head movement in fMRI time-series analysis [51], and geometric uncertainties in localizing organ shape or motion during image-guided radiotherapy [30]. Additionally, DIR uncertainties also affect the decisions made by the end-users of DIR, which include both image analysts and physicians. An erroneous DIR may result in catastrophic outcomes for a patient during image guided therapies where DIR is often used for the purpose of guidance [40]. Furthermore, established methods for estimating the overall accuracy of registration processes from fiducial or target registration errors cannot be simply generalized from rigid to deformable transformation models [204]. Instead, in order to calculate the accuracy of DIR, the ground truth (GT) transformation, a dense 3D-3D mapping, is required.

Unfortunately, GT data for DIR are either scarce, non-existent, or impossible to obtain for most clinical applications. Obtaining GT data is particularly difficult in clinical applications where volumetric medical images from complementary imaging modalities are fused for the purpose of image-guided interventions or therapies. Such applications widely vary from multimodal imaging for radiation treatment [30] to computer assisted surgery [106]. The lack of a GT thus restricts the validation process to

testing on synthetically generated data or simplified approximations of DIR errors [19, 181].

Even though errors cannot be equated to uncertainty, the two have been found to be strongly correlated [75, 102] and RU thus remains as a valuable measure of quality for DIR. The importance of understanding and communicating the uncertainty associated with DIR software is a timely issue. Recently, the Therapy Physics Committee of the American Association of Physicists in Medicine (AAPM) commissioned a task group to review the current approaches and solutions for image registration in radiotherapy. Among their clinical recommendations, Brock et al. [19] advocated for a better understanding of the basic components of the registration algorithm; end-to-end tests of imaging, registration, and treatment systems using a physical phantom; and comprehensive commissioning of image registration using digital phantom data. Although we support these recommendations, we emphasize that phantom-based studies shed limited insights on the validity of DIR software, primarily due to the fact that phantom-based analyses are often an oversimplification of real world situations where noise, distortion, and complex anatomical variations typically occur.

1.5.3 Related Works

Our survey of the prior works relating to computation and propagation of RU is naturally divided into two parts. In the first part of this section, we survey and outline the limitations of current mathematical methods for estimating RU and, in the second part, we discuss notable works that propose application specific end-to-end frameworks for propagating the effects RU.

Registration uncertainty computation:

Current approaches to estimating RU may be divided into three groups: (i) characterizing uncertainty from contextual image information [182, 196], (ii) frequentist approaches involving multiple registrations [75, 76, 94, 199], and (iii) Bayesian approaches involving model inference on the posterior of the deformation parameters [147] and at the regularization level [161]. Among these methods, Watanabe and Scott [199] stands out as a suitable base for a solution in the context of IGT since: it can be implemented on top of existing DIR software; it is computationally more efficient than the Bayesian approach of Risholm et al. [148]; and it is capable of representing RU through ellipsoidal spatial confidence regions in the pixel-domain of the target image, which facilitates intuitive visualization and RU propagation. Despite these important advantages, some associated limitations exist which we address below.

It may be argued that a notable limitation of the cited frequentist methods, in contrast to the Bayesian approach of Simpson et al. [161], is that the computation of RU is evaluated from changes in image similarity subject to local random deformations without an explicit consideration for the global influence of regularization. Though the effects of regularization on RU have been shown to be significant, e.g., especially the case for complex intra-subject brain registration [164], the amount of computation required for inference on regularization parameters is costly. This cost hinders the applicability of such Bayesian methods in intraoperative applications. On the other hand, the frequentist RU computation methods are often pleasingly parallel and extendable to commonly used medical image registration software such as

`elastix` [91]. We elaborate on this point further in Section 5.3.5.

Another limitation of the frequentist approach of Watanabe and Scott [199], the context-based method of Wang et al. [196], and the generative Bayesian approaches of Risholm et al. [147] and Simpson et al. [161] is that they are designed for unimodal registration tasks; wherein two images from the same modality, or even from the same patient, are registered together. In Watanabe and Scott [199] for example, RU is estimated from intensity information of the moving image alone, which may not be representative of the true RU. Extending the generative Bayesian RU estimation approaches to multimodal DIR is also challenging as doing so would essentially require a model for generating the target imaging modality from the moving imaging modality.

Registration uncertainty propagation:

Compared to the amount of literature on RU estimation methodologies, research into RU propagation in an end-to-end fashion is surprisingly scarce. Within the broad context of fractionated radiotherapy, which also includes external beam radiotherapy, the effects of DIR error propagation have been mainly studied on dose accumulation. Most notably, the Bayesian RU estimation method of Risholm et al. [147] was applied to oropharyngeal radiotherapy in Risholm et al. [148], while Murphy et al. [119] and Tilly et al. [186] implemented a frequentist approach to quantify RU during radiotherapy of the prostate. In Murphy et al. [119], principal component analysis was applied to multiple registrations of CT image pairs to obtain decorrelated modes of error, from which sample deformations (used to compute RU) were drawn. In Tilly et al. [186], a synthetic simulation framework was used to study the sensitivity of planning parameters to DIR. Radiation planning parameters are typically derived from dose-volume histogram (DVH) that provide a summary of absorbed radiation over the entire volume of a structure, which include the target volume and the organs at risk (OAR) [120, 133]. Cumulative DVH is important for planning and postoperative analysis as it has been shown to correlate with patient toxicity outcomes [55]. For hollow elastic OAR, in addition to DVH parameters, it is also important to study the spatial distribution of the accumulated doses as it is indicative of the formation of radiation hot spots and, thus, potential resulting complications [209]. To the best of our knowledge, the effects of RU have only been studied on the entire dose volume [147] or the resulting DVH. There is still a need for planning parameters that can capture the spatial distribution of total radiation dose while accounting for the quality of DIR.

Given the challenges described above, there remains a need for a RU estimation and propagation solution that is supported by mathematical formalism while being implementable as an end-to-end framework in real world applications. To address the need for RU-encoding methodology, in Chapter 5, we present a mathematical framework for estimating RU and propagating the effects of the computed uncertainties from the registration stage through to the proceeding dosimetric evaluations and visualizations.

1.6 Thesis Contributions

Our peer-reviewed contributions presented in this thesis are organized into four chapters; the first two of which pertain to our first research question, the latter two to the second.

1.6.1 Motion-Based Localization of Vasculature

Our methods presented in Chapter 2 and Chapter 3 demonstrate how motion cues may be extracted and used to locate blood vessels in IGT applications. In support of our contributions, we provide publicly available MATLAB executables[1] of our PBMS and KMVS methods to allow our fellow researchers to evaluate and incorporate our proposed methodologies within other application domains.

Phased-Based Motion Segmentation of Occluded Vasculature

We propose an automatic PBMS method, which leverages subtle motion cues from medical video data to localize blood vessels that are hidden under layers of connective tissues [5, 7]. We also present:

- Evaluations of our PBMS method on a retrospective study of fifteen clinical RAPN procedures. To the best of our knowledge, we are the first to attempt the task of localizing occluded vasculature in endoscopic video without the use of additional hardware or preoperative scans. In this challenging context, we demonstrate quantitatively promising vessel localization performance, i.e., a mean area under the receiver operating characteristics curve (AUC) of 0.72.
- Evaluations of our high-level variational scene segmentation method [126] (which integrates our PBMS alongside other image-based cues and patient-specific priors, i.e., shape and deformation) and demonstrate a 45% increase in pixel-wise accuracy (for localizing renal vasculature in context of RAPN) compared to our original PBMS method.
- A preliminary clinical user study involving four surgeons and our findings regarding how our PBMS visualization techniques may be improved in the future.

Kinematic Model-Based Motion Segmentation

We propose KMVS [7, 8], an extension of PBMS, that is designed to localize vasculature from dynamic medical image sequences by leveraging: (i) the estimation of local motion vectors and (ii) a novel PRMM that enables the modelling of divergent (radially moving) motion patterns. We also present:

- Implementation of a parallelizable technique for the computation of motion vectors, which estimates motion via the changes in the monogenic representation of image information.
- Four alternative implementations of our KMVS method using different motion computation techniques and discussions regarding the advantages of the different implementations.
- Evaluations of the four implementations of our KMVS method on a synthetic dataset and two real DUS datasets of the CCA and report differences in performance, in terms of segmentation accuracy and computation time, compared to the PBMS method. Compared to PBMS, our fast tuned optical flow implementation increases the average AUC from 0.82 to 0.99 on our in-house data and from 0.83 to 0.98 on a publicly available dataset.

[1] MATLAB executables are available for download from https://bisicl.ece.ubc.ca/software/radialDistension.html

1.6.2 Computation, Propagation, and Visualization of Navigation Uncertainties

In Chapter 4 and Chapter 5, we present our contributions towards automatic augmentation of intraoperative data through the computation, propagation, and visualization of different sources uncertainties that result from the use of computer-assisted surgical and radiological guidance tools.

Uncertainty-Encoded Augmentation of the Surgical Scene

We propose an endoscopic scene augmentation method for facilitating the registration of probabilistic preoperative CT segmentations with stereo endoscopic video data [4]. We also propose:

- An uncertainty-encoded computational stereopsis technique for extracting probabilistic surface information from stereo endoscopic data.
- Application of our framework to an *ex vivo* lamb kidney phantom to simulate the tumor demarcation stage of RAPN interventions.
- Uncertainty-encoded visualization techniques for depicting probabilistic tumor margins onto the endoscopic scene and discussions regarding the potential advantages of our proposed visualizations compared to existing crisp (deterministic) techniques.

Encoding Deformable Image Registration Uncertainties for Scene Augmentation

We propose a mathematical framework for estimating RU from DIR and subsequently propagating the effects of the computed uncertainties from the registration stage through to the visualizations, organ segmentations, and dosimetric evaluations [9]. We also propose:

- A method for computing RU that is designed to: (i) interface with existing multimodal DIR software, which we deploy using `elastix`, and (ii) represent RU in a parametric manner using structure tensors.
- A weighted averaging technique for propagating the effects of RU, onto volumetric segmentation and dose data, to produce a probabilistic map of aligned segmentation and dose information subject to the estimated RU.
- Evaluation of our framework on a retrospective study consisting of 37 multi-fraction cervical cancer brachytherapy (MFCCB) patients, in context of which we present preliminary evidence that our proposed framework may be advantageous. Specifically, we show that (i) the effect of RU on dose accumulation provide useful insights for quality control and post-treatment analysis; (ii) RU propagation improves the transfer of delineations from one fraction to the next; and (iii) RU can be used to generate visualizations that reflect the quality of DIR that may prove to assist physicians in making decisions based on registered image data.

Chapter 2

Phase-Based Motion Segmentation of Occluded Vasculature

"We, on the other hand, must take for granted that the things that exist by nature are, either all or some of them, in motion."
— Aristotle

In this chapter, we present our contributions towards automatic segmentation of visually occluded vasculature from video data. The methodology presented herein was originally published in Amir-Khalili et al. [5, 7], Nosrati et al. [126]. To aid in vessel discovery, in Amir-Khalili et al. [5, 7], we proposed a novel automatic method to segment hidden vasculature by labeling minute pulsatile motion that is otherwise imperceptible to the naked eye. Our segmentation technique extracts subtle tissue motions using a technique adapted from phase-based video magnification [195], in which we measure motion from periodic changes in local phase information. Based on measuring local phase through spatial decomposition of each frame of the endoscopic video using complex wavelet pairs, our approach assigns segmentation labels by detecting regions exhibiting temporal local phase changes matching the heart rate. Our proposed phase-based motion segmentation (PBMS) method presented in this chapter is extended in Chapter 3 to increase its specificity to outliers.

2.1 Localizing Vasculature using Temporal Information

Periodic pulsations of major blood vessels are within a narrow temporal passband centered around the heart rate of the patient. With high definition surgical video data, one can observe the pulsations of the vessels as faint movements on the surface of the connective tissue that covers them. Our goal is to automatically process every frame in the surgical video and label pixels that exhibit this characteristic motion. We denote our labels

$$L(\mathbf{x},t) : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow l \in [0,1], \tag{2.1}$$

where l is a normalized fuzzy value that is proportional to the magnitude of pulsatile motion measured at the pixel $\mathbf{x} = (x_1, x_2)^\intercal, \mathbf{x} \in \Omega$ in the 2D spatial domain $\Omega \subset \mathbb{R}^2$ at a given point in time $t \subset \mathbb{R}^+$. Similarly,

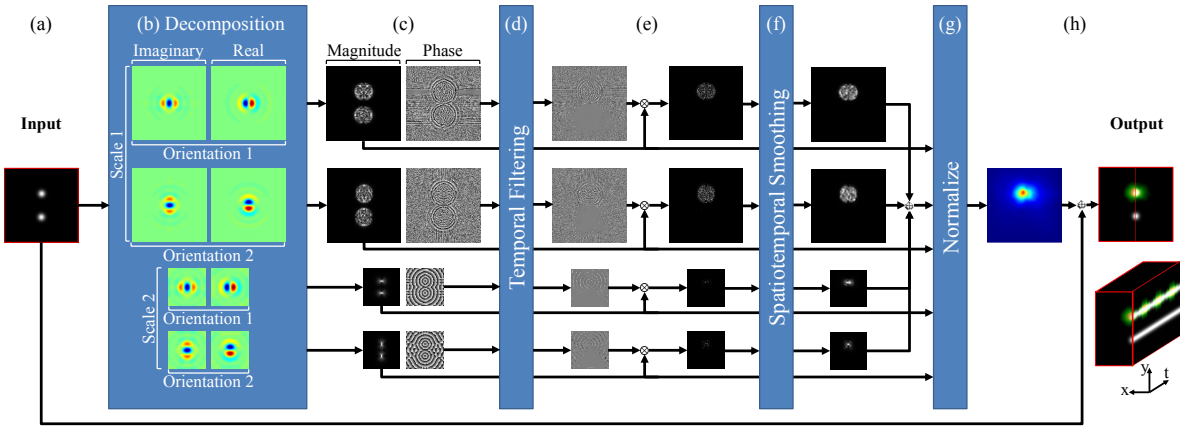


Figure 2.1: Overview of our proposed method: (a) a synthetic input video composed of one circle that pulsates (top) and another that remains stationary (bottom). (b) Steerable filter bank with illustrated impulse responses decompose the information inside each frame into (c) magnitude and local phase at different scales and orientations. (d) The phase information of all frames of the video is temporally filtered using an ideal bandpass filter centered on the frequency of the pulsating circle. (f) A median filter is applied to (e) the magnitude weighted response of the filtered phases to remove phase noise. (g) The results are then combined to generate the fuzzy labels L and (h) added back to the input video as an overlay. A spatiotemporal cross section of the video (lower right) illustrates four pulsations across 30 frames of the synthetic video. High resolution images are available in the digital copy.

the color endoscopic video signal is defined as

$$\mathbf{V}(\mathbf{x},t) : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbf{v} \subset \mathbb{R}^3, \tag{2.2}$$

in the continuous time domain, where $\mathbf{v} = (v_r, v_g, v_b)^\intercal$ represents the vector encoded values of red v_r, green v_g, and blue v_b color channels at every pixel $\mathbf{x}$. In the discrete time domain, each j^{th} frame $j \subset \mathbb{N}$ of a video recorded at $1/T$ frames per second is defined as $\mathbf{V}(\mathbf{x}, jT) = \mathbf{V}(\mathbf{x},t)$. The motion estimation techniques presented in this thesis, extract motion from the scalar valued (grayscale) representation of the video defined hereby as a function

$$f : (\mathbf{x} \subset \mathbb{R}^2, j \subset \mathbb{N}) \rightarrow \mathbb{R}, \tag{2.3}$$

mapping a pixel $\mathbf{x}$ in the 2D spatial domain of each frame j to an intensity value $f = 0.299v_r + 0.587v_g + 0.114v_b$. In the remainder of this section, we elaborate on our methodology for generating segmentation labels L from local phase measurements, starting with the relationship between phase and motion.

2.1.1 Extracting Motion-Based Cues from Time Varying Local-Phase Information

The shift property of the Fourier transform, $f(t-t_0) \Leftrightarrow F(\omega)e^{-i\omega t_0}$, states that motion is related to the change in phase, however the explicit computation of motion vectors from phase with techniques

like Gautama and Van Hulle [54] is computationally expensive. To avoid this cost, we aim to exploit the relationship between phase and motion to segment the pulsating regions of interest in the video sequence. The overview of our method is shown in Figure 2.1 and it is described in the following 1D intuitive manner, without loss of generality. First, consider a video, a simplified version of the one shown in Figure 2.1a, denoted as $f(x,t)$ representing a function mapping given pixel $x \subset \mathbb{R}$ at time t to an intensity value. Suppose that this video is a sequence composed of the 1D image $f(x)$ that is translated by a displacement vector $d_x(t)$ along the x direction as a function of time t such that $f(x,t) = f(x+d_x(t))$. To extract this motion, we decompose each frame of the video into spatial sub-bands

$$f(x,t) = f(x+d_x(t)) = \sum_{\omega=-\infty}^{\infty} A_\omega e^{i\omega(x+d_x(t))} \tag{2.4}$$

with each sub-band representing a complex sinusoid $S_\omega(x,t) = A_\omega e^{i\omega(x+d_x(t))}$ at spatial frequency ω. The phase of each sub-band is defined as $\phi_\omega(x,t) = arg(S_\omega) = \omega(x+d_x(t))$. Since $\omega d_x(t)$ is the only motion related component of the phase that varies with time, we can isolate it from the DC component ωx by applying a DC-balanced temporal bandpass filter with a wide enough passband to capture all temporal variations in $d_x(t)$.

Multi-Scale Steerable Analytic Decomposition

Generally, motions in a video are not merely a simple global translation. The displacement vector $d_x(t)$ is, in fact, $d_x(x,t)$ since it varies as a function of both time and space. In 1D, local phase can be measured by constructing the analytic signal. The analytic signal is constructed from quadrature filters, i.e., 1D Hilbert pair of bandpass filters. The estimation of local phase is more complex in 2D images and there are thus many approaches to extend the analytic signal to 2D. One approach is to use a steerable complex pyramid decomposition [132] to extract local motion information from a sequence of grayscale images. To measure these spatially localized variations, rather than using a Fourier series expansion, we decompose each frame of the video using a spatial filter bank (pyramid) consisting of a cascade of Gabor odd and even symmetric filter pairs (analogous to Hilbert transform filter pairs in 1D) with limited spatial support (note that the impulse responses of these wavelets have been enlarged for clarity in Figure 2.1b). In the steerable pyramid, the spatial extent of each filter is determined by the scale or spatial passband of the Gabor wavelets and, at each scale, the filters are designed to measure motion along a certain direction or orientation in 2D space (note that in Figure 2.1b, a pyramid consisting of two scales and two orientations is used). In the 1D example, if the local motion is from a single sinusoid with spatial frequency ω along the x direction, we would only need a single pair of Gabor wavelets to extract the motion from the change in local phase.

With the complex steerable pyramid, the analytic signal is estimated at different scales $s = \{1,...,S\}$, along $n = \{1,...,N\}$ different orientations from the complex response $\mathbf{h}(\mathbf{x},j;s,n)|_{s,n} : \mathbb{R}^3 \rightarrow \mathbb{C}$ to a set of steerable filters $\mathbf{b}(\mathbf{x};s,n)$. The real and imaginary parts of $\mathbf{h}(\mathbf{x},j;s,n)$ correspond to a pair of even- and odd-symmetric filter responses that are analogous to a one dimensional Hilbert transform along the given orientation. The orientations are sampled evenly such that the local orientation θ_n (where $\measuredangle\mathbf{x} =$

$(\cos\theta_n, \sin\theta_n)^\mathsf{T}$) is determined by $\theta_n = \pi n/N$, where N is the total number of orientations used in the pyramid. This steerable method measures the magnitude of local phase $\phi(\mathbf{x}, j; s, n) = \arg(\mathbf{h}(\mathbf{x}, j; s, n))$ projected onto N different angles $\theta_n, n = \{1, ..., N\}$.

Spatiotemporal Filtering

Local, or instantaneous, phase is calculated from the argument of the response to the wavelet pair (Figure 2.1c). We then estimate local motion $d_x(x,t)$ from the change in local phase by applying a DC gain-balanced temporal bandpass filter to the obtained local phase values. We filter the local phase measurements using an ideal bandpass filter:

$$z(j) = 2\tau_H \texttt{sinc}(2\tau_H j) - 2\tau_L \texttt{sinc}(2\tau_L j), \tag{2.5}$$

where τ_L is the temporal low frequency cut-off and τ_H is the high frequency cutoff and the `sinc` functions are the time domain representations of `rect` functions in the temporal frequency domain that construct an ideal bandpass. The response of the temporal bandpass filter is $\phi_z(\mathbf{x}, j; s, n) = \phi(\mathbf{x}, j; s, n) * z(j)$. We tune the passband of the filter to the typical heart rate of a patient so that we then can simply and effectively isolate components of the local motion that are synchronous with the heart rate and hence to vascular pulsation. In this work we have set the passband of the temporal filter wide enough such that it can separate pulsatile motion from breathing motion in all of the fifteen cases. Future development should involve a tighter estimate of the patient's heart rate to improve the results. Such estimates may be recorded directly from the patient's heart rate monitor or obtained from the anesthetist.

To generate fuzzy segmentation labels from the computed local motion, ϕ_z is first attenuated in regions where the magnitude response (A_ω in the 1D case) of the spatial sub-band is weak. This is done by computing the product between the bandpassed phases and the normalized magnitude of the spatial filter response vectors $|\mathbf{h}(\mathbf{x}, j; s, n)|$ to obtain $\hat{\phi}_z(\mathbf{x}, j; s, n) = |\mathbf{h}(\mathbf{x}, j; s, n)| \phi_z(\mathbf{x}, j; s, n)$ (Figure 2.1e). Since local phase measurements ϕ_ω are wrapped between the interval $(-\pi, \pi]$, and since z in Equation 2.5 acts as a derivative, the jumps in wrapped phase become impulse noise in $\hat{\phi}_z$. We remove this noise from the product Q_ω using a spatiotemporal median filter (Figure 2.1f). For faster performance, the spatiotemporal median filter is replaced with a spatial pseudo-median filtering process using efficient 2D morphological opening $\circ$ and closing $\bullet$ operations as follows

$$\tilde{\phi}_z = \hat{\phi}_z \circ E + \hat{\phi}_z \bullet E - \hat{\phi}_z, \tag{2.6}$$

where E is a 2×2 square structuring element.

Multi-Scale Motion-Based Segmentation

The denoised product $\tilde{\phi}_z(\mathbf{x}, j; s, n)$ is averaged across all spatial sub-bands (Figure 2.1g) and all filter orientations to obtain our final fuzzy labels

$$L = \frac{1}{M} \sum_{\forall s,n} \frac{|\hat{\phi}_z(\mathbf{x}, j; s, n)|}{2\pi\omega_s}, \tag{2.7}$$

where ω_s is the spatial frequency of scale s and M is a normalizing factor to fix the range of phase-based motion segmentation (PBMS) labels $L \in [0, 1]$. The resulting sequence of fuzzy labels L may be displayed as an overlay or separately to highlight this pulsatile motion (Figure 2.1h). In a real-time application, the ideal temporal filter z may be replaced with an infinite impulse response filter.

2.2 Automatic Vessel Localization during Robot Assisted Partial Nephrectomy

Approximately 64,000 new cases of kidney cancer, commonly renal cell carcinoma, were projected to occur in the U.S. in 2017 [158]. This constitutes double the number of cases reported in 2005 and has not changed since 2014 [157]. Kidney resection, also known as a nephrectomy, remains the only known effective treatment for this type of localized cancer [33]. Robot-assisted partial nephrectomy (RAPN) refers to nephron-sparing techniques performed with surgical robots in which only the cancerous cells are excised and the kidney is reconstructed to retain functionality.

The RAPN procedure is organized into five main stages according to Gill et al. [57]: (i) Bowel mobilization; (ii) Hilar dissection and control; (iii) Identification and demarcation of tumor margins; (iv) Resection of tumor; and (v) Reconstruction of the kidney (renorrhaphy). Hilar dissection stands out as a daunting stage requiring significant expertise since improper clamping due to overlooked accessory renal vessels can cause significant bleeding during resection [165].

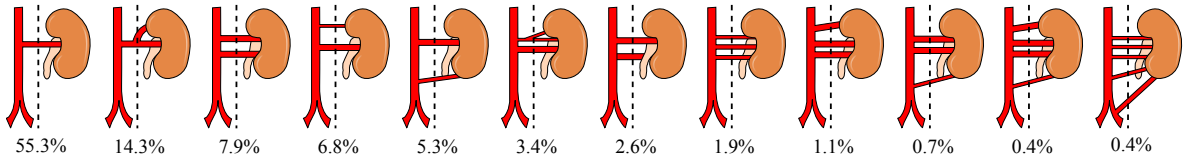


Figure 2.2: Variation of renal artery structure and corresponding percentage of occurrence in 266 kidneys adapted from [151]. In each case all vessels that cross the dotted line must be clamped or ligated to minimize intraoperative hemorrhaging.

Hilar dissection is a delicate procedure during which the surgeon dissects through the Gerota's fascia and removes the connective tissue that surrounds the renal artery (RA) and renal vein (RV). This task is complex due to substantial natural variability in patient vasculature (Figure 2.2) and the amount of perinephric fat surrounding the kidney. Access to the hilum grants the surgeon control over the flow of blood into and out of the kidney, which is critical as warm ischemia is required during the excision of the tumor to minimize internal hemorrhaging. In some cases, accessory vessels that

branch off from the RA or the abdominal aorta (AA) are accidentally missed as they lie hidden behind a thick layer of perinephric fat. In one study of 200 laparoscopic partial nephrectomy cases by world leading surgeons [140], seven incidents of intraoperative bleeding were reported as a result of inadequate hilar control, two of which were directly caused by missed accessory vessels. Although the number of incidents is relatively low, other studies by [151, 192] observed the existence of accessory vessels in more than 35% of patients. These accessory vessels also prolong the hilar dissection stage as the surgeon must locate them prior to resection. If the surgeon's level of experience is limited, the incidence of bleeding and overall dissection time may be much higher. The implications are many, aside from obvious complications that would arise from internal hemorrhaging, as bleeding may also jeopardize the surgical outcome by occluding the surgeon's view while the tumor is being resected.

Nephrogenesis and Kidney Migration: Development of Accessory Renal Arteries

Kidneys are primary retroperitoneal organs developed from intermediate mesoderm. Kidney development, also called nephrogenesis, proceeds through a series of three successive mesenchyme-to-epithelial transformation phases: pronephros, mesonephros, and metanephros. These three phases follow a craniocaudal developmental cascade starting with the development of the pronephros in the neck region of the embryo. The pronephros extends from the sixth to the fourteenth somites and consists of 6-10 pairs of tubules [31]. These tubules spill into a pair of primary ducts that are formed at the same level and extend caudally into the cloaca. The pronephros is a vestigial structure that is nonfunctional in mammals and disappears completely by the fourth week of human embryonic life as the mesonephros develops. Mesonephros develops by the formation of mesonephric tubules from the intermediate mesoderm and it is the principal excretory organ during early four to eight weeks of embryonic life. It gradually degenerates, although parts of its duct system (Wolffian duct) become associated with the male reproductive organs [180]. Metanephros arises caudal to the mesonephros at five weeks of development and it ultimately serves as the permanent and functional filtration components of the kidneys. The ureteric bud arises as a diverticulum from the Wolffian duct, close to the entrance to the cloaca and grows towards the metanephric mesenchyme. As the cephalic end of the ureteral bud grows inside metanephric mesenchyme, it expands within the growing mass of metanephrogenic tissue to form the renal pelvis and the primary collecting ducts of the kidney [180].

As the kidney develops in the elongating fetus, it moves cephalad relative to the bladder to its mature location (in the retroperitoneum just caudal to the diaphragm). During its migration, the kidney takes new arterial supply from the aorta and new venous drainage into the inferior vena cava (IVC). Occasionally, caudal branches of these vessels persist as the kidney ascends. These persisting branches form accessory renal arteries (Figure 2.2). Accessory renal arteries may arise from the aorta adjacent to the main RA, distal to the ostium of the main RA, or even from the iliac artery.

According to Chavan et. al. [26], when multiple arteries occur, each artery supplies a distinct segment of the kidney. As there is no collateral perfusion, occlusion of one artery will result in infarction of the associated kidney segment. Multiple renal veins, draining into the IVC, are almost as frequent as multiple renal arteries on the right side and are infrequent on the left side. Unlike the arteries, how-

ever, the renal veins interconnect (anastomose) within the kidney. If one renal vein is occluded, the remaining renal veins will continue to drain the entire kidney. Because of this fact, aside from the context of RAPN, the identification of accessory renal arteries is extremely important to kidney transplant procedures [22, 67].

Locating the Renal Hilum and Accessory Renal Arteries

Similar to the kidneys, the AA, IVC, and branching renal vessels are also primary retroperitoneal structures. Depending on the patient's amount of visceral adiposity, the vasculature at the renal hilum may be difficult to locate as all of the associated structures lie behind the peritoneum on the posterior abdominal wall.

There are two established approaches to perform RAPN interventions [57], the choice between which depends on location of the cancerous mass. Posterior or posterolateral tumors are approached retroperitoneoscopically, while anterior, anterolateral or lateral tumors are approached transperitoneally. Additionally, upper pole apical tumors are better approached by transperitoneal laparoscopy. In RAPN interventions, the transperitoneal approach is more common than the retroperitoneal approach [41].

Locating the renal hilum retroperitoneoscopically is further complicated by the fact that the retroperitoneal space is relatively small during the hilar dissection stage of the intervention. If the renal hilum cannot be located during the retroperitoneal approach, the scope is reinserted to identify the psoas muscle. The psoas muscle is then crossed from lateral-to-medial in a cephalad direction and a search is conducted for arterial pulsation near its medial border. Pulsations of the fat-covered RA is usually identifiable on the surface of the peritoneum [122]. While performing this search, some of the more prominent accessory arteries (such as aberrant branches from the aorta, superior messnteric or iliac arteries) may be identified.

In comparison, locating the hilar vessels transperitoneally is relatively easier. If RAPN is being performed on the left kidney, the dissection is pursued cephalad along the gonadal vein and the ureter, which run parallel and anterior to the psoas major muscle. By following them cephalad, the renal hilum can be identified by the deep pulsations of its artery under the renal pole. The gonadal vein typically drains directly into the left renal vein, further aiding with localization and dissection. On the right, the IVC is first identified and the dissection is then pursued cephalad until either the ureter or the right renal vein is exposed. The right renal hilum can then be identified from the pulsations.

As neither approach guarantees localization of accessory vessels, surgeons often make use of preoperative medical images for identifying troublesome accessory vessels [117]. Contrast enhanced angiography images are usually required as part of surgical planning for kidney transplant procedures, but are very rarely acquired prior to RAPN.

Even with high-resolution scans and segmented preoperative plans available to them, surgeons are still burdened with the complex and error-prone task of mentally transferring these abstractions onto the surgical site during the operation. Reducing the difficulty of navigation has been attempted by various approaches that rely on multi-modal registration to align the preoperative surgical map of the vessels onto the surgeon's endoscopic view, e.g., Amir-Khalili et al. [4], Estépar and Vosburgh [38], Hamarneh

et al. [64], Nosrati et al. [124], Pratt et al. [135], Puerto-Souza et al. [137], Su et al. [177], Teber et al. [185]. Registering and augmenting preoperative segmentations into intraoperative video is an excellent idea. However, such techniques have major limitations including selection of sensitive parameters [124], manual alignments [135, 177], use of invasive fiducials [185], and high computational complexity that prohibits practical real-time operation [4, 38, 64, 137]. These limitations stem from the difficulty of registering intraoperative video data with 3D anatomy that deforms and changes due to factors such as cutting, retraction, and breathing. Furthermore, these methods do not specifically focus on augmenting the location of vasculature.

Recent methods that focus more specifically on the detection of vasculature include the use of hardware solutions such as near infrared fluorescence imaging [187] or algorithmic methods that only use color intensity information from the endoscope to highlight vasculature based on perfusion models [29]. Solutions that use near infrared fluorescence are not widely accessible as they are cost restrictive, requiring additional equipment and expert clinicians to control the dosage of fluorescent agents. On the other hand, simple algorithmic methods fail to identify vessels that are hidden under a layer of fat.

Intraoperative ultrasound (US) imaging is another hardware solution employed during the tumor demarcation and excision stages of RAPN; mainly to resolve uncertainties in the location of tumor boundary [57]. Recent advancements in the field of US imaging, i.e., 'pick-up' transducers [153], motivate the use of US during hilar dissection, but such US guidance techniques also incur additional costs in terms of an increase in required personnel (as some robotic surgeons are not trained to operate and interpret US), operating time, and equipment upkeep. Even with Doppler US imaging, the localization of complex vascular structures is further ameliorated by the fact that the laparoscopic US probes currently available in the market can only acquire 2D images and, depending on the availability of picture-in-picture visualization, the surgeon may have to look at a separate screen to view the US images.

In summary, we have established that hilar dissection is a critical stage during RAPN, often complicated by the presence of accessory renal vasculature. The presence of accessory renal vessels is common as they are caused by natural variations during fetal development. Current laparoscopic hilar dissection techniques employed during RAPN interventions do not focus on the localization of these accessory vessels. Furthermore, existing vascular imaging techniques are rarely used due to prohibitive costs and additional risks to patients. There remains a need for cost effective alternatives to current methods of imaging and locating renal vasculature. As a result, the application of vessel localization during the hilar dissection stage of RAPN is a suitable and challenging application for the evaluation of our proposed phase-based motion segmentation (PBMS) methodologies.

2.3 Experiments

In this section, we detail the *in vivo* dataset and the parameters used for the qualitative, quantitative, and clinical evaluation of our method. Results are illustrated *in situ* following the description of the experiments in each subsection. Extended discussions of all experimental findings are carried out in the next section.

2.3.1 Experimental Setup and Data Acquisition

Video sequences from fifteen clinical RAPN interventions were used for validation. All endoscopic video data were acquired by a da Vinci Si Surgical System. High-definition (1080i) videos were downsampled to 480×270 pixels, to reduce computation times and memory requirements of the overcomplete wavelet representation. The publicly available code from Portilla and Simoncelli [132] was used to generate sixteen complex steerable filter pairs (four orientations at four scales). The number of scales was set to four since, at the downsampled resolution, this number restricts the spatial extent of the wavelets such that most of the large structures in the videos are detected without blurring the motion of the smaller structures. In Section 2.3.2, we point to cases where the detection can be improved with a higher number of scales. The number of orientations was chosen such that diagonal motion is detected more accurately without increasing the computational and memory complexity of the algorithm. Increasing the orientations would improve the detection but it would increase the over completeness of the pyramid representation at a faster rate than an increase in number of scales. Readers are referred to Wadhwa et al. [195] for more details regarding over completeness.

The passband of the temporal filter was set between $\tau_L = 60$ to $\tau_H = 120$ beats per minute. Average runtime of our unoptimized MATLAB code to process these four second clips (120 frames) was 45 seconds (35 seconds with the pseudo-median filter). All results shown were obtained using the spatiotemporal median filter with a window size of $3 \times 3 \times 3$.

To provide an objective framework for validation, we compared the segmentations obtained through our guidance system against the manually obtained ground truth vasculature. To generate the ground truth, we segmented the kidney, tumor/cyst, AA, IVC, RA, RV, and accessory vessels (Figure 2.3) from the preoperative CT data using the publicly available ITK-SNAP semi-automatic segmentation tool [207]. The resulting meshes were then manually aligned onto the first frame of each endoscopic scene (Figure 2.4a) by rigidly transforming the models using a 6-degrees of freedom (DOF) 3D mouse to adjust all DOF contemporaneously. Anatomical landmarks such as the contour of the kidney, visible parts of the vessels, tumor, liver, spleen, ribs, and the surgical dissection planes were used to guide the registration process. Initial alignments were made by a graduate student with extensive knowledge of renal anatomy, and finalized by an early career urologist to ensure correctness. The segmentations and alignments were done prior to performing the vascular motion segmentation. Examples of the registered ground truths are presented in Figure 2.4b. Small observable discrepancies between the aligned ground truth segmentation and the endoscopic view are attributed to non-rigid deformations of the organs and vasculature caused by deformation during insufflation, retraction, or the mobilization of organs during the dissection, which our rigid registration does not take into account. By comparing the observable discrepancies against the known size of visible structures obtained from the CT images, we estimate an average of 1–3 mm of alignment error in most cases and a maximum of 4–7 mm in cases where organs have been considerably retracted by the surgical instruments or mobilization of other organs such as in Cases 2, 3, 4, 5, 9, and 14.

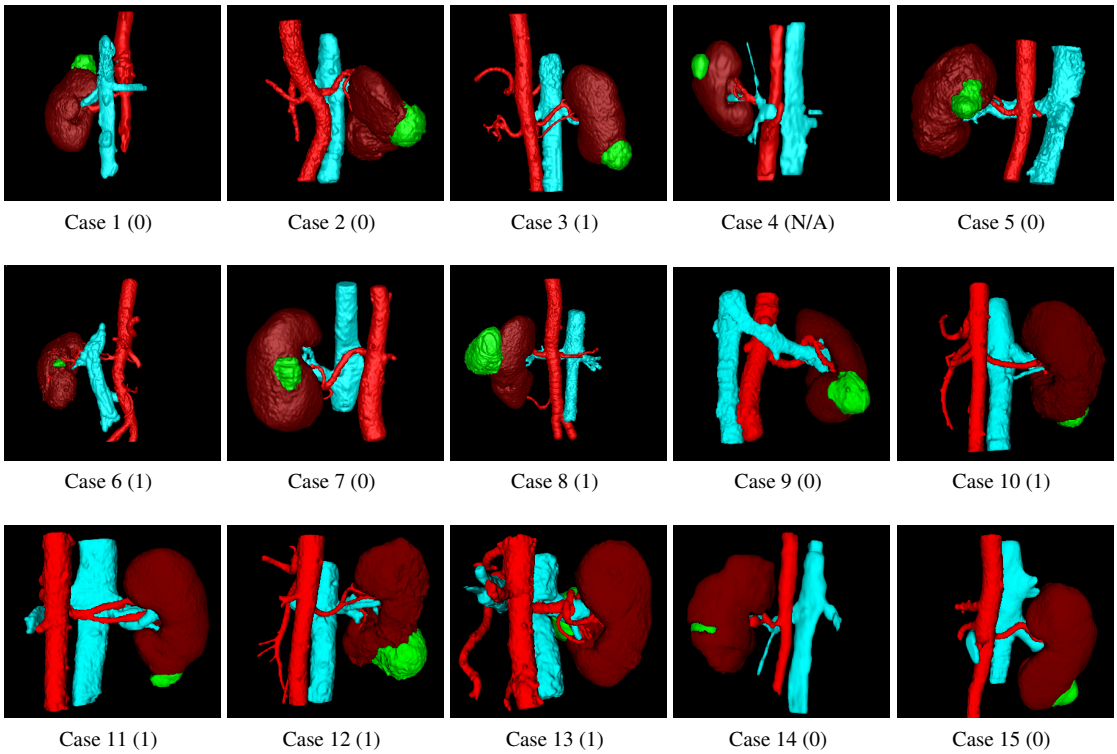


Figure 2.3: Manually segmented computed tomography (CT) scans of each surgical case with the number of accessory vessels included in parenthesis, showing kidney (brown), tumor/cyst (green), veins (cyan), and arteries (red).

Data Selection Criteria

The sequences that constitute our dataset, represent patients with varying visceral adiposity, RENAL Nephrometry Score (e.g. cyst or tumor, endophytic or exophytic), and vasculature. To explore the advantages and limitations of our method, we included challenging conditions, such as heavy presence of specular noise (Cases 1, 2, 5, 9, 10, and 12), the endoscope being close (Cases 3, 10, and 14) or far (Cases 1, 5, and 6) from the tissue surface, vessel occlusion by other organs (Cases 4, 6, 8, and 10), retraction of blood vessels (Cases 3, 5, 9, 13, and 14), and tool motion (Cases 7 and 12). Though not all the clinical examples were successful (Cases 7 and 8), the cases cover a wide range of typical scenes and potential problems that can arise.

2.3.2 Localizing Hidden Vasculature using Motion-Based Cues

In Figure 2.4 we illustrate the qualitative performance of our proposed method. In our experiments, we observed that although venous and arterial structures pulsate at the same frequency, their pulsations are not always in-phase. In fact, the temporal resolution of the surgical endoscope (30 frames per second) was able to observe the motion of the IVC and RV an average of six frames ahead of the AA and RA. To present this temporal phenomenon in Figure 2.4, we manually extracted two frames of the segmented

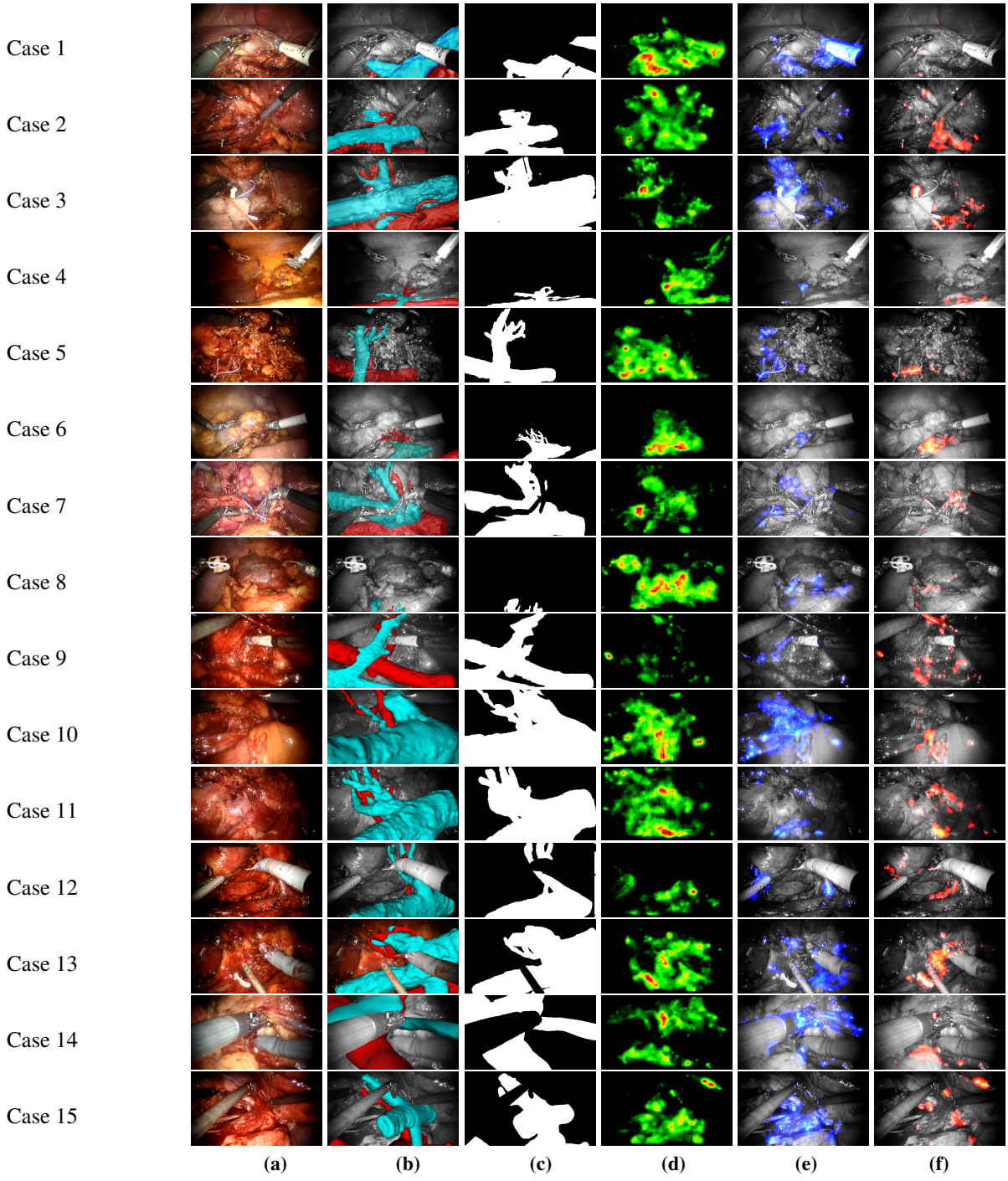


Figure 2.4: Exemplar video frames with the proposed automatic localization of veins and arteries showing: (a) the first frame of the sequence, (b) manually localized venous (cyan) and arterial (red) structures, (c) the binary mask used for quantitative evaluation, (d) temporal mean of the segmentations, and exemplar frames at the time of (e) venous and (f) arterial pulsation.

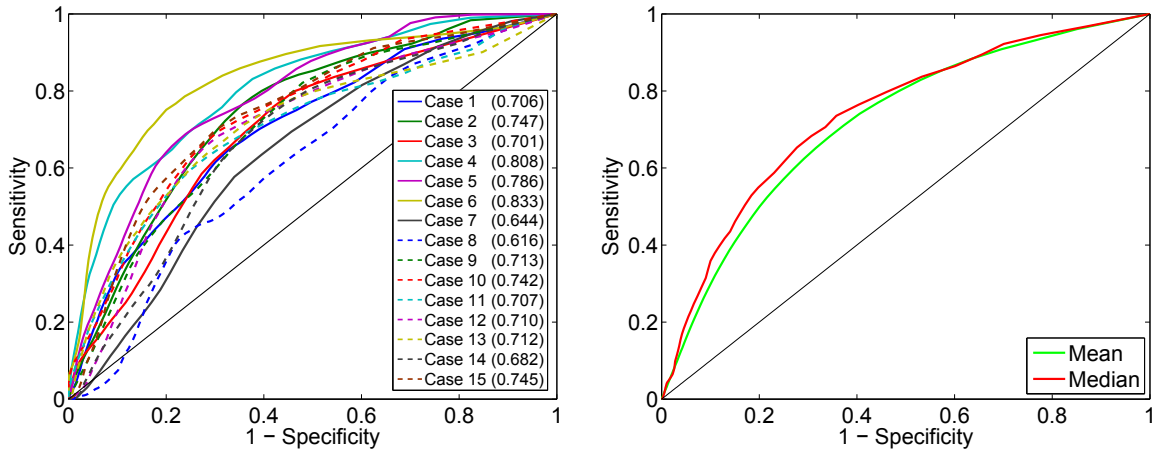


Figure 2.5: Quantitative results on fifteen RAPN cases. Left: ROC of all cases with the associated area under curve in parenthesis. Right: median (red) and mean (green) of all ROC. Mean area under all ROC is 0.72 with a standard deviation of 0.05.

video that highlight venous vs. arterial pulsation.

The two exemplar frames were manually extracted from the same cardiac cycle. The fifteen four-second clips contain between four to eight cardiac cycles; this number varies depending on the heart rate of the patient during acquisition. Each motion label frame within a cardiac cycle, containing 15–30 frames, was visually compared to the ground truth to identify a pair of frames that best represent the venous and arterial structures. The ground truth was not altered during this process. Compared to the reference in Figure 2.4b, the motions highlighted in Figure 2.4e correspond to the cyan structures (venous) and Figure 2.4f corresponds to the red structures (arterial).

Currently our method is not able to automatically differentiate between the locations of veins and arteries. Therefore, to quantify a measure of detection for such pulsating structures, the automatic segmentations were first binarized (at a fixed threshold throughout the sequence) and combined across the frames of the video to generate a single binary image containing the union of all pulsatile motion regions in the video. The resulting image was then compared to a binary mask (Figure 2.4c) of the reference manual segmentation, combining all vasculature into a single mask. Figure 2.5 illustrates the segmentation performance of all cases, at different threshold values, via their receiver operating characteristics (ROC). The areas under these ROC curves are presented in the legend of in Figure 2.5.

Clinical User Study

We performed an initial user-centric assessment of the proposed technique by analyzing feedback from different surgeons at the Hamad Medical Corporation. We recruited four surgeons (two early career surgeons, and two experienced surgeons) and asked them to watch short surgical scenes (video clips) of hilar dissection from six different patients. Each surgeon was presented two types of video clips for each patient and was asked to identify the vasculature in the scene. The first set of videos consisted of clips from the original unprocessed surgical scenes, while the second set of videos consisted of the same clips

but augmented using our proposed method. To prevent the surgeon's performance from being biased by their viewing of the original and augmented videos of the same patient, we separated the viewing dates by one week and randomized the order of viewing for both original and augmented scenes. In addition, the surgeons were given the option of viewing the history video, defined as a video clip of 5–10 minutes in duration representing the history of the surgery preceding the selected short scene. The availability of this option was requested by the surgeons as it mimics real surgical scenarios. During a real operation, the events and actions leading to the current scene provide context for the short (4 seconds) video clips, thus implicitly providing the surgeon with vital information such as where the vessels may be based on how the organs have been retracted and positioned. In each experiment, we measured the time it took for the surgeon to annotate the clip and saved the surgeon's annotations on the first frame of each scene. We then compared the annotations with the ground truth data of that first frame (binary mask in Figure 2.4b) using the Dice similarity coefficient (DSC).

In general, the proposed method was found to improve vessel detection, mainly by reducing the detection-time for early career surgeons. More specifically, the proposed method reduced vessel detection time by 22% on average for early career surgeons, though it did not seem to affect detection-time for expert surgeons. Our proposed overlay only increased the average DSC for all users by a marginal 0.024. Upon debriefing, both early career and experienced surgeons confirmed that they were relying on visual cues and prior knowledge for locating vessels and that the overlays were mainly used to confirm their own localizations, which in and of itself was reported as an added benefit by the surgeons. The feedback from the surgeons was generally positive, with the exception of one experienced surgeon who stated that the visualizations of the segmentations were difficult to interpret. Nevertheless, the average DSC for all cases (with and without augmentation) performed by all participants was 0.13 with a standard deviation of 0.11, whereas our method performed with a DSC of 0.50 with a standard deviation of 0.24. Note that to compute the DSC for our method, we thresholded our fuzzy labels at 0.10 for all cases, this threshold was chosen as it gave the highest average DSC across all cases. Our interpretation of these findings is reported in the following section.

Discussions

In this section, we analyze the significance and implications of our experimental findings. We were careful in the selection of our cases and in ensuring that the alignment of the ground truth is, to the best of our ability, without error. The reported misalignment errors in ground truth (up to 4–7 mm in cases with visible non-rigid deformations) may appear to be large to some readers, but it is rather small compared to the size of the structures (up to 30 mm in diameter for the IVC) that we are trying to detect. This small error does not have a pronounced impact on our interpretation of the results.

Our quantitative assessment indicates a mean area under the ROC curve of 0.72, midway between the noise baseline of 0.5 and the highest possible value of 1.0 (indicating perfect detection). Our detection rate may be slightly higher than the presented value due to the aforementioned misalignment of the ground truth. A conservative value of 0.72 suggests that our generated labels are discriminative and suitable for integration (as an additional data term) into existing vessel segmentation techniques that use

other visible features such as color, shape and texture.

To put the reported detection rates into perspective and to identify opportunities for achieving higher accuracy, we manually chose two exemplar frames (depicting venous and arterial structures) from the resulting segmentations and compared the results to our ground truth. The following observations were made for each of the fifteen cases.

Case 1 The RV and IVC are correctly labeled in spite of heavy occlusion. The small RA is also identified at the hilum, even though some false positives were detected on the tools. We attribute the false positives to the fact that the vascular pulsations cause the end effectors of the surgical tools to vibrate through the abdominal wall.

Case 2 The ground truth location of the RA (Figure 2.4b) is slightly misaligned due to retraction by the surgical instrument. The detected locations of the AA, RA, IVC and RV are correct despite the retraction.

Case 3 All structures are correctly identified, including the portion of the IVC that is occluded by the surgical gauze and the small accessory RA to the left of RV.

Case 4 All structures are identified but it is difficult to differentiate between arterial and venous structures; possibly because of the heavy occlusion by the perinephric fat. Possible misalignment of the ground truth (suprarenal vein) due to mobilization of the spleen.

Case 5 All structures are identified. Some false positives are present during arterial pulsation (Figure 2.4f). There may be a misalignment in the ground truth as retraction has shifted the abdominal aorta up.

Case 6 Branching of the RA is detected on both sides of RV and the pulsation of heavy vascular region has caused the tumor to pulsate in the center of the frame.

Case 7 The RV is correctly identified but noticeable amount of false positives is observed due to the motion of the tools in the scene. Ideally, surgical instruments should remain motionless during the acquisition of the video.

Case 8 Like Case 7, Case 8 also posed a big challenge as the vasculature is heavily occluded by the bowel and many false positives are detected in fluid filled cavities to the left of the cyst.

Case 9 All structures are roughly identified. A specific patch of specular reflections on top of the RA in the left side of the frame (Figure 2.4d) has skewed the normalization of the labels. Specular highlights pose a great challenge to endoscopic video analysis, and although the specular patches are not large at the hilum compared to other organs such as the liver, their presence does have noticeable effects, i.e., regions with highlights are emphasized more (stronger response) in comparison to their neighbors. Cases 10 and 12 are also notable examples of this.

Case 10 The RA and RV are correctly identified. A large part of the bowel is occluding the large (30 mm) IVC making it difficult to detect.

Case 11 The RA is correctly identified but the pulsations of the RV are missed. We attribute this to the large (20 mm) size of the structure. In such cases, the number of scales of the spatial steerable pyramid may be increased to five levels.

Case 12 Both RA and RV are clearly detected in Case 12; small motion artefacts are present on the suction tool.

Case 13 All structures are identified. This case provides a good visualization of the phase difference between arterial and venous pulsations.

Case 14 In addition to identifying the RV our method was also capable of localizing the gonadal vein present on the right side of the image (Figure 2.4e).

Case 15 All structures are identified. There is a Heavy presence of false positives caused by specular highlights on the top right corner.

In summary, our method was qualitatively successful in all cases, except for Cases 7 and 8, both of which are associated with the lowest area under curve values. Camera and tool movement may seem to be a big challenge but coping with them is rather simple in the context of hilar dissection during RAPN. Hilar dissection stage is not time-critical in comparison to the subsequent resection and reconstruction stages. Our surgeons have confirmed that they can afford a four-seconds pause during the operation while our system acquires the required information for processing. The surgical instruments can also be moved out of sight during the acquisition. The most challenging sources of false positives are the specular highlights and pulsatile vibrations in background structures. All organs inside a live human exhibit minute pulsatile motion in sync with the heart rate. These minute vibrations are very small compared to the motion of major vasculature, yet their apparent motion is magnified with the presence of specular highlights or fluids. In the future, we plan to focus mainly on an adaptive estimation for noise (to boost its robustness to specular reflections) and automating the process of differentiation between veins and arteries.

Our initial user study was insightful, a Wilcoxon signed-rank test indicated that the segmentation performance of our method (mean DSC of 0.50) was statistically significantly higher than the segmentations performed by the surgeons (mean DSC of 0.13), $p < 0.001$. This difference does not indicate that the ground truth is inaccurate or that the surgeons disagree with it. The performance of the surgeons appears to be poor since surgeons were not able to locate large segments of the vasculature that were hidden/occluded. This difference in DSC and the fact that the performance of the surgeons only improved marginally imply that, perhaps with more training and a better visualization technique, our method has the potential to improve the surgeon's performance even further. We plan to address this need by developing new (clinically appropriate) visualizations, user inputs that grant the surgeon the ability to control the opacity of the displayed overlay, and a spotlight option that enables the surgeon to choose the region within which the segmentations are overlaid. Although our initial trial is promising, further studies with more participants and data are required to quantify the clinical impact and effectiveness of our method in finding accessory vessels.

2.3.3 Embedding Motion-Based Cues in a High-Level Segmentation Framework

The PBMS technique, presented in this chapter, which was first proposed in Amir-Khalili et al. [5, 7], was then incorporated into the high-level endoscopic video segmentation framework of Nosrati et al. [125] and presented in Nosrati et al. [126]. In Nosrati et al. [126], a variational technique is proposed to augment the surgical scene by segmenting visible as well as occluded structures in the intraoperative endoscopic view. This approach to scene augmentation, estimates the 3D pose and deformation of anatomical structures segmented from 3D preoperative data in order to align to and segment corresponding structures in 2D intraoperative endoscopic views. To estimate pose and segment the highly noisy and cluttered environment of an endoscopic video, this framework leverages both preoperative data, as a source of patient-specific prior knowledge, as well as the vascular motion cues presented in this chapter, and endoscopic visual cues by training a random decision forest. A tissue-specific physically-based deformation model is also employed to handle the non-rigid deformation of different structures. Finally, to make the non-rigid deformation of each structure closer to reality, the Houndsfield unit (HU) value of each structure in the preoperative CT is used to assign a specific tissue-specific (i.e. heterogeneous) stiffness constraints to each deformable model. The utility of this technique was validated on the same fifteen challenging clinical cases that was summarized in Section 2.3.1, demonstrating a 45% improvements in accuracy compared to the standard PBMS method.

The detailed methodology of this high-level segmentation framework is out of the scope of this thesis and was thus omitted. Readers are referred to Nosrati et al. [126] for more information.

Results

The default parameters suggested in our previous works [5, 7] were used to detect the vascular motion cues. The parameters of learning based colour and textural feature extraction method is detailed in Nosrati et al. [126]. The segmentations obtained through the high-level segmentation system of Nosrati et al. [126] was validated with the same ground truth presented in Amir-Khalili et al. [5, 7] (Figure 2.6 and Figure 2.7). Note how the noisy segmentations in Figure 2.6c and Figure 2.7c are improved in Figure 2.6d and Figure 2.7d by incorporating the learned intensity cues and preoperative prior information. The quantitative performance comparisons between the original Amir-Khalili et al. [5, 7] approach and the high-level segmentation method of Nosrati et al. [126] is presented in Table 2.1.

In the high-level segmentation implementation, the average run-time of our unoptimized MATLAB code to process the vessel pulsation in a four-second clip (120 frames) was 65 seconds. The run-time for pose estimation and segmenting the structures depends on the initial pose of the organs. The average run-time to find the pose and segment the structures during system initialization is $\sim$16 seconds on a standard 3.40 GHz CPU.

Discussions

The results on *in vivo* clinical cases of partial nephrectomy illustrate the potential of the proposed framework for augmented reality applications in minimally invasive surgery (MIS).

The observable differences between the ground truth and the results presented in Figure 2.6 and

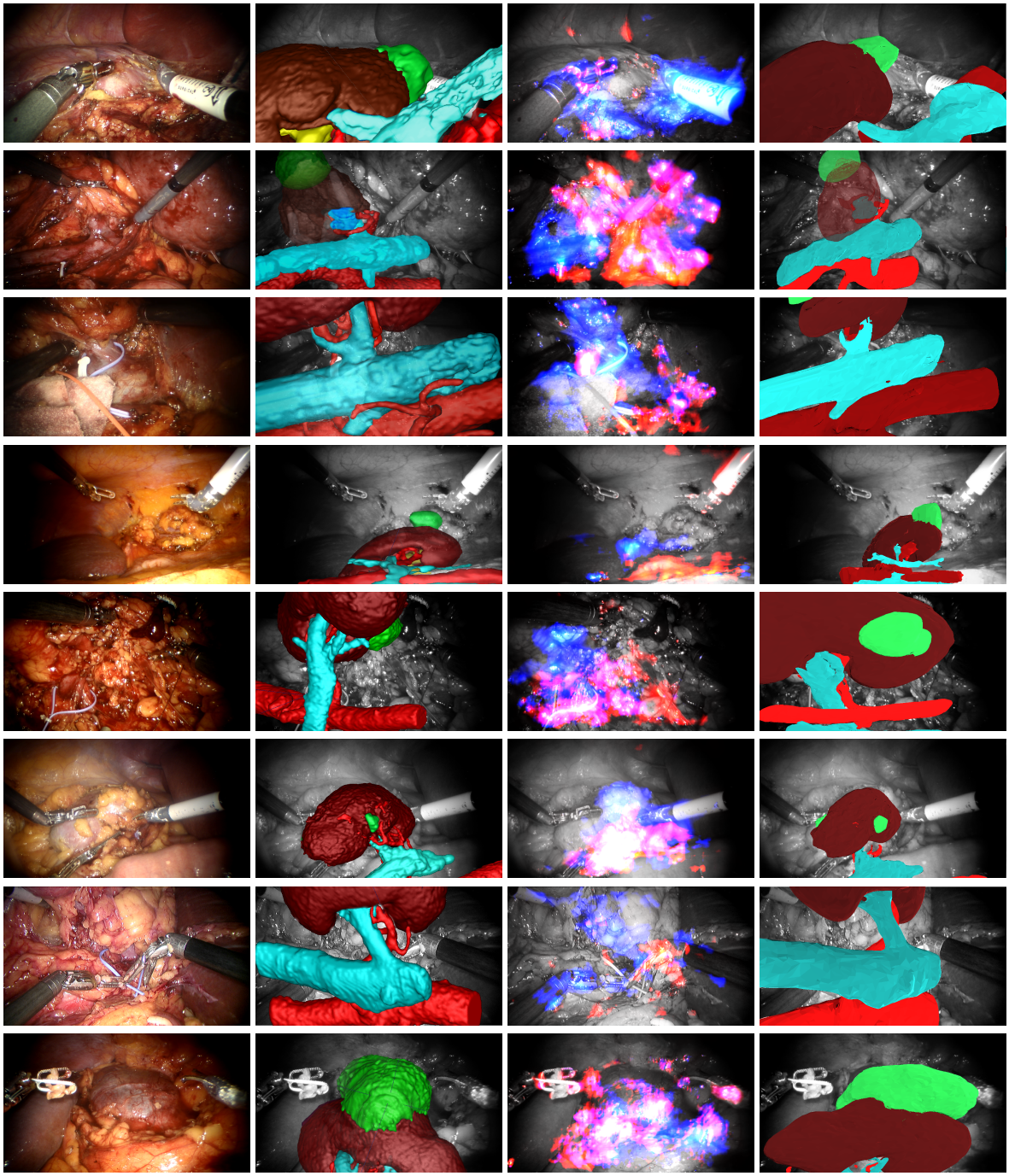


Figure 2.6: First eight cases of the qualitative comparison of the proposed high-level segmentation method against the simple PBMS method. (a) Original endoscopic image. (b) The ground truth of venous (cyan), arterial (red), kidney (brown) and tumor (green). (c) Segmentation results of vessels using PBMS. (d) Segmentation results of the proposed high-level framework.

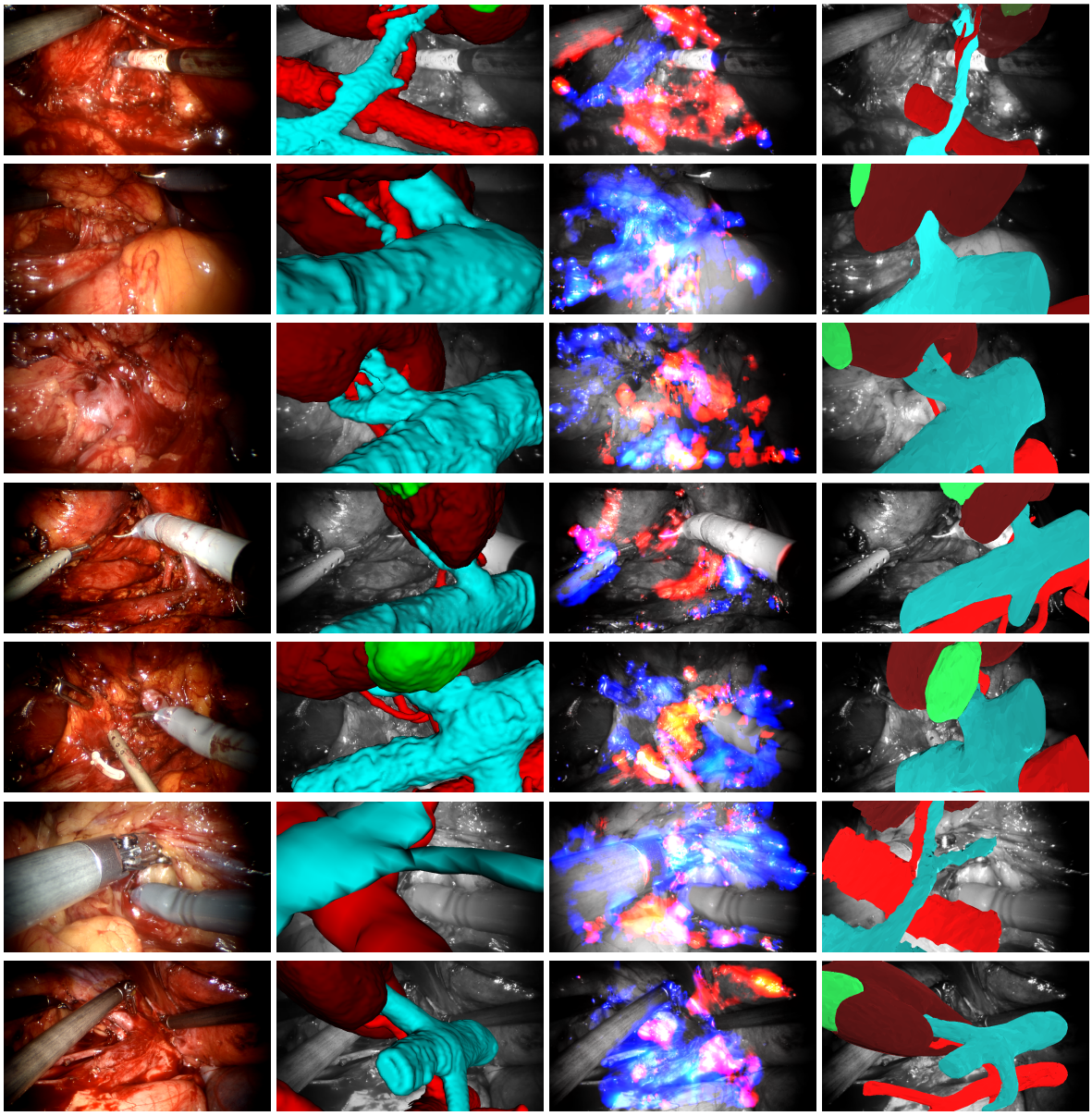


Figure 2.7: Final seven cases of the qualitative comparison of the proposed high-level segmentation method against the simple PBMS method. (a) Original endoscopic image. (b) The ground truth of venous (cyan), arterial (red), kidney (brown) and tumor (green). (c) Segmentation results of vessels using PBMS. (d) Segmentation results of the proposed high-level framework.

Figure 2.7 are attributed party to the local optimization framework and also partly to the error in the alignment of the ground truth. As mentioned in Section 2.3.1, due to the fact that the preoperative model was rigidly aligned to the endoscopic video, an alignment error of 4–7 mm exists in cases where the organs have been considerably retracted by the surgical instruments or mobilization of other organs.

Table 2.1: Quantitative comparison for kidney and vessel segmentation of the proposed high-level framework vs. direct phase-based motion segmentation (PBMS). Results are presented in terms of accuracy, Dice similarity coefficient (DSC), true positive and false positive rates. Note that the fuzzy PBMS results were thresholded at 0.50 to generate the following comparisons.

Method	**DSC**		**True Positive Rate**		**False Positive Rate**		**Accuracy**	
	Kidney	Vessel	Kidney	Vessel	Kidney	Vessel	Kidney	Vessel
PBMS	-	0.41	-	**0.74**	-	0.40	-	0.60
High-Level Framework	0.70	**0.61**	0.70	0.56	0.07	**0.06**	0.88	**0.87**

We believe that despite the visible differences between the two, this variational segmentation framework may be one step closer to an ideal solution compared to the ground truth as this framework allows for non-rigid modes of vibration. Generating a ground truth that accounts for the non-rigid deformations due to mobilization and retraction requires volumetric intraoperative imaging such as cone beam CT or possibly implanting fiducials. The use of such imaging techniques is not feasible as it exposes the patient and clinicians to ionizing radiation and implanting fiducials is intrusive and invasive and hence not recommended.

There are several directions to extend this work. This variational framework is highly parallelizable and we do not foresee any obstacles towards a graphical processing unit (GPU) implementation for real-time pose estimation and endoscopic video segmentation. In addition, we believe that leveraging stereo views as well as encoding depth information into the proposed energy functional can improve the performance.

Another promising avenue for research would be to explore the use of an additional shape variation component that is orthogonal to the restricted shape model, as described by Andrews and Hamarneh [10], since this allows for exploring larger shape variability without noticeable increase in complexity. Although the modes of vibration are limited to less than three times the corresponding eigenvalue (to avoid any irrational shape deformation), similar projection from two different 3D deformations are possible. This is due to the fact that geometric information is inevitably lost during the 3D to 2D transformation. We believe that this limitation may be addressed by leveraging stereo views and is thus another interesting future direction that warrants further investigation. Furthermore, given that a local optimization technique was used in this framework, by leveraging state-of-the-art convexification techniques [10, 11, 108, 123] we can further reduce the sensitivity of this framework to initialization. Finally, improved estimates of elasticity parameters (e.g., using elastography imaging) will likely constrain the space of non-rigid deformations more accurately [154].

2.4 Summary

In this chapter, we have presented an automatic method for localizing and labeling regions in endoscopic video that contain occluded vessels. Our method extends Eulerian phase-based video motion processing techniques to detect and label small motions that are barely visible on the surface of the perinephric fat.

To the best of our knowledge, we were the first to attempt the challenging task of localizing occluded vasculature in endoscopic video without the use of additional hardware or preoperative scans. We validated our novel method qualitatively in a retrospective *in vivo* study to verify its application in a clinical setting. Using manually aligned preoperative models of the *in vivo* patient data as ground truth, we performed conservative quantitative validation of our method to report well known measures of detection, i.e., the area under the ROC curve. Furthermore, we conducted a preliminary clinical study, and received very enthusiastic feedback from the surveyed urologists.

The results from our evaluation and the user study demonstrated that our vessel localization method is suitable for integration alongside existing techniques (e.g., as an additional cue) that use other visible features such as color, shape and texture. In further support of this claim, we presented the results of a high-level segmentation framework—an extension of Nosrati et al. [124]—that integrates both preoperative data, as a source of patient-specific prior knowledge, as well as our vessel localization cues and endoscopic visual cues in order to accurately segment the highly noisy and cluttered environment of an endoscopic video. With the help of the proposed motion cues, the presented high-level framework boasts state-of-the-art performance in localizing the kidney and vasculature from real surgical scenes of RAPN interventions; producing a mean DSC of 0.70 and 0.61, and accuracy of 0.88 and 0.87 for localizing the kidney and blood vessels respectively.

We have therefore provided evidence in support of our hypothesis that an understanding of temporal anatomical behavior and mathematical modeling of this behavior can be used to improve localization and navigation in the context of image-guided interventions, i.e., RAPN. To bolster our findings, in addition to more extensive user studies involving *in vivo* experiments and animal trials, further evaluation of the limitations of our methodologies in the context of RAPN are to be conducted with controlled *ex vivo* phantom experiments. Specifically, there is a clear need to study the sensitivity of our methods to various physiological parameters, which include: blood pressure, amount of occlusion (thickness of tissues covering the vessel), vessel diameter, and vascular distensibility. This study would involve the construction of a realistic *ex vivo* phantom, such as the one developed by Schneider [154]—however, to mimic realistic pulsatile vascular motion in the phantom, we recommend the use of a pulsatile pump [25, 112, 190] in place of the peristaltic pump used in Schneider [154]. Indeed, the construction of said *ex vivo* phantom is challenging, but such phantom experiments are indispensable as they may be used to prototype and evaluate different visualization techniques. Furthermore, these phantom experiments may also be used to explore other opportunities to improve our motion segmentation methodologies, i.e., more advanced endoscopic camera hardware and the use of both stereo endoscopic cameras to improve vessel localization [110].

In the following chapter, we present extensions to our proposed automatic vessel localization methodology to increase its computational performance and sensitivity to outliers, and hence extend its utility to other dynamic imaging modalities and clinical image-guided interventions that demand real-time computational performance.

Chapter 3

Kinematic Model-Based Vessel Segmentation

"All parts of the material universe are in constant motion and though some of the changes may appear to be cyclical, nothing ever exactly returns, so far as human experience extends, to precisely the same condition."
— Joseph Henry

This chapter details the extension of our phase-based motion segmentation (PBMS) method presented in Chapter 2 to the more robust kinematic model-based vessel segmentation (KMVS) method, which was originally disseminated in Amir-Khalili et al. [6, 8]. The KMVS method leverages the local orientation, in addition to magnitude of motion, and demonstrates that the extended computation and utilization of motion vectors can improve the segmentation of vascular structures. We implement our KMVS method using four alternatives to magnitude-only motion estimation by using traditional optical flow and by exploiting the monogenic signal for fast flow estimation.

3.1 Leveraging Motion Vector Computation for Vessel Segmentation

Our goal is to leverage an understanding of the kinematics of vascular structures to perform image segmentation. These kinematics are observed through the change in intensity information of anatomical structures captured in a video or dynamic sequence of frames. In this paper, each j^{th} frame of such sequence is defined as a scalar valued (grayscale) function $f : (\mathbf{x} \subset \mathbb{R}^2, j \subset \mathbb{R}) \rightarrow \mathbb{R}$ mapping a pixel $\mathbf{x} = (x_1, x_2)^\intercal$ in the 2D spatial domain of each frame to an intensity value. Depending on the spatiotemporal resolution of the sequence and the specific vascular anatomy being imaged, some kinematics of the vasculature may be observed by the naked eye, the most notable characteristic being the periodic motion of the vascular walls induced by the pulsatile flow of blood in the vessel. The visibility of this phenomenon—or the magnitude of observable displacement—varies depending on the radius, thickness, and viscoelastic properties of the vascular walls as well as the flow rate and pressure of blood inside the vessel [198]. In previous publications [5, 7], we demonstrated that advanced Eulerian motion estimation techniques may be used to observe this phenomena, even in situations where the motions are

subtle and imperceptible to the naked eye, by observing the temporal change in f at every pixel $\mathbf{x}$.

There are different ways to identify periodic motions occurring at a given pixel. For blood vessels, these periodic pulsations are within a temporal passband centered on the heart rate of the patient. A trivial way to identify this periodic motion is to apply a temporal bandpass filter to the raw intensity information f at every pixel $\mathbf{x}$ independently. This naïve approach is prone to error as it does not consider the motion of neighboring pixels, is sensitive to noise, and it cannot estimate the magnitude of motion, which is required to attenuate the effects of noise. In our previous PBMS approach [5, 7], presented in Chapter 2, we overcame these limitations by (i) estimating the magnitude of motion via the change in local (spatial) phase information, and (ii) measuring this change over multiple spatial scales and orientations to encode the motion information from neighboring pixels.

In the original PBMS formulation, the objective was to generate segmentation labels for every frame in the sequence. In this work, however, we aim to leverage all temporal information to generate a single segmented frame

$$L(\mathbf{x}) : \mathbb{R}^2 \rightarrow l \in [0,1]. \tag{3.1}$$

To do so, the denoised product $\hat{\phi}_z(\mathbf{x}, j; s, n)$ is averaged across all scales s, orientations n, and frames j to obtain our final fuzzy labels

$$L_P(\mathbf{x}) = \frac{1}{M_P} \sum_{\forall j,s,n} \frac{|\hat{\phi}_z(\mathbf{x}, j; s, n)|}{2\pi\omega_s}, \tag{3.2}$$

where ω_s is the spatial frequency of scale s and M_P is a normalizing factor to fix the range of PBMS labels $L_P \in [0,1]$.

The simple averaging across all orientations only considers the weighted mean magnitude of motion. An alternative approach, presented in Amir-Khalili et al. [6, 8] and described in detail below, is to compute the motion vectors entirely (magnitude and orientation) and to use a kinematic model that is more specific to vessel-like structures, which radially distend and contract in time.

3.1.1 Localizing Vasculature from Divergent Motion Patterns

The PBMS method, described in the previous Chapter 2, leverages the pulsatile temporal motion of structures to approximate the location of blood vessels. In addition to pulsatile motion, the geometry of the vessel is also an integral part of its distinguishing kinematics. Blood vessels are tubular structures. When these structures are subjected to a pulsatile flow, the vascular walls undergo radial and longitudinal displacements [198]. The radial displacement component manifests as the expansion and contraction of vessel walls in medical images. These motions are unique to vasculature, unlike the longitudinal motions that also occur in the surrounding tissues. The complete computation of local motion vectors, that encode orientation and magnitude, allows us to model the characteristic pulsatile radial motion of a blood vessel more accurately.

The major components and differences between PBMS and KMVS are illustrated in Figure 3.1. In the KMVS pipeline, the simple multi-scale motion-based segmentation step is replaced with a complex spatiotemporal pulsatile radial motion model (PRMM). Although the parameters of the temporal band-

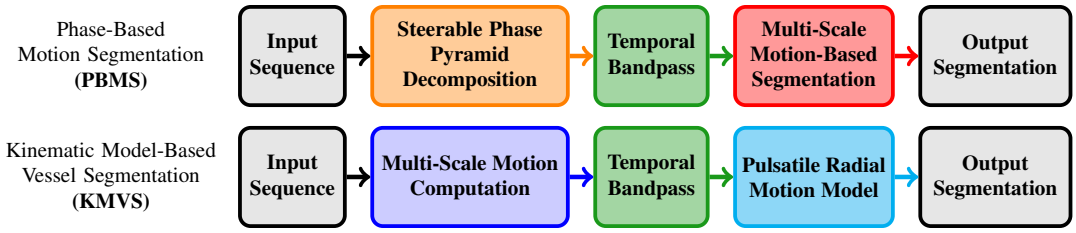


Figure 3.1: Overview of the PBMS and KMVS segmentation pipelines presented in this paper. The same filter parameters are used to create the temporal bandpass filters (green) in both pipelines but the dimensionality of inputs and outputs of the filters are different. The steerable phase pyramid decomposition (orange) decomposes the input sequence $f : \mathbb{R}^3 \to \mathbb{R}$ into the overcomplete representation of phase magnitude $\phi : \mathbb{R}^{3+N+S} \to \mathbb{R}$ measured along N orientations and S scales, while the output of the alternative multi-scale motion computation step (blue) is the dense flow field $\mathbf{d} : \mathbb{R}^3 \to \mathbb{R}^2$. In the PBMS pipeline, the bandpassed phase magnitudes are combined using a weighted average in the multi-scale motion-based segmentation stage (red) to generate fuzzy segmentation labels; whereas, in the pulsatile radial motion modelling stage (cyan) of the KMVS pipeline, the divergence operator is applied to the bandpassed flow fields to extract a more accurate fuzzy segmentation.

pass used in both approaches are the same, the dimensionality of inputs to the filter are different. Lastly, the KMVS approach requires an estimate of complete motion vectors. These motion vectors may be computed using traditional optical flow (OF) or monogenic flow (MF).

Optical Flow

OF methods compute apparent local motions of objects in a sequence of images by imposing the brightness consistency assumption. This constraint

$$f(\mathbf{x}, j) = f(\mathbf{x} + \Delta\mathbf{x}, j + \Delta j) \tag{3.3}$$

assumes that the scene depicted at different points in time does not change in intensity, i.e., objects merely translate across frames. In the discrete time domain of the sequence, this constraint defines motion vectors $\mathbf{d}(\mathbf{x}, j)$ that relate two frames in a sequence

$$f(\mathbf{x}, j+1) = f(\mathbf{x} - \mathbf{d}(\mathbf{x}, j), j), \tag{3.4}$$

which can also be expressed as a continuous problem

$$\frac{df(\mathbf{x}, j)}{dj} = \frac{\partial f}{\partial \mathbf{x}}\frac{d\mathbf{x}}{dj} + \frac{\partial f}{\partial j} = \nabla_{\mathbf{x}} f\mathbf{v} + \nabla_j f = 0, \tag{3.5}$$

where $\frac{\partial f}{\partial \mathbf{x}} = \nabla_{\mathbf{x}} f$ is the spatial gradient of the frame, $\frac{\partial f}{\partial j} = \nabla_j f$ is the difference between the intensities of two frames, and $\mathbf{v}$ is the velocity, or flow, of motion.

A simple solution to this problem, initially proposed in the context of computational stereopsis, is

to assume that the motion is constant over a local window. This assumption leads to an overcomplete system of equations that can be solved using least squares iteratively by warping the image at each iteration [103]. Additionally, large displacements may be accounted for by performing the iterative optimization over multiple spatial scales as well, by subsampling the frames in the spatial domain. Computational stereopsis problems comprise a subset of OF problems, where correspondence between two rectified stereo images are constrained to only one dimension [68]. Once the motion of objects within the scene are generalized to 2D, the simple localized least squares solution becomes ill-conditioned and the problem of motion estimation becomes more difficult due to the aperture problem. An alternative way to overcome this problem, proposed by Horn and Shunck [73], is to impose a global smoothness constraint over the flow vectors $\mathbf{v}$ and solve the problem globally.

Both of these local and global approaches of computing flow have drastically advanced since their inception and have also been successfully combined into a unified framework [20], which combines the advantages of both approaches, i.e., robustness to noise and ability to yield dense flow fields. In our analyses, we opted to use two modern implementations of OF [100, 178] implemented in MATLAB for the purpose of computing the dense motion vector $\mathbf{d}(\mathbf{x}, j)$.

One of the major drawbacks of OF lies in the brightness consistency assumption. This assumption is sensitive to smooth contrast variations (temporal changes in lighting conditions) and other similar situations where pixel intensities cannot be considered as reliable features. These scenarios are abundant in medical image sequences; examples include: moving light sources in endoscopic video (EV), specular noise or non-Lambertian reflections, and local brightness variations caused by complex acoustic beam propagation in dynamic ultrasound (DUS). In context of 3D DUS imaging, [1] attribute brightness consistency violations (temporal variations in the local echo strength) in part to changes in the angle between connective tissue fibers and beam propagation direction, and the limited acquisition frame rates of 3D DUS. 2D DUS benefits from higher frame rates, but suffers from artifacts caused by the out-of-plane motion of structures within the 2D field of view. Structures with varying thickness and acoustic properties may travel through the field of view during acquisition and cast a time varying acoustic shadow on surrounding tissues. This results in local attenuation or amplification of 2D B-mode intensity values which are not necessarily correlated to the relevant in-plane motions. On the other hand, it has been argued that phase-based computation of motion vectors is more robust to this type of noise and has the added advantage of producing subpixel accuracy without explicit subpixel reconstruction or feature localization [47, 195]. This was the original motivation behind our choice to utilize a phase-based [195] approach in our PBMS pipeline compared to seminal gradient-based Eulerian video magnification (EVM) approaches [202].

Monogenic Flow

The monogenic (MON) signal is another 2D extension of the analytic signal (similar to the steerable phase-based method presented in Section 2.1.1) and it provides an efficient framework for extracting the local orientation θ and the local phase ϕ features from an image. By measuring the temporal change in θ and ϕ in a sequence of images, we can estimate motion [2, 42]. The MON signal is constructed

from a trio of bandpass filters with finite spatial support. This trio is commonly referred to as spherical quadrature filters (SQF) [43]. To estimate the motion of both small and large structures in each frame, we generate different SQF by tuning the spatial passband of the filters to varying scales.

Each set of SQF comprises an even (symmetric) radial bandpass filter and two odd (antisymmetric) filters. The odd filters are computed from the Riesz transform, a 2D generalization of the Hilbert transform, of the radial bandpass filter [43]. In the literature, many different bandpass filters have been proposed to construct the SQF including: first order Gaussian [18], Cauchy [18], Poisson [44], and difference of Poisson filters [200]. We employ Log-Gabor [93] bandpass filters as they suit the natural statistics of an image [46, 168] and maintain zero DC gain at lower spatial scales. For every scale s, the even Log-Gabor component of the SQF is expressed as

$$B_e(\mathbf{u};s) = \exp\left(\frac{-\left[\log\left(\frac{|\mathbf{u}|}{\omega_s}\right)\right]^2}{2\left[\log k\right]^2}\right), \tag{3.6}$$

in the frequency domain $\mathbf{u} = (u_1, u_2)^\intercal$, where k and ω_s are parameters of the filter. The parameter $k = \sigma/\omega_s$ is a fixed constant representing the ratio of the standard deviation σ of the Gaussian describing the Log-Gabor filter's transfer function in the frequency domain to the filter's center frequency ω_s. At each scale s, the center frequency is defined as $\omega_s = (\lambda_0 2^{(s-1)})^{-1}$, where λ_0 is an initial minimum wavelength. The radial bandpass filter B_e is symmetric as it is only a function of the magnitude of the frequency $\mathbf{u}$. Using the Riesz transform, we compute the two odd components (B_{o1} and B_{o2}) associated to this SQF as

$$B_{o1}(\mathbf{u};s) = i\frac{u_1}{|\mathbf{u}|}B_e; \quad B_{o2}(\mathbf{u};s) = i\frac{u_2}{|\mathbf{u}|}B_e. \tag{3.7}$$

In the spatial domain, the components of the MON signal $(h_e, \mathbf{h}_o)$ are obtained by convolving the SQF with a given frame of the sequence such that

$$\begin{aligned} h_e(\mathbf{x}, j; s) &= \mathscr{F}^{-1}\left[B_e(\mathbf{u};s)F(\mathbf{u}, j)\right] \\ h_{o1}(\mathbf{x}, j; s) &= \mathscr{F}^{-1}\left[B_{o1}(\mathbf{u};s)F(\mathbf{u}, j)\right] \\ h_{o2}(\mathbf{x}, j; s) &= \mathscr{F}^{-1}\left[B_{o2}(\mathbf{u};s)F(\mathbf{u}, j)\right] \\ \mathbf{h}_o(\mathbf{x}, j; s) &= (h_{o1}(\mathbf{x}, j; s), h_{o2}(\mathbf{x}, j; s))^\intercal, \end{aligned} \tag{3.8}$$

where $F(\mathbf{u}, j) = \mathscr{F}[f(\mathbf{x}, j)]$ is the frequency domain representation of the frame.

From the SQF responses Equation 3.8, the phase vector $\mathbf{r}$ is then defined as the continuous representation of local orientation θ and local phase information ϕ such that

$$\mathbf{r}(\mathbf{x}, j; s) = \phi(\cos\theta, \sin\theta)^\intercal = \frac{\mathbf{h}_o}{|\mathbf{h}_o|}\arg\left(h_e + i|\mathbf{h}_o|\right). \tag{3.9}$$

Local motion may then be calculated by first computing the components of a 3D rotation that relates the

response of two adjacent frames in the video

$$\begin{aligned}\Delta h_e &= h_e(\mathbf{x}, j; s)h_e(\mathbf{x}, j+1; s) + \mathbf{h}_o(\mathbf{x}, j; s)^\intercal \mathbf{h}_o(\mathbf{x}, j+1; s) \\ \Delta \mathbf{h}_o &= h_e(\mathbf{x}, j; s)\mathbf{h}_o(\mathbf{x}, j+1; s) - h_e(\mathbf{x}, j+1; s)\mathbf{h}_o(\mathbf{x}, j; s)\end{aligned} \tag{3.10}$$

and then computing the phase differences $\Delta\mathbf{r}$ by substituting Equation 3.10 into Equation 3.9. Given a local neighborhood $\mathcal{N}$, the local displacement $\mathbf{d}_\mathcal{N}(\mathbf{x}, j; s)$ is calculated from

$$\sum_{\mathbf{x}\in\mathcal{N}} [\nabla^\intercal \mathbf{r}(\mathbf{x}, j; s)]\, \mathbf{d}_\mathcal{N}(\mathbf{x}, j; s) = \sum_{\mathbf{x}\in\mathcal{N}} \Delta\mathbf{r}(\mathbf{x}, j; s) \tag{3.11}$$

where $\nabla^\intercal$ is the divergence operator. The derivation of Equation 3.9, Equation 3.10, and Equation 3.11 from the response to the SQF falls outside of the scope of this paper and can be found in the original MF paper [42]. To improve the estimate for the displacement vector $\mathbf{d}_\mathcal{N}(\mathbf{x}, j; s)$, we compute the mean of this value across all scales s. The computed $\mathbf{d}_\mathcal{N}(\mathbf{x}, j)$ is an estimate of the true motion vectors $\mathbf{d}(\mathbf{x}, j)$ that relate two frames in a sequence $f(\mathbf{x}, j+1) = f(\mathbf{x} - \mathbf{d}(\mathbf{x}, j), j)$. This is the same motion vector that is computed by traditional OF techniques.

Pulsatile Radial Motion Model

Let $\mathbf{d}(\mathbf{x}, j)$ define a motion field containing the motion vectors estimated for all adjacent frames inside a given sequence using either OF or MF. We first isolate the motions that are in sync with the heart rate by applying the same ideal temporal bandpass filter (Figure 3.1) described in Equation 2.5. We define the temporally bandpassed motion vectors as $\mathbf{d}_z(\mathbf{x}, j) = \mathbf{d}(\mathbf{x}, j) * z(j)$.

Temporal filtering alone does not distinguish between structures that distend radially and tissues that translate at pulsatile frequency. Pulsating vessels are subject to periodic expansion and contraction, and the key insight is that this radial motion results in motion vector fields that are oriented away and then toward the center-line of the vessel during expansion and contraction respectively. Such vector fields thus exhibit high divergence along the center-line of the structure as illustrated in Figure 3.2.

In physical terms, divergence measures the extent to which a point source in the vector field behaves as a sink or source and it is defined as the sum of the partial derivatives of the vector field

$$\nabla^\intercal \mathbf{d}_z(\mathbf{x}, j) = \frac{\partial \mathbf{d}_z}{\partial x_1} + \frac{\partial \mathbf{d}_z}{\partial x_2}. \tag{3.12}$$

Due of the tubular geometry of vessels, the radial motions at the center-line of the vessels are weaker compared to the vascular walls. We account for this by computing the divergence across multiple spatial scales; the motion field at each scale denoted $\mathbf{d}_z(\mathbf{x}, j; s)$. At each scale, we downsample the vector field by a factor of two using bilinear interpolation. The resulting vessel labels are computed to be

$$L_K(\mathbf{x}) = \frac{1}{M_K} \sum_{\forall j,s} |\nabla^\intercal \mathbf{d}_z(\mathbf{x}, j; s)|, \tag{3.13}$$

where M_K is a normalizing factor to fix the range of KMVS labels $L_K \in [0, 1]$.

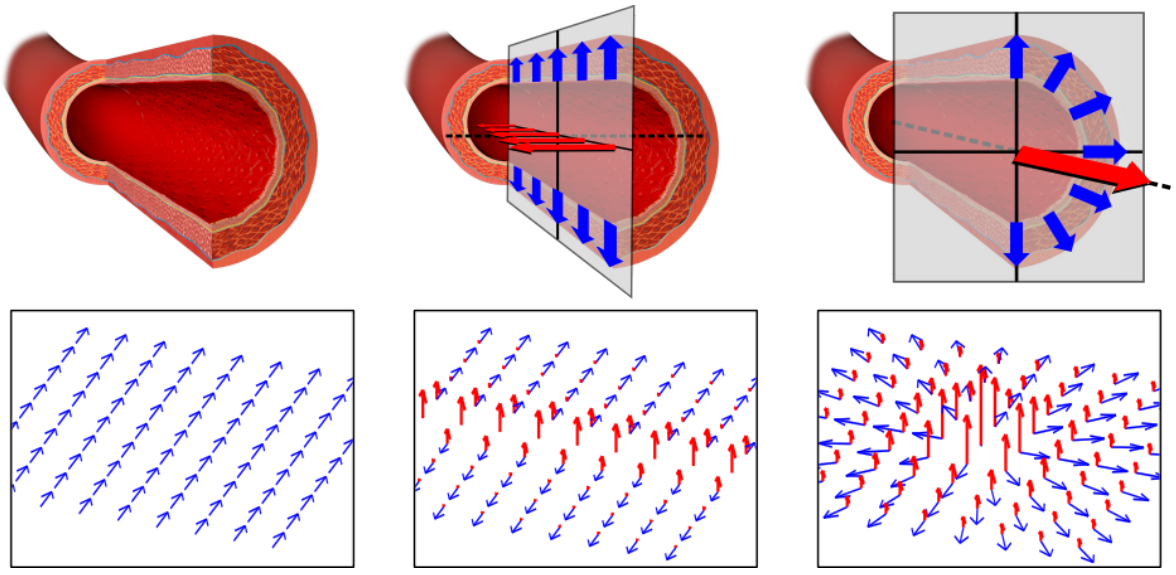


Figure 3.2: Depiction of simple 2D motion vector fields (blue arrows) and corresponding scalar divergence value plotted on an orthogonal axis (red arrows). The occurrence of divergent vascular wall motions are illustrated on longitudinal (b) and transverse (c) cross sections of an artery. Vector fields that purely translate in one direction (a) are incompressible or divergence free, whereas radial distension motion (b and c) exhibit high divergence along the center of expansion due to the tubular shape of the vessel.

3.2 Experiments

In this section we present two experiments to evaluate the performance of the different approaches presented in the methodology section. In each experiment, we compare the original PBMS implementation against four different implementations of KMVS. The four implementations of the KMVS consist of different motion estimation techniques, which includes two OF techniques [100, 178], the simple MF method presented in Section 3.1.1, and a more robust alternative to compute MF [2].

3.2.1 Implementation Details

The different implementations of PBMS and KMVS are denoted PB, OF1, OF2, MF1, and MF2. A summary of the different parameters used in each implementation is provided in Table 3.1. In our experiments, we use the default parameters of PB [7] and the `classic+nl-fast` setting of OF1 [178]. In the OF2 [100] implementation, we manually tuned the parameters to increase the speed and performance of the algorithm. The minimum wavelength was set to $\lambda_0 = 8$, weight of the regularization α was decreased to 0.022, the scale multiplier (inverse of downsampling ratio) was set to match other implementations $\delta = 2$, number of outer N_o and inner N_i fixed point iterations were reduced to 1, and

Table 3.1: Summary of parameters used in all experiments.

Name	Motion Estimation		Kinematic Model	
	Method	Parameters	Method	Parameters
PB	Amir-Khalili et al. [5, 7]	$s = 4$, $n = 4$	Multi-Scale Motion-Based Segmentation	$s = 4$
OF1	Sun et al. [178]	`classic+nl-fast`	PRMM	$s = \lfloor \log_2(L) \rfloor - 1$
OF2	Liu [100]	$\lambda_0 = 8$, $\alpha = 0.022$, $\delta = 2$, $N_o = 1$, $N_i = 1$, $N_s = 20$		
MF1	Amir-Khalili et al. [6]	$\lambda_0 = 2$, $k = 0.05$, $s = \lfloor \log_2(L) \rfloor - 1$, $\mathscr{N} \sim 7 \times 7$ box		
MF2	Alessandrini et al. [2]	mode = `lucas_kanade`, filter_type = `loggabor`, orient_mode = `robust`, freq_mode = `robust`, $\lambda_0 = 2$, $k = 0.05$, $\delta = 2$		

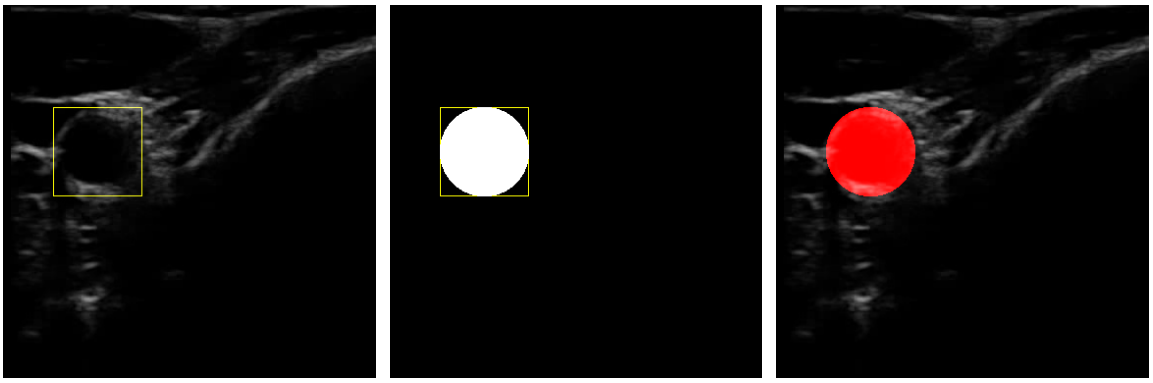
the default number of successive over relaxation iterations $N_s = 20$ was found to be sufficient for our purposes. The parameters for the MF1 [6] method were set to values used in the original publication: λ_0 is set to 2, ratio k for the Log-Gabor filter was set to 0.05, and a 7×7 box filter was used to average the displacements over the neighborhood of $\mathscr{N}$. The number of scales are set such that $s = \lfloor \log_2(L) \rfloor - 1$ where L is the smallest image dimension. MF2 [2] was setup to use the same filter parameters as MF1 but the frequency and scale computation mode was set to `robust`. The motion estimation mode was changed to `lucas_kanade` as the spatially affine transformation model was not performing well on our dataset. The remaining parameters were kept at their default values.

All of the methods described were implemented in MATLAB 2013b[1] running on a workstation with an Intel 3.7 GHz Xeon E5-1620 processor and 8 GB of RAM. The source code for OF1, OF2, and MF2 is publicly available online.

3.2.2 Materials and Experimental Setup

Although our methods are applicable to other imaging modalities, ultrasound (US) is ideal for validation as it can image vessels in the transverse and longitudinal axes, it has high temporal resolution, and the vessels can be manually delineated with accuracy and regarded as ground truth for evaluation. We validate the performance of PBMS and different implementations of KMVS on a set of synthetic computational phantoms and two DUS datasets of CCA scans. The phantom dataset is designed to mimic the kinematics of pulsating vascular structures imaged along the transverse and longitudinal slices and is described in further detail in Section 3.2.3. The first DUS dataset, hereby referred to as the UBC dataset [6], was acquired in-house and consists of eight sequences from three volunteers with six scans acquired along the transverse and two along the longitudinal axis of the CCA captured at 30 frames per second. The first frame of each sequence was manually segmented for quantitative analysis. These

[1] MATLAB executables are available for download from https://bisicl.ece.ubc.ca/software/radialDistension.html



(a) US with bounding box **(b)** Ellipsoidal mask **(c)** Ground truth overlay

Figure 3.3: Processing steps used to generate ground truth data for the SPLab dataset. Each frame in the dataset is accompanied by a set of coordinates that define a bounding box, overlaid in yellow, around the common carotid artery (CCA) (a). We generate an ellipsoidal mask (b) inside this bounding box and use it as an approximate ground truth overlaid in red (c).

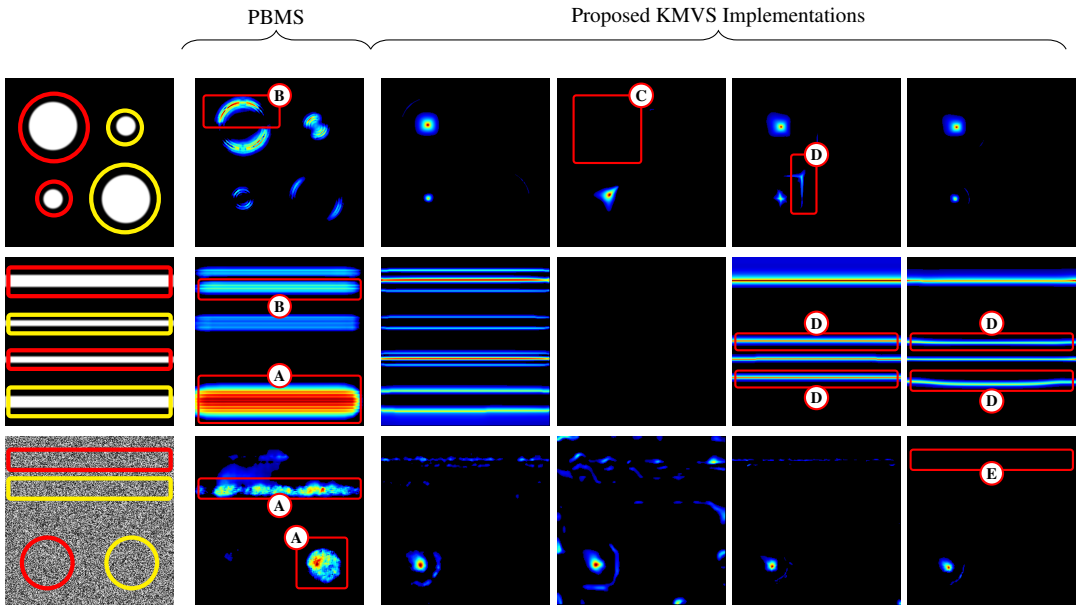
eight cases and the synthetic dataset were previously presented in our KMVS paper [6] and are used to select the parameters presented in Table 3.1. Our parameters were selected based on visual assessment of the resulting segmentation and motion estimation quality and the quantitative segmentation accuracy metrics employed in our results. A secondary publicly available DUS transverse scans of CCA, referred to as the SPLab dataset [142–145], is used to further corroborate our findings. We selected a total of 35 sequences from the dataset with each sequence containing four to eight frames captured at an estimated three frames per second. Every frame in the SPLab dataset is accompanied by image coordinates that define a tight bounding box (Figure 3.3a) around the CCA. These coordinates are used to generate an ellipsoidal mask (Figure 3.3b) for the first frame of the selected sequences to serve as an approximate ground truth (Figure 3.3c), in lieu of manual segmentations.

3.2.3 Localizing Vasculature in Dynamic Ultrasound Sequences

Phantom Experiment

We use three computational phantoms, in a two-frame matching experiment, to compare the effectiveness of our MF and OF based segmentation techniques to the PB method. Temporal filtering was not used in this experiment. Our phantoms consist of: pulsating and translating circles (top row in Figure 3.4 and Figure 3.5), pulsating and translating tubular structures (middle row in Figure 3.4 and Figure 3.5), and a noise pattern that undergoes a combination of pulsating and translating motions in shape of circles and tubes (bottom row in Figure 3.4 and Figure 3.5). The fuzzy automatic segmentation results obtained with all five implementations presented in Section 3.2.1 are presented in Figure 3.4 and the intermediary motion computation results of the KMVS implementations are presented in Figure 3.5.

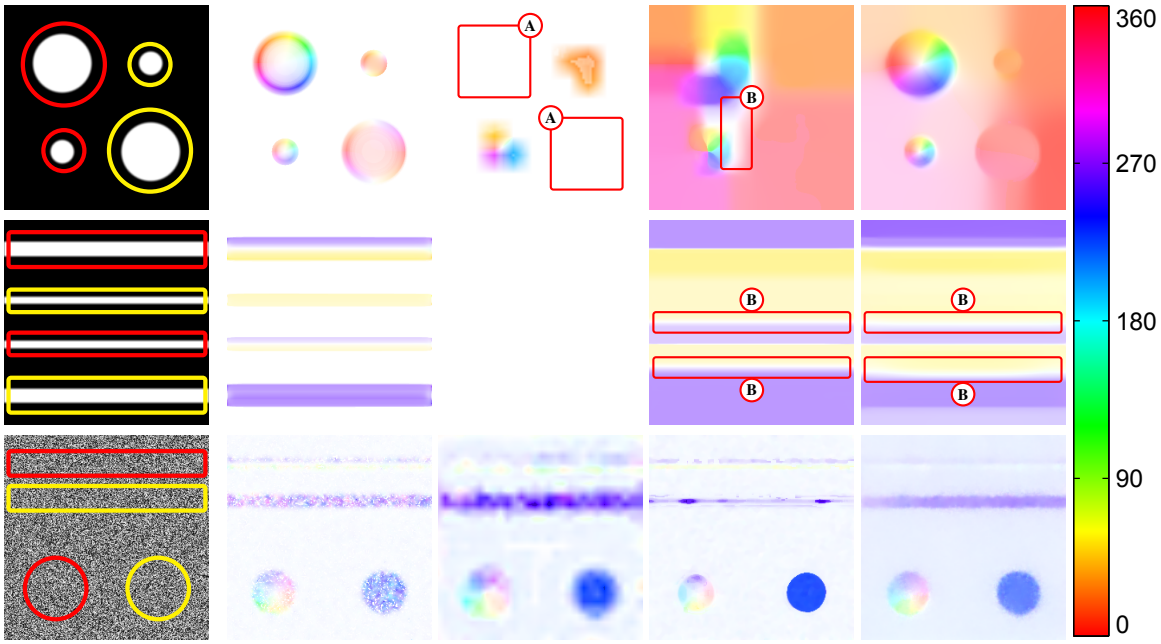
In our phantom experiments (Figure 3.4), the PB implementation is only capable of detecting the



(a) Phantom **(b)** PB **(c)** MF1 **(d)** MF2 **(e)** OF1 **(f)** OF2

Figure 3.4: Qualitative phantom experiments illustrating the results of PB, OF1, OF2, MF1, and MF2 implementations. Top row: phantom with circular shapes of varying sizes. Middle row: tubular shapes of varying sizes. Bottom row: A randomly generated texture subjected to local deformations that include four translating/pulsating circular/tubular structures. The first frame of each phantom sequence is depicted in column (a) with red contours overlaid to indicate structures that distend or contract radially and yellow contours for structures that are subject to translation only. The motion of these structures, estimated in the intermediate flow computation stage of KMVS, are presented separately in Figure 3.5. The color-coded fuzzy segmentation results of different implementations are presented in columns (b-f). The colors range from blue to red, representing weak to strong response to detected vascular structures. All segmentations are thresholded at 0.3 for visibility; responses below this threshold are colored black. The KMVS results, columns (c-f), exhibit more accurate segmentation responses (red) at the center-line of pulsating structures compared to the PBMS method presented in column (b). In this experiment, the qualitative results favor the MF1 implementation of KMVS in column (c). Refer to Section 3.2.3 for detailed explanation of annotations Ⓐ to Ⓔ.

translating structures (Ⓐ in Figure 3.4c) and the edges of the larger pulsating structures (Ⓑ in Figure 3.4c) that appear as local translations. Among the proposed KMVS implementations, MF2 (d) fails to localize the large pulsating structure in the circle phantom (Ⓒ in Figure 3.4) and any of the structures in the tube phantom. This is due to the fact that the MF2 method fails to estimate the motion of large circular structures in the circle phantom (Ⓐ in Figure 3.5c) and any of the tubular structures in the tube phantom (Figure 3.5c). The algorithm cannot detect the motion of the larger circles due to the fact that default window sizes, over which the motion is computed, are too small. Increasing the window size improves the performance when computing the motion of the larger circles, but it increases the computation time and adversely effects the overall performance on the real data presented in the following

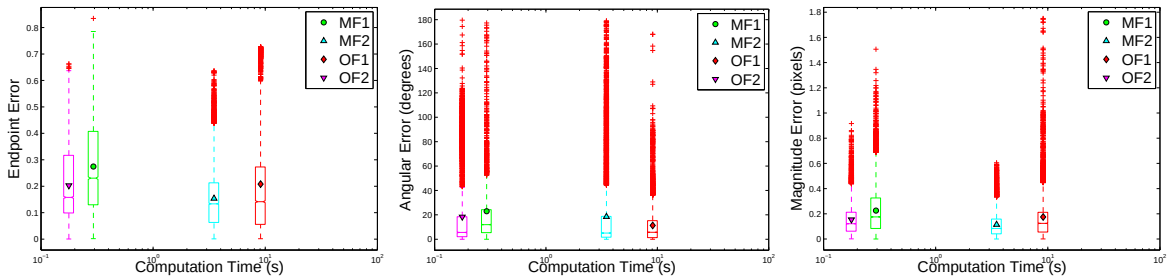


(a) Phantom **(b)** MF1 Motion **(c)** MF2 Motion **(d)** OF1 Motion **(e)** OF2 Motion

Figure 3.5: Intermediate results of the motion estimation stage of different KMVS implementations applied to the phantom dataset. The final segmentation results are presented in Figure 3.4. The PBMS method is not included in this comparison as it does not explicitly compute the flow vectors. Column (a): the first frame of each phantom sequence with red contours overlaid to indicate structures that distend or contract radially and yellow contours for structures that are subject to translation only. Columns (b-e): the estimated 2D motion vectors of each implementation is color-coded such that hue represents the direction and saturation represents the relative magnitude of motion. The color-bar to the right illustrates the mapping of hue to the orientation of motion measured in degrees. Refer to Section 3.2.3 for detailed explanation of annotations Ⓐ and Ⓑ.

section, thus the default values were kept. As for the poor performance on the tube phantom, upon inspecting the algorithm, we noted that this may be due to a stability check performed by the algorithm during motion estimation. We found that relaxing the stability threshold results in some motions being correctly measured on the tube phantom but we ultimately decided to leave the threshold unchanged as lower limits resulted in more errors on other phantoms.

The results of motion estimation presented in Figure 3.5 are visibly different, especially for the circle and tube phantoms, as both of the OF codes used rely on Horn and Schunck [73] style regularization (flow field smoothness constraints) in regions that do not contain salient textures (Ⓑ in Figure 3.5). This regularization sometimes results in falsely occurring divergent behavior and, as a result, false positives in the final segmentation (Ⓓ in Figure 3.4). Although similar false positives are observed on the tube phantom results for MF1 (Figure 3.4c), the strength of the response is lower compared to the response at the center-line. We thus chose to use the third phantom to perform further quantitative comparison



Method	Mean endpoint error (95, 99.3 percentiles)	Mean angular error (95, 99.3 percentiles)	Mean magnitude error (95, 99.3 percentiles)
MF1	0.27(0.64, 0.74)	$23^o(90^o, 97^o)$	0.23(0.59, 0.86)
MF2	**0.16 (0.42, 0.52)**	$19^o(85^o, 160^o)$	**0.11 (0.30, 0.50)**
OF1	0.21(0.71, 0.72)	**11 ^{o}(40^o, 82^o)**	0.17(0.51, 0.93)
OF2	0.20(0.46, 0.57)	$18^o(93^o, 114^o)$	0.15(0.40, 0.59)

Figure 3.6: Boxplots of endpoint errors, angular errors in degrees, and magnitude errors in pixels for each motion estimation methods in the KMVS pipeline. The corresponding mean of the errors are depicted with filled markers and black outline. The errors are computed at every pixel of the noise texture phantom (bottom row of Figure 3.5) by comparing the estimated motions resulting from a two-frame matching experiment against corresponding ground truth motion vector values used to create the phantom. The mean and maximum errors at the 95 and 99.3 percentiles over all pixels are tabulated, with the best performing methods presented in bold. All differences in motion estimation performance were found to be significant, $p < 0.0001$ according to Wilcoxon signed-rank tests using Bonferroni adjusted alpha levels of 0.00056 per test $(0.01/18)$. The OF2 method is the fastest in this experiment and ranks as second best in motion computation performance for all three error metrics.

between the OF and MF estimation modules (presented in Figure 3.6) by computing the error in flow endpoint ε_E to the ground truth $\mathbf{d}_{GT}$ defined as

$$\varepsilon_E(\mathbf{x}, j) = \|\mathbf{d}(\mathbf{x}, j) - \mathbf{d}_{GT}(\mathbf{x}, j)\|_2 \tag{3.14}$$

and the constituting errors in flow orientation ε_O and magnitude ε_M defined as

$$\begin{aligned} \varepsilon_O(\mathbf{x}, j) &= \cos^{-1}\left(\frac{\mathbf{d}(\mathbf{x},j)\cdot\mathbf{d}_{GT}(\mathbf{x},j)}{\|\mathbf{d}(\mathbf{x},j)\|\|\mathbf{d}_{GT}(\mathbf{x},j)\|}\right) \\ \varepsilon_M(\mathbf{x}, j) &= \left|\|\mathbf{d}(\mathbf{x}, j)\| - \|\mathbf{d}_{GT}(\mathbf{x}, j)\|\right|. \end{aligned} \tag{3.15}$$

A number of outliers were observed in the boxplots of orientation ε_O and magnitude ε_M errors presented in Figure 3.6. This is indicative of the fact that the errors follow a heavily skewed distribution. To give an accurate representation of the distribution of errors for each method, we report the the mean error and the maximum errors at the 95 and 99.3 percentiles over all pixels in Figure 3.6. As these error distributions are skewed, we tested the differences between each method for statistical significance using a Wilcoxon signed-rank tests with Bonferroni adjusted alpha levels of 0.00056 per test $(0.01/18)$ to correct for multiple comparisons and observed that all differences were significant $p < 0.0001$.

In quantitative comparisons of flow estimation methods, OF2 is the fastest and ranks as second best in motion computation performance for all three error metrics. However, combined with the PRMM, the motions extracted from OF2 fail to detect the small pulsating tubular structure (Ⓔ in Figure 3.4f) in the noise phantom. Reducing the regularization weight α to 0.01 improves the detection of the small pulsating tubular structure, however we observed that regularization weights outside the range of $0.015 < \alpha < 0.1$ consequently result in noticeable reduction in motion estimation performance on synthetic data and segmentation accuracy on UBC dataset experiments presented in the following section.

Real Data Experiment

Initial real data evaluation is conducted on the UBC dataset, consisting of eight 30- to 40-frame DUS sequences of the CCA acquired along the transverse and longitudinal axes, where the vessel appears as a pulsating ellipsoid and tube respectively. Unlike the previous two-frame phantom experiment, experiments with real data require the temporal filtering stage to remove the high frequency noise and the low frequency motions (caused by breathing and small movements of the probe) that occur in the sequence. The passband of the temporal filter is tuned to the patient's approximate heart rate, denoted τ_r. The parameters were set such that $\tau_L = \tau_r/2$ and $\tau_H = 2\tau_r$. The PBMS segmentation method and the four implementations of KMVS (Section 3.2.1) are then applied to the dataset using the same temporal filter parameters across all methods. All of the resulting fuzzy segmentation labels are shown in Figure 3.7. To clarify the advantages of each approach as a trade-off between accuracy and computation time, in Figure 3.8, we present quantitative analysis of segmentation error using the ground truth segmentations of the DUS sequences. The area under the receiver operating characteristics curve (AUC) for each case (thresholding the fuzzy segmentations from 0 to 1) is reported as a measure of segmentation accuracy, in which the value of 1 indicates perfect segmentation and 0.5 is the noise threshold.

Once the parameters of each implementation have been tuned on the phantom and UBC datasets, the real data experiments are repeated using the larger publicly available SPLab dataset. Only the first four frames of each sequence, recorded at an estimated three frames per second, was used to generate the results presented in Figure 3.9, Figure 3.10 and Figure 3.11. Due to the small number of frames and low frame rate of each sequence in this dataset, the temporal bandpass filter becomes a highpass filter that passes all of the temporal frequencies except for the zero frequency, DC gain, component.

In addition to the quantitative evaluation of segmentation errors using AUC, we also compare the results of each implementation by computing the Dice similarity coefficient (DSC) between the computed segmentations and corresponding ground truth of each dataset. To compute the DSC, the fuzzy segmentation labels generated using each implementation were binarized at a threshold of 0.5. The mean DSC and AUC of all cases in each dataset is computed and tabulated in Table 3.2. The summarized results indicate that the OF2 implementation of KMVS is the best performing method in terms of DSC and AUC. Detailed analysis of the presented results is carried out in the following section.

Discussions

Our experimental results show that all of the KMVS implementations outperform the PBMS approach

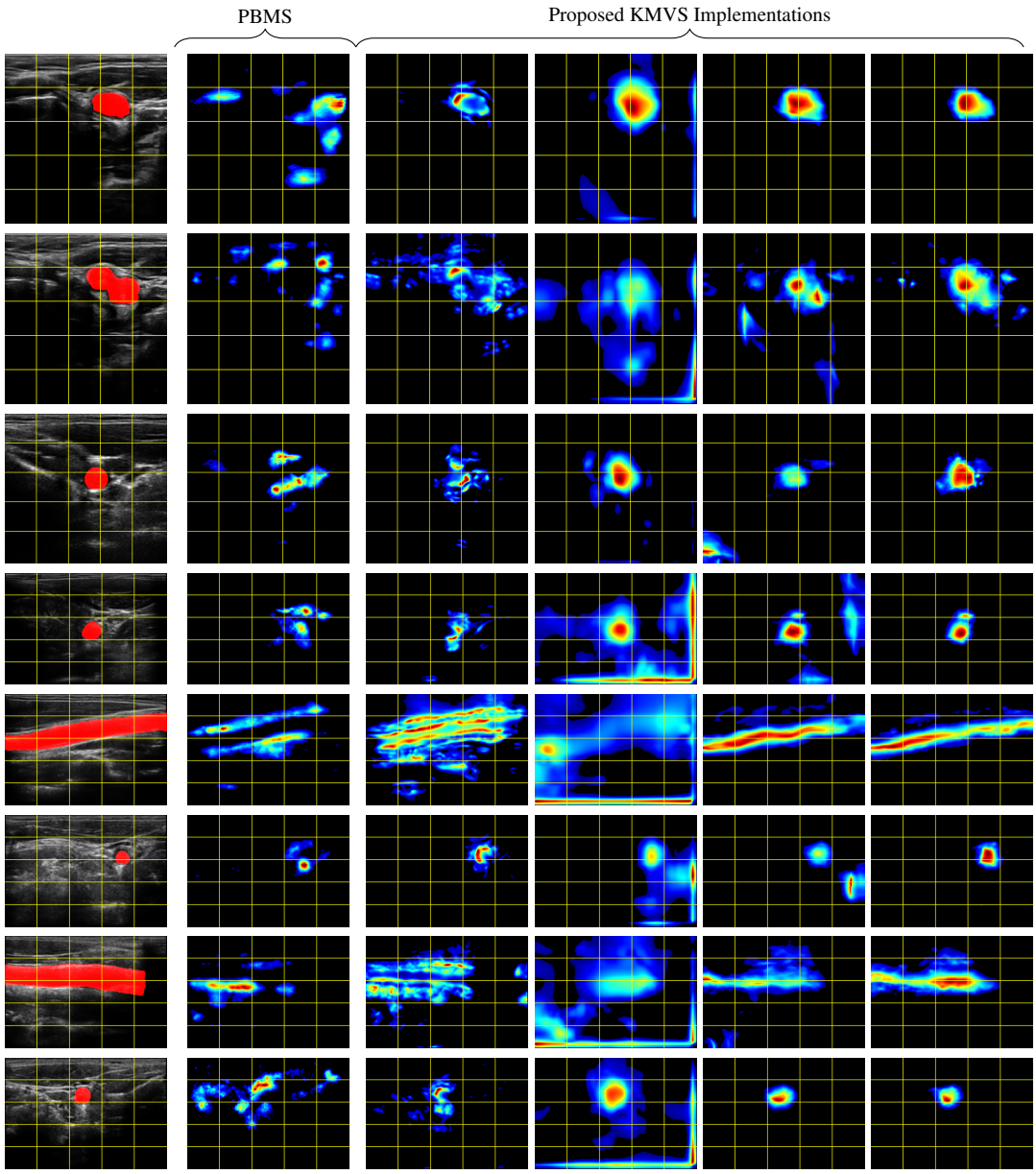
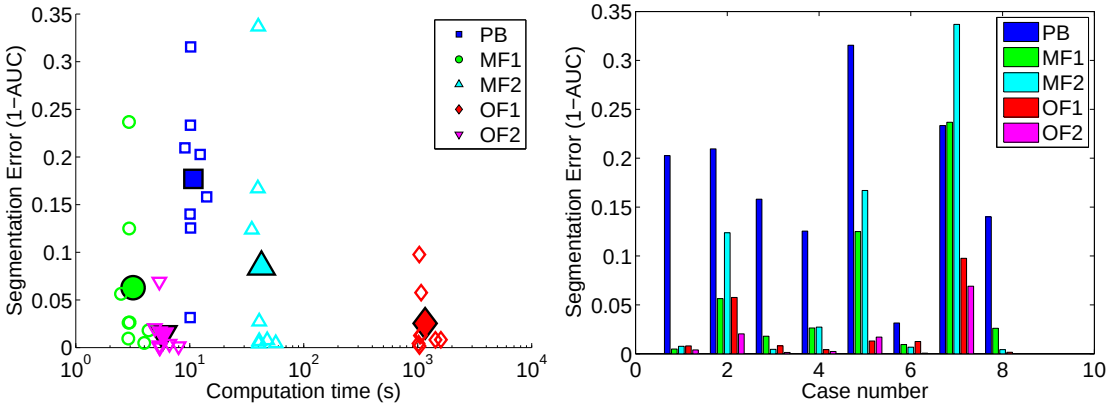


Figure 3.7: Qualitative results of our experiment on the UBC dataset with yellow grid-lines superimposed to facilitate correspondence. Column (a): first frame of DUS sequences of CCA in axial and longitudinal axes including the bifurcation of internal and external carotid arteries. The corresponding US ground truth for the vessel is shown in red. Columns (b-f): color-coded fuzzy segmentation results of different implementations. The colors range from blue to red, representing weak to strong response to detected vascular structures. Segmentations are thresholded at 0.3 for visibility; responses below this threshold are colored black.



Method	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Mean
PB	0.20	0.21	0.16	0.13	0.32	0.032	0.23	0.14	0.18
MF1	0.0049	0.056	0.018	0.026	0.13	0.0095	0.24	0.026	0.063
MF2	0.0077	0.12	0.0046	0.027	0.17	0.0068	0.34	0.0042	0.085
OF1	0.0081	0.058	0.0082	0.0042	**0.013**	0.013	0.098	0.0015	0.025
OF2	**0.0040**	**0.020**	**0.0014**	**0.0023**	0.017	**0.0006**	**0.069**	**0.0003**	**0.014**

Figure 3.8: Quantitative performance of our experiment on the UBC dataset. Left: Performance of each segmentation method illustrating the trade-offs between computation time and segmentation performance; corresponding averages depicted with filled markers and black outline. Right: Bar chart of segmentation errors, grouped according to case number. Segmentation error values (1-AUC) are tabulated with the best performing methods in bold.

Table 3.2: Quantitative summary of segmentation performance of the experiments performed on real data presented as mean Dice similarity coefficient (DSC) and mean area under the receiver operating characteristics curve (AUC) across all cases of the UBC and SPLab datasets. The fuzzy segmentations were thresholded at 0.5 in order to compute the DSC. The OF2 implementation of the proposed KMVS pipeline boasts the highest performance (shown in bold) compared to PBMS and other implementation of KMVS.

Name	Method	UBC Dataset		SPLab Dataset	
		DSC	AUC	DSC	AUC
PB	PBMS	0.11±0.10	0.82±0.084	0.20±0.096	0.83±0.070
OF1	KMVS	0.60±0.15	0.97±0.034	0.53±0.28	0.93±0.12
OF2		**0.72±0.097**	**0.99±0.023**	**0.67±0.21**	**0.98±0.027**
MF1		0.49±0.095	0.94±0.080	0.42±0.18	0.95±0.051
MF2		0.36±0.18	0.92±0.12	0.44±0.26	0.88±0.16

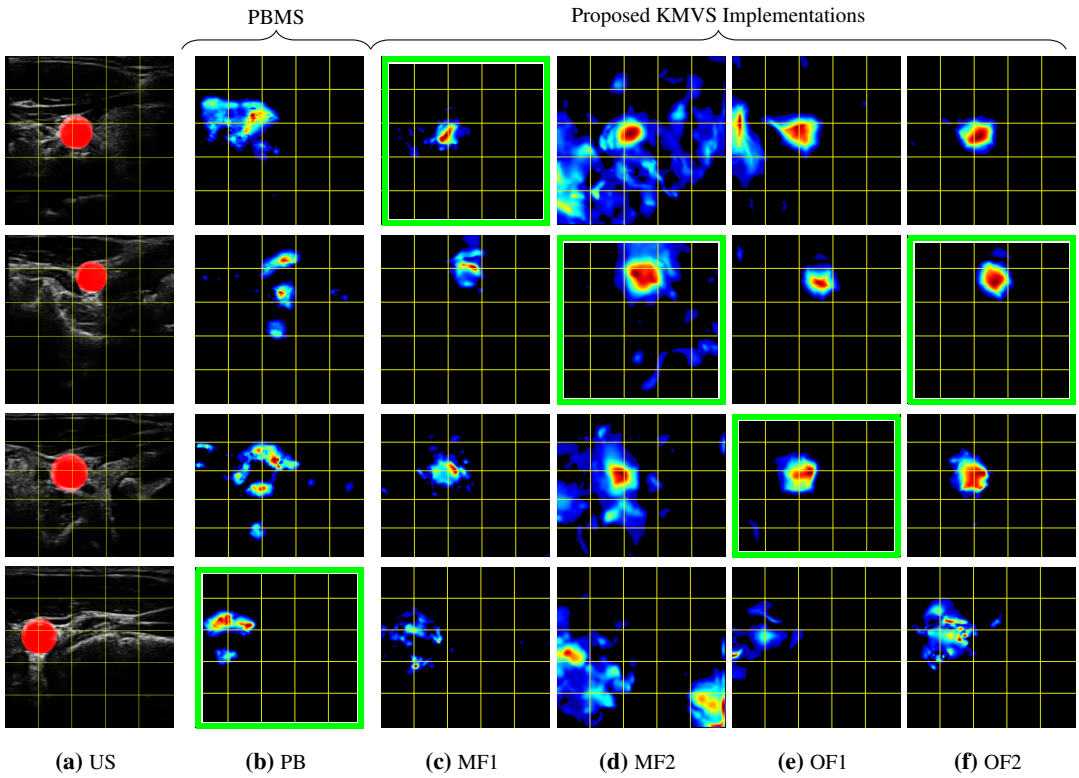


Figure 3.9: Four exemplar cases illustrating the best results of our experiment with the publicly available SPLab dataset with yellow grid-lines superimposed to enhance correspondence. Column (a): first frame of DUS sequences of CCA in axial and longitudinal axes including the bifurcation of internal and external carotid arteries. The corresponding US ground truth for the vessel is shown in red. Columns (b-f): color-coded fuzzy segmentation results of different implementations. The colors range from blue to red, representing weak to strong response to detected vascular structures. All segmentations are thresholded at 0.3 for visibility; responses below this threshold are colored black. The best performing out of all 35 cases are framed in green.

in terms of segmentation accuracy. Compared to PB, with the addition of our proposed PRMM, our MF and OF pipelines are more specific to motion of the CCA and resilient to motions that occur on the surrounding soft tissues. In the phantom experiments (Figure 3.4), we explicitly showed how the PB implementation is only capable of detecting the translating structures (Ⓐ in Figure 3.4) and the edges of the larger pulsating structures (Ⓑ in Figure 3.4) that appear as local translations. On the other hand, the KMVS segmentation results presented in columns (c-f) of Figure 3.4 are closer to center-line of the pulsating structures. This trend is also evident in the qualitative results of the real data experiments presented in Figure 3.7, Figure 3.9 and Figure 3.10, and is further substantiated by the quantitative analysis of segmentation accuracy presented in Figure 3.8, Figure 3.11 and Table 3.2. However, contrary to our initial conclusions on real data experiments that favored MF methods over OF [6], we show that it is possible to outperform our proposed MF1 approach using the tuned OF2 algorithm.

Both OF implementations presented in this paper tend to perform well on DUS sequences as the

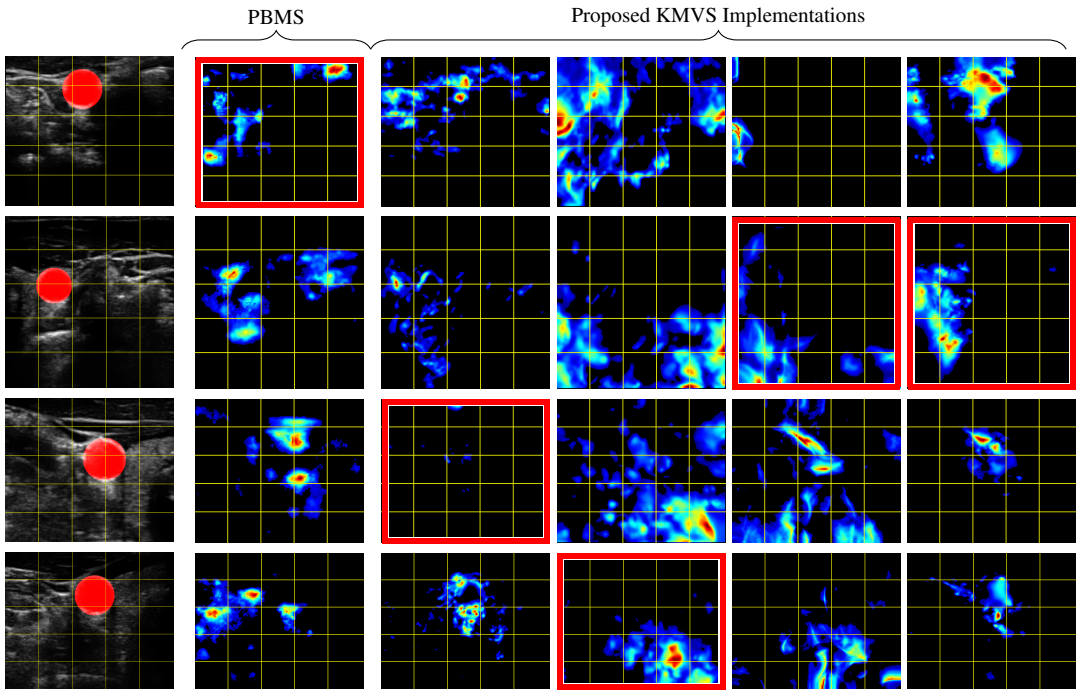


Figure 3.10: Four exemplar cases illustrating the worst results of our experiment with the publicly available SPLab dataset with yellow grid-lines superimposed to enhance correspondence. Column (a): first frame of DUS sequences of CCA in axial and longitudinal axes including the bifurcation of internal and external carotid arteries. The corresponding US ground truth for the vessel is shown in red. Columns (b-f): color-coded fuzzy segmentation results of different implementations. The colors range from blue to red, representing weak to strong response to detected vascular structures. All segmentations are thresholded at 0.3 for visibility; responses below this threshold are colored black. The worst performing out of all 35 cases are framed in red.

global flow smoothness constraint enables OF to approximate the motion of tissues in locations that are void of salient image information, i.e., the center of the vessel. This added constraint increases the computational complexity of the algorithm but, by comparing OF1 to a manually tuned OF2 (Figure 3.8 and Figure 3.11), we demonstrate that it is possible to obtain low segmentation error without increasing the run-time of OF motion computation. The motion estimation errors presented in Figure 3.6 further confirm that the manually tuned OF2 implementation produces errors that are comparable to the automatically tuned OF1 implementation, while maintaining a run-time that is faster by almost two orders of magnitude across all phantom and real data experiments. The only notable difference between the two OF methods is that the mean angular error of OF2 is greater than that of OF1 and, as a result, OF1 is better at localizing the small pulsating tubular structure in our noise phantom compared to OF2 (Ⓔ in Figure 3.4f). Regardless, OF2 outperforms OF1 on all real DUS data experiments (Figure 3.8 and Figure 3.11). From this observation, we conclude that although the computationally more expensive

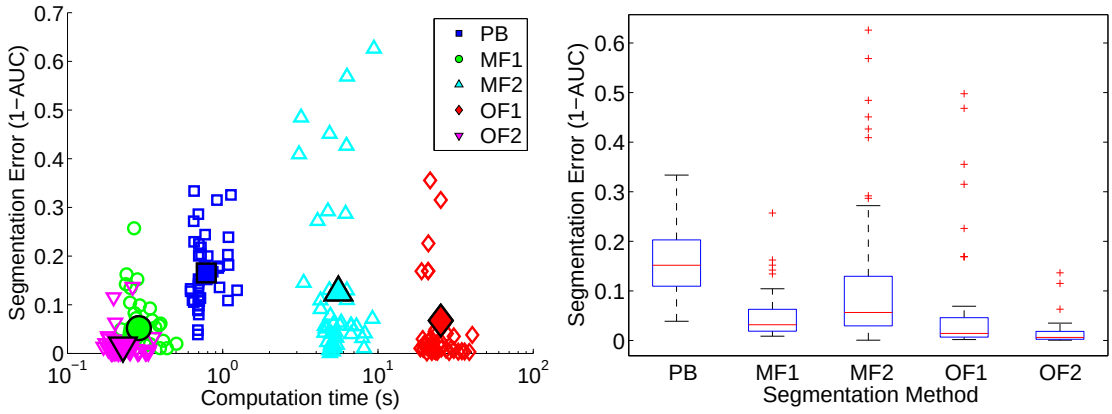


Figure 3.11: Quantitative performance of experiment on 35 sequences from the publicly available SPLab dataset. Left: Performance of each segmentation method illustrating trade-offs between computation time and segmentation error; corresponding averages depicted with filled markers and black outline. Right: Boxplots of segmentation errors for each method.

OF1 method generates better estimates of motion vectors, the much faster OF2 implementation performs better in tandem with our proposed PRMM technique and thus generates more accurate vessel localizations in DUS sequences.

The quantitative results presented in Figure 3.8 and Figure 3.11 also show that the MF1 method can, on average, achieve comparable accuracy with the OF1 method. The comparable performance of MF1 to other flow estimation methods is further confirmed by the analysis of the endpoint errors presented in Figure 3.6. In terms of run-time, both OF2 and MF1 implementations are suitable for diagnostic applications, e.g., CCA segmentation, as they are projected to perform in near real-time with our imaging setup given an efficient implementation and a specialized workstation. MF1 is slower than OF2 in the two-frame phantom experiment (Figure 3.6) and the experiment on the SPLab data (Figure 3.11), but it performs faster on the UBC data (Figure 3.8). This is attributed to the computational overhead cost of building the SQF for each sequence. The MF1 method is designed in such a way that the filter bank containing the SQF pyramid is constructed only once per sequence and, as a result, this one-time computational cost dominates the overall run-time in cases where the number of frames are few, such as in the phantom and SPLab data experiments. In the UBC data experiments, where each sequence contains 30–40 frames, our MF1 implementation is faster than OF2. MF1 is also parallelizable as most of the computations performed are point-wise (pixel-wise) operations, which do not need a Horn and Schunck [73] style global smoothness constraint. By precomputing the filter bank and porting the code to run on a graphical processing unit (GPU), it is possible to achieve further performance gains and enable the method to run in real-time on sequences with larger images [3, 127].

Our UBC dataset was recorded at 30 frames per second while the SPLab dataset was recorded at an estimated three frames per second. This translates to a recording window of 1 to 1.3 seconds in duration for both datasets. We found this recording length to be empirically sufficient for localizing the CCA as it captures the periodic behavior of a patient with a resting heart rate as low as 60 beats per minute or 1

Hz. The discriminative performance of our proposed methods would increase with a larger number of samples, however with the current temporal filtering scheme, a larger recording window would result in a delay (lag time) greater than the current 1.3 seconds, which may not be reasonable for some clinical applications. If the acquisition is not corrupted by prominent motion artifacts, e.g., probe movement or swallowing, we hypothesize that the periodic behavior of the CCA may be detected with a shorter recording window, up to twice the heart rate frequency, which amounts to a 0.5 second window. In real-time applications, our method may be combined with fast tracking algorithms and on-line model fitting techniques such as extended Kalman filtering [111] to compensate for this lag.

The proposed temporal filtering scheme is not well suited for applications that impose hard constraints on temporal performance and lag times, such as intraoperative image-guidance systems. The ideal filter employed in our paper is non-causal and assumes knowledge of future inputs to the filter. This ideal filter approach gives a single response over the 1 to 1.3 seconds recording window. For real-time applications, causal filters such as infinite impulse response filter used in Wu et al. [202] and general biquad filter employed by McLeod et al. [110], or the aforementioned extended Kalman filtering [111] approach are more suitable alternatives that can be easily incorporated into our proposed pipeline.

The motion estimation method used in MF2 was originally developed to extract motion from medical imaging modalities, e.g., DUS and dynamic magnetic resonance (MR) imaging, in which the brightness consistency assumption does not hold [2]. Despite performing well on the synthetic phantom experiments, producing the lowest endpoint error among all methods (Figure 3.6), MF2 does not perform as well as other KMVS methods on the real datasets. We hypothesize that this might be due to the fact that MF2 was specifically designed to compute large myocardial motion as opposed to the subtle motion of vascular walls, i.e., the CCA.

Although the OF2 pipeline outperforms both MF methods in our real data experiments, monogenic methods have other advantages that may be exploited to improve vessel localization. In the domain of US image processing, SQF have been shown to improve the extraction of structures, during radio frequency (RF) signal to B-mode conversion, by demodulating the RF in a 2D context [194]. This implies that MF methods may be used to extract local motion information from raw RF data, allowing for a direct implementation in the native representation of acquired DUS data.

Another advantage is that the MF implementations presented may be extended to three spatial dimensions [1]. Our approach is not yet able to cope with gross out-of-plane motion, common during 2D+time DUS acquisitions, but such problems would not exist once the method is extended to process 3D+time sequences. A study on 3D steerable wavelets and the monogenic signal [27] concluded that the steerable approach of Portilla and Simoncelli [132], which was used in the PBMS method, is hard to extend into 3D due to its invertibility but, on the other hand, the MF formulation can be extended to process 3D+time sequences [1, 27]. Similarly, 3D+time extension of the OF implementations are also possible, e.g., 3D OF between two volumes followed by divergence computation of the estimated 3D vector field and multi-scale averaging. This will allow our proposed MF and OF implementations to be extended to volumetric medical images such as 3D+time computed tomography (CT) fluoroscopy, Cine MR imaging, and 3D DUS. Due to the aforementioned fact that the MF1 implementation largely

consists of point-wise operations, we hypothesize that a 3D implementation of MF1 would outperform 3D Horn and Schunck [73] style OF methods in terms of speed. We consider this to be an important advantage of our MF implementation as 3D+time data of vasculature is prevalent in cardiac motion analysis [1, 2], imaging of coronary arteries [90], monitoring the distensibility of vasculature [23, 65], and other applications; with more applications likely to emerge as the availability and quality of 3D DUS continues to improve over time.

3.3 Summary

In this chapter, we assessed the performance of vessel localization using automatic segmentation algorithms that model pulsatile radial motion—a more robust extension of the PBMS methodology presented in the previous chapter. We presented four implementations of a low-level motion-based segmentation pipeline, which detects the characteristic pulsatile radial motion of vascular structures through the analysis of divergent motion vector fields extracted from a sequence of frames. Our proposed methods are focused solely on fast and accurate extraction and modeling of motion vector fields, without the aid of learning and appearance models, such that our segmentation labels may be incorporated alongside complimentary low-level intensity based features into existing high-level discriminative segmentation frameworks [13, 71].

In each implementation, we explored alternative optimized and off-the-shelf techniques for performing the motion estimation stage of the proposed pipeline. Through evaluation of proposed methods, using synthetic and real DUS sequences of the CCA, we conclude that coupling a tuned OF motion estimation method with our PRMM provides the best overall performance compared to other candidates. Our experiments on two real DUS datasets show that the performance of the old PBMS method can be increased using the tuned OF2 implementation of KMVS from an average AUC of 0.82 to 0.99 on the UBC dataset and from 0.83 to 0.98 on the SPLab dataset. Similarly, binarizing the resulting fuzzy segmentation labels at 0.5 yields a notable increase in DSC from 0.11 to 0.72 on the UBC dataset and 0.20 to 0.67 on the SPLab dataset. Furthermore, the tuned OF2 implementation of the pipeline performs the fastest on short sequences compared to all other implementations, while our MF1 implementation boasts the fastest run-time on longer sequences due to the precomputation of the SQF.

The computational speed of our proposed MF1 motion estimation method can also greatly benefit from a parallel implementation that takes advantage of GPU computing [3]. The monogenic framework used to extract motions in the MF1 implementation may also be able to estimate motion from native RF data of DUS sequences [194] and can be extended to process volumetric dynamic (3D+time) data efficiently [1, 27]. The Horn and Schunck [73] style OF methods, on the other hand, do not scale as well in comparison as the spatial dimensionality of the data increases from 2D to 3D. As a result, point-wise motion estimation methods such as MF1 are poised to outperform competing methods as the use of volumetric dynamic imaging of vasculature becomes more prevalent in image-based medical diagnosis and interventions.

Chapter 4

Uncertainty-Encoded Augmentation of the Surgical Scene

"Information is the resolution of uncertainty."
— Claude Shannon

In Chapter 2 and Chapter 3, we demonstrated the utility of temporal modelling in automated analysis of dynamic medical image sequences. This chapter presents our preliminary phantom experiments towards addressing our second research question; that is, to explore avenues in which navigation uncertainties may be computed, propagated, and visualized for interventions that target deformable tissues. Our contributions investigate a proof-of-concept augmentation method for encoding navigation uncertainties during image-guided minimally invasive robot-assisted partial nephrectomy (RAPN). The presented framework encodes uncertainties through the computation and visualization of uncertainties that may occur during preoperative computed tomography (CT) segmentation and computational stereopsis steps of RAPN navigation.

4.1 Towards Probabilistic Tumor Demarcation in Image-Guided Surgical Interventions

In Chapter 2 of this thesis, we motivated the benefits of medical robotic technologies in the context of nephron sparing interventions. We stated that the RAPN procedure is organized into five main stages: (i) Bowel mobilization; (ii) Hilar dissection and control; (iii) Identification and demarcation of tumor margins; (iv) Resection of tumor; and (v) Reconstruction of the kidney. In this chapter, we focus on the crucial step of tumor identification, during which the surgeon localizes the kidney tumor mass and identifies the resection margins. This step is important to properly plan and speed up the succeeding stage of tumor mass excision during which blood flow can only be safely obstructed for a limited time. More importantly, the accuracy of this step is necessary not only to preserve kidney function by sparing as much healthy tissue as possible, but also to avoid tumor recurrence by excising all cancerous tissue.

The tumor identification step is usually performed with the help of multimodal sources of infor-

mation at surgeon's disposal: preoperative scans (typically 3D CT and/or magnetic resonance (MR)) and intraoperative data (2.5D stereo endoscopic data and, when available, laparoscopic 2D/3D ultrasound (US)). Currently, these rich and complementary sources of information are just displayed on the surgeon's console in a tiled fashion, i.e., side-by-side, or even sometimes on a separate screen of a workstation nearby. These typical display setups require substantial additional effort from the surgeon to piece together a 3D mental map of the surgical scene that integrates all information together in order to localize the tumor and adjacent tissue. Hence, an augmented reality view, in which the endoscopic video stream is overlaid with highlighted kidney and tumor boundaries, can substantially reduce the effort required by the surgeon to achieve accurate and quick tumor excision.

The fusion of data in RAPN for the purpose of augmenting the surgical scene is, however, challenging considering the different reference frames of acquisitions, the heterogeneity of the imaging modalities, and the likely motion and deformation of organs when the patient's abdomen is insufflated during surgery.

Among the proposed computer-assisted image-guided therapy (IGT) solutions, seminal efforts necessitated the use of invasive fiducials to provide further details regarding the current position and shape of the kidney and tumor [12, 185, 206]. At the time of the publication of our proposed contributions, state-of-the-art IGT navigation systems that did not require the use of invasive tools mostly relied on an initial manual rigid alignment of preoperative segmentations to stereo data. This alignment was then followed by a motion tracking component that can be deterministic [174, 177], probabilistic [56], or biomechanically constrained [134]. Other notably relevant works focused on developing more robust [15] and real-time [150, 175] stereo surface reconstruction algorithms to which they aimed to register preoperative segmentations.

Computation of a 3D surface from stereo video data had shown to be an important aspect of the IGT guidance systems but the efforts pertaining to 3D computational stereopsis were limited to deterministic results. In computer vision research, attempts had been made to model the encoding of 3D geometry in a pair of 2D images using Bayesian frameworks [14]. Probabilistic stereo reconstruction had been proposed [74, 87] and probabilistic scene analysis had also been used to detect smooth problem areas prior to matching [172]. The goal of these approaches were to use this probabilistic model to arrive at a globally optimal depth reconstruction. However, none of the probabilistic reconstruction methods had attempted to project the probabilistic cost (disparity) map directly into 3-space to create a probabilistic representation of the surface. This probabilistic surface is an important source of navigational uncertainty in RAPN that has not been computed and accounted for.

Moreover, all of the surveyed computer-assisted IGT navigation systems proposed for tumor demarcation in the context of RAPN were limited to the visualization of a crisp segmentation only [135]. Such crisp visualizations, which do not encode uncertainties, render the surgeon susceptible to the varying levels of confidence in what is overlaid on the screen. Like the computational stereopsis stage, segmentations are hardly ever perfectly accurate for many possible reasons: graded composition [191], image acquisition artifacts, inter-expert segmentation variability, and fuzzy image segmentation [58, 210]. These uncertainties can be important in subsequent analyses and decision-making [191, 197].

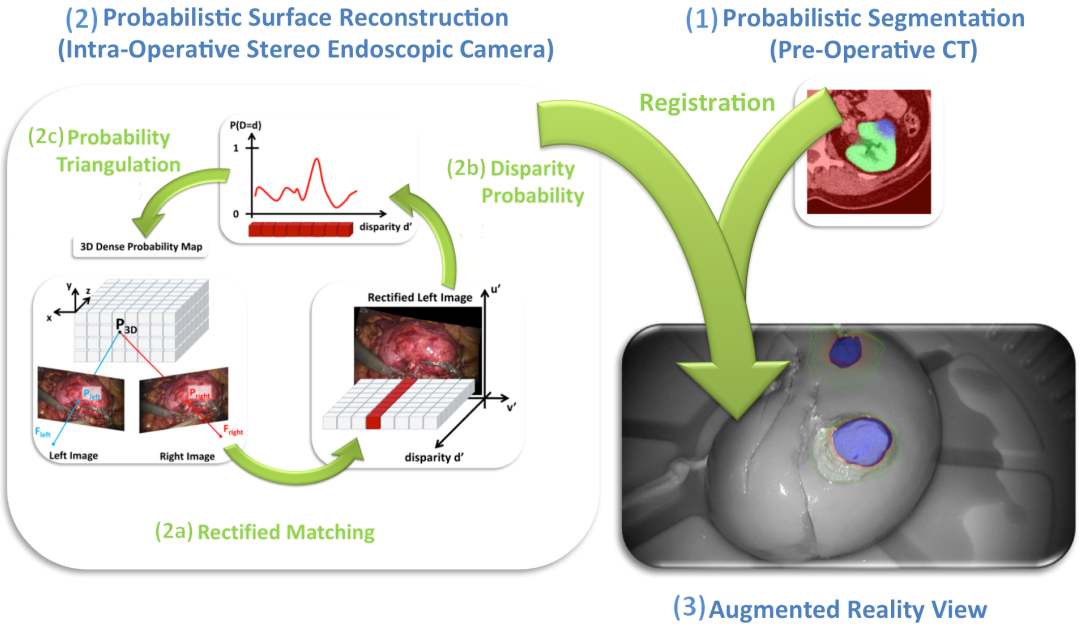


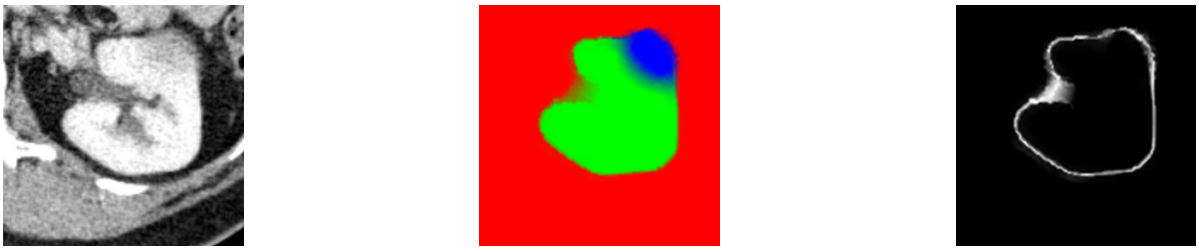
Figure 4.1: Proposed augmented reality framework for kidney tumor demarcation. 3D dense probabilistic map of surface reconstruction from stereo endoscopic data and probabilistic segmentation of preoperative CT are co-registered to create a stereo endoscopic view augmented with tumor boundaries and corresponding uncertainties to guide resection planning.

We propose to provide a visualization of uncertainties at the kidney and tumor boundaries as a visual cue to assist the surgeon in finding the optimal resection strategy. This is similar in concept to what is currently being explored in radiotherapy for brain tumors when extrapolating glioma invasion with variable margins [92].

Our visual cues are derived from shape boundary uncertainties in the probabilistic segmentation of the preoperative CT. To do so, as summarized in Figure 4.1, we extract 1) the kidney/tumor boundaries in the CT prior to the operation and 2) corresponding uncertainty information from the stereo-endoscopic views intraoperatively using (2a) computation stereo matching techniques [68], (2b) converting matching weights into probability values, and (2c) triangulating the surface probabilities into the same domain as the CT. Finally, 3) we register the preoperative boundary uncertainties to the live probabilistic reconstruction from stereo. We apply our method to an *ex vivo* lamb kidney to create an uncertainty-encoded augmented reality view. We compare our results to standard guidance methods that use crisp segmentations and present the potential added benefits of our method and its utility for resection planning.

4.2 Uncertainty-Encoded Probabilistic Resection Margins

We first describe the probabilistic segmentation of the preoperative CT that provides uncertainties about the boundary localization of kidney and tumor. Secondly, we perform a probabilistic 3D surface reconstruction from stereo endoscopy to which the probabilistic segmentation is directly registered.



(a) (b) (c)

Figure 4.2: Probabilistic preoperative CT segmentation. (a) Original CT. (b) Membership probabilities of kidney (green), tumor (blue), and background (red). (c) Background boundary location probability (0 in black and 1 in white).

4.2.1 Probabilistic Segmentation of Preoperative Image Volumes

The probabilistic segmentation of the preoperative CT is based on the random walker segmentation algorithm [58, 167] that generates membership probabilities of three manually seeded regions: background (BG: red), kidney (KD: green), and tumor (TM: blue) as shown in Figure 4.2b.

We denote the resulting multi-label probabilistic CT segmentation by:

$$P_{seg}^{CT} : \Omega \subset \mathbb{R}^3 \rightarrow \mathbf{p} \in \mathscr{S}^2 \, ,$$

where $\mathbf{p} = [p_{BG}, p_{KD}, p_{TM}]$ belongs to the simplex of order 2, and Ω is the region of interest of the CT. From this multi-label probabilistic segmentation, we can extract the membership probability map of background P_{BG}^{CT}, kidney P_{KD}^{CT} and tumor P_{TM}^{CT} regions.

We also compute the likelihood $P_{surface}^{CT}$ of the *surface* union of kidney and tumor in the preoperative CT (Figure 4.2c) by combining the membership probabilities of being *inside* the kidney P_{KD}^{CT} and inside the tumor P_{TM}^{CT} as follows:

$$P_{surface}^{CT} = 1 - \frac{|(P_{KD}^{CT} + P_{TM}^{CT}) - 0.5|}{0.5} \, . \tag{4.1}$$

4.2.2 Probabilistic Stereo-Endoscopic Surface Reconstruction

We propose an extension of traditional stereo surface reconstruction from a single crisp surface [68] to a probabilistic representation of surfaces in 3-space.

Dense Matching of Left and Right Stereo Images

Using polar rectification [131] with the camera calibration parameters, the 2D dense matching of left and right stereo images is simplified to a 1D matching along parallel epipolar lines in the left and right rectified images. We use the normalized cross-correlation (NCC) ratio on greyscale images as a matching similarity metric. This metric has the advantage of being less prone to changes in illumination. In contrast with current state-of-the-art methods, e.g., Stoyanov et al. [175], Bernhardt et al. [15], and Röhl et al. [149], instead of computing one set of robust and optimal matches, we retain *all* possible

matches with their associated disparity (displacement $d \in \mathbb{Z}$ between matching points along the same horizontal line of the rectified images) and similarity measure ($c \in [-1,1]$).

Construction of a 3D Probabilistic Voxel Map

In order to facilitate the preoperative to intraoperative registration detailed in Section 4.2.3, we first create a 3D probabilistic voxel map in which each voxel stores the probability of being at the surface of the stereo endoscopic scene. To achieve this, we compute the disparity probability values by converting the NCC profile $\mathbf{c} = [c_1, c_2, \cdots, c_{N_d}]$ computed previously at every pixel $(u,v) \in \Omega_{2D} \subset \mathbb{R}^2$ in one of the rectified images for different disparities $d \in \mathscr{D} = \{d_1, d_2, \cdots, d_{N_d}\}$, where N_d is the number of inter-digit signed disparities. Basically, the NCC profiles are stacked into a 3D correlation map:

$$NCC_{3D}^{stereo} : (u,v,d_i) \in \Omega_{3D} \rightarrow c_i \in [-1,1] \tag{4.2}$$

and converted into a 3D probabilistic voxel map using the Gibbs measure as follows:

$$P_{3D}^{stereo}(u,v,d_i) = \frac{\exp\left(-\beta\left(\max_d\left(NCC_{3D}^{stereo}(u,v,d)\right) - NCC_{3D}^{stereo}(u,v,d_i)\right)\right)}{W(\beta)}\ , \tag{4.3}$$

where $W(\beta) = \sum_d \exp\left(-\beta\left(\max_d\left(NCC_{3D}^{stereo}(u,v,d)\right) - NCC_{3D}^{stereo}(u,v,d_i)\right)\right)$ is the partition function, and β is a free parameter.

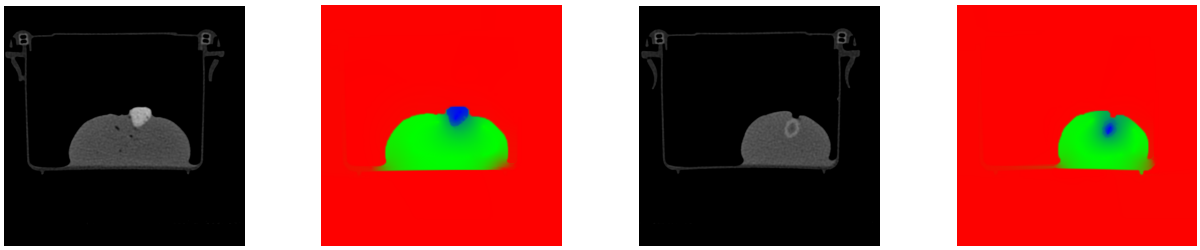
Finally, the 3D position of each matched pair of points in the stereo views is triangulated with the camera projection matrices to transform P_{3D}^{stereo} into a probabilistic voxel map $P_{surface}^{stereo}$ in real world 3D space:

$$P_{surface}^{stereo} : (x,y,z) \in \Omega_{3D} \rightarrow [p, 1-p] \in \mathscr{S}^1\ , \tag{4.4}$$

where $p \in [0,1]$ is the likelihood of the tissue surface being reconstructed at the discretized position (x,y,z) in real world 3D space Ω_{3D} with our computational stereopsis method.

4.2.3 Registration of Stereo Camera and Preoperative Segmentations

We initialize the registration of the CT to the stereo camera in a semi-automatic manner using manually matched landmarks between original CT, left and right camera views. In this first step, we use a similarity transformation to model the combination of Equation 4.1 a rigid transformation to cope with different reference frames between stereo camera and CT acquisitions and Equation 4.2 a global scaling to cope with ambiguities resulting from possible camera calibration errors. The resulting transformation is then refined with an automatic similarity registration of $P_{surface}^{CT}$ to $P_{surface}^{stereo}$ obtained respectively from Equation 4.1 and Equation 4.4. Finally, a non-linear registration step of these two volumes with a B-Spline transformation model is performed to cope with deformations occurring between the preoperative CT acquisition and the surgical scene. We used `elastix` [91] with the sum of squared differences similarity metric for the two last automatic registration steps.



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

Figure 4.3: Transverse slices of CT volume depicting our *ex vivo* lamb kidney phantom with (a) an exophytic and (c) an endophytic artificial tumor. (b) and (d) are probabilistic Random Walker segmentations of (a) and (c), respectively. Tumor labels are colored blue, kidney is colored green, and the background is red.

4.3 Experiments

4.3.1 Materials and Experimental Setup

For validation purposes, we fabricated an *ex vivo* phantom using a lamb kidney and implanted artificial tumors inside it. Different materials (chewing gum and olive pit) were used to emulate low and high contrast kidney-tumor boundaries within the CT. The chewing gum was placed on the surface of the kidney to emulate a partially exophytic tumor/cyst (Figure 4.3a) and the olive pit was planted deep inside the kidney (close to the renal pelvis) representing a completely endophytic tumor (Figure 4.3c).

A 16 slice Siemens SOMATOM Sensation CT scanner was used to acquire a high resolution CT volume of the phantom. The resulting volume is composed of 130 (0.600 mm thick) transverse slices of 512×512 pixels (0.215 mm pixel spacing). Stereo endoscopy data was captured with a calibrated da Vinci S Surgical System at full HD 1080i resolution.

4.3.2 *Ex vivo* Lamb Kidney Study

The Random Walker segmentation algorithm was applied with manual seeding of each label in the CT volume. The probabilistic labeling corresponding to the two simulated tumors is illustrated in Figure 4.3b and Figure 4.3d. Note that the diffusion of uncertainties in the endophytic case is more visible compared to the exophytic tumor; this is a direct result of weaker contrast (CT intensity values: difference in pit/gum composition) at the kidney-tumor boundary. We were careful to keep the distances between the manually placed seeds and the visible boundaries constant to decrease the influence of seed placement on the resulting segmentations.

As illustrated in Figure 4.4a, our phantom is quite smooth and lacks unique features on its surface. This results in a largely uncertain reconstruction from our stereo matching algorithm, which in turn causes the registration to be sensitive to the initial pose estimation. Successful registration was achieved after estimating the pose (Figure 4.4c) using only four manually selected corresponding surface landmarks. The outcome of the registration was verified visually (Figure 4.4) by projecting the kidney and

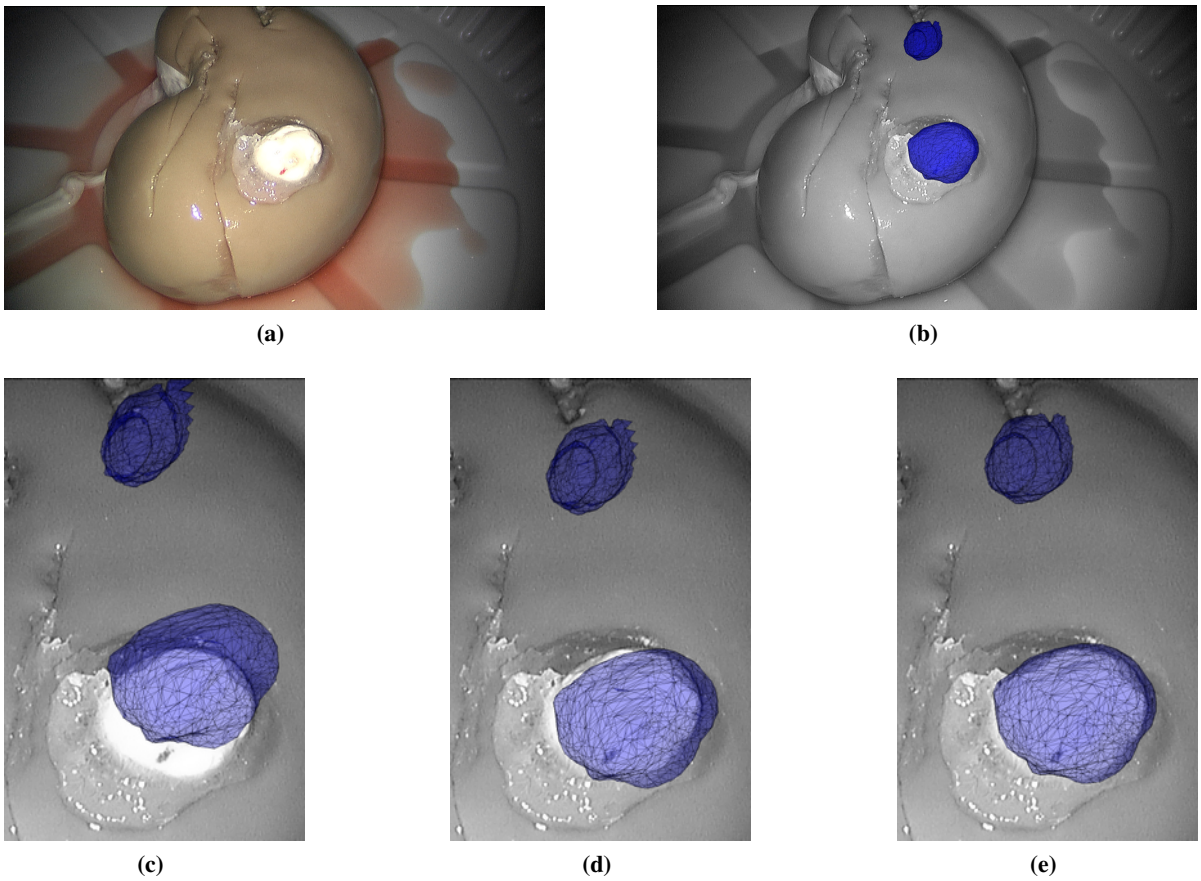


Figure 4.4: Results of registration: (a) Original left stereo camera view, (b) final registered crisp mesh of the tumors (blue) projected on top of the image. Close-up views depicting intermediate results of the registration: (c) pose estimation, (d) automatic similarity transformation, and (e) non-rigid registration.

tumor surfaces on both left and right endoscopy views. The small amount of error (< 1 mm) observed in the resulting registration is due to the error in reconstruction which is attributed to lack of texture on the phantom.

In order to verify the usefulness of probabilistic boundary visualization, we present four visualization scenarios. In the **first case** (Figure 4.4b), we generate a crisp mesh model of the tumor by thresholding the probabilistic segmented CT volume to extract the most probable kidney-tumor boundary. In our **second case**, we project the previously generated mesh onto a 2D plane (normal to the camera) and extract its contour (Figure 4.5a). This particular approach does not provide the surgeon with much additional information. Without any visible information (e.g. in the endophytic case) the surgeon's confidence regarding the visualized crisp boundary is, at best, changing isotropically away from the contour (as emulated in Figure 4.5b). **Third case**, we isotropically dilate the 3D thresholded volume of the tumors by 1mm increments and overlay the corresponding projected 2D contours (Figure 4.5c). The resulting 2D contours dilate anisotropically as they are influenced by the orientation and shape of the tumor in 3-space. **Fourth case**, we propose thresholding the probabilistic volume at

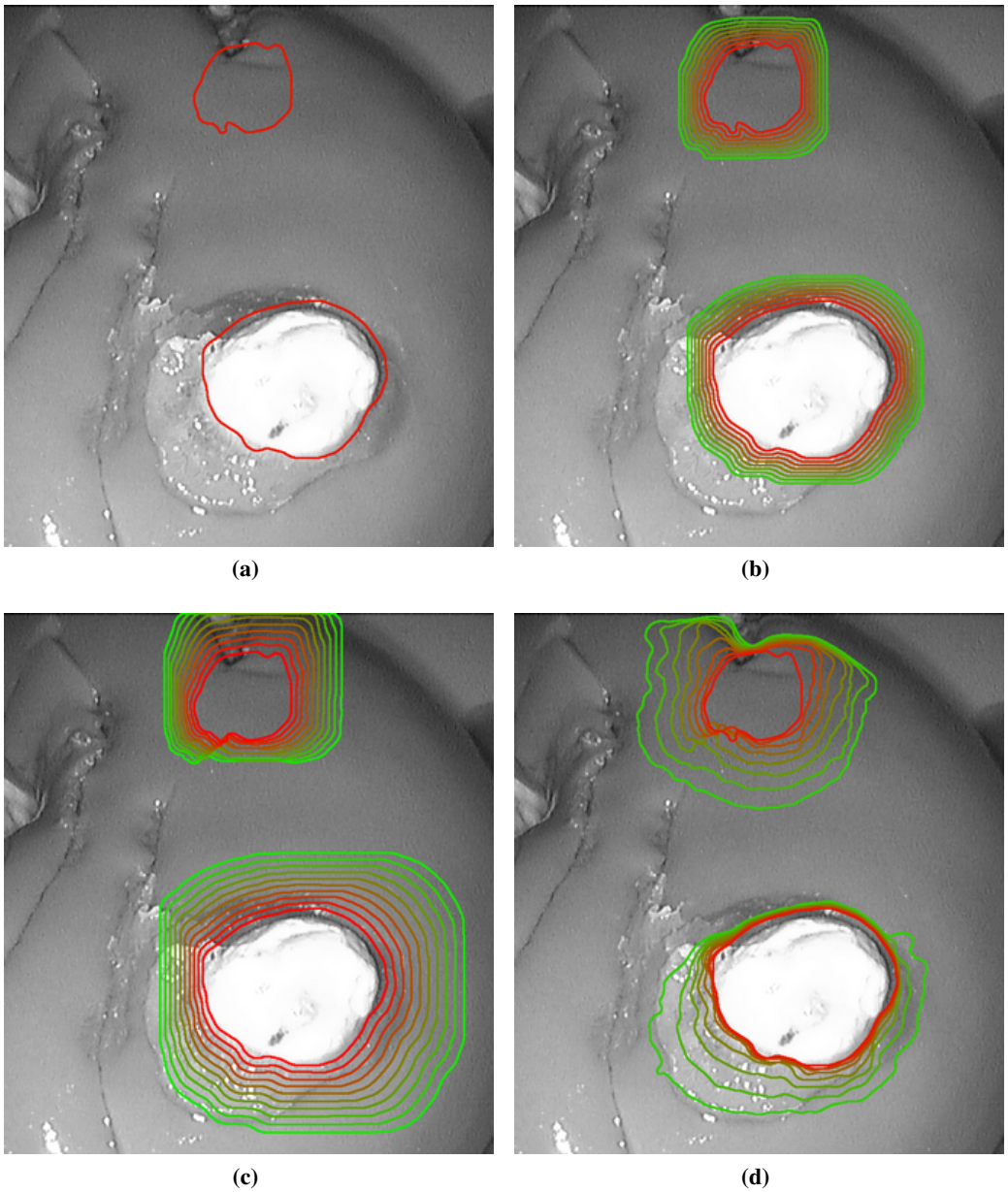
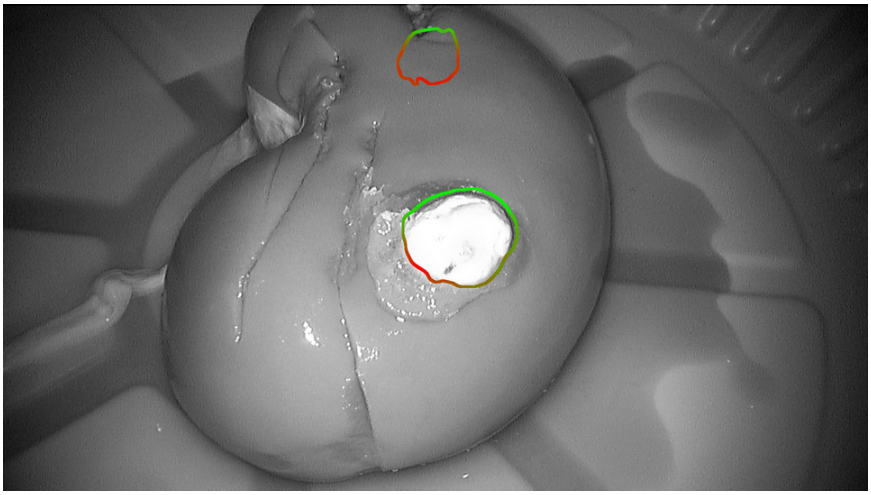
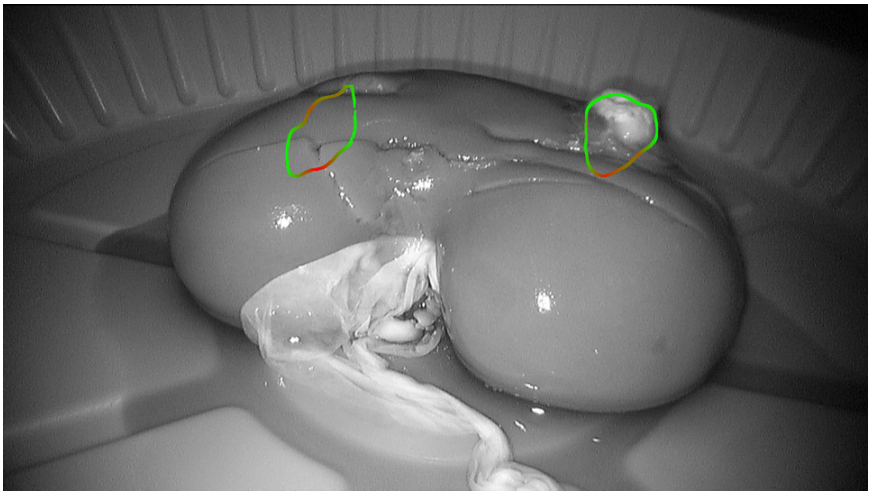


Figure 4.5: Augmented reality view of *Ex vivo* lamb kidney with endophytic and exophytic artificial tumors showing different visualization scenarios: (a) crisp contour of projected mesh, (b) isotropic 2D diffusion of the crisp contour, (c) 2D projections of the crisp mesh dilated in 3D by 1mm increments, (d) 2D projections of 3D isoprobabilities from 0.5 to 0.15. Contours range from the most probable boundary (red) to the most conservative boundary (green).

increasingly conservative confidence intervals instead of isotropic dilation to obtain isoprobability contours (Figure 4.5d). In this case, we are essentially guiding the dilation of resection boundaries using the underlying uncertainty information extracted during the probabilistic segmentation of the CT. These results are consistent with our initial observation that the diffusion of uncertainties are greater in the endophytic case (pit/gum difference).



(a)



(b)

Figure 4.6: (a) Top and (b) side views of *ex vivo* lamb kidney augmented with uncertainty-driven tumor boundary localization. Uncertainty is encoded into the tumor boundary ranging from certain (green) to uncertain (red).

We presented the four cases to expert urology surgeons. The general consensus was that the information presented in the fourth case (Figure 4.5d) is promising. A valid critique was made regarding the number of contours being overlaid on the endoscopy view: it obstructs the kidney more than the simple crisp solution (Figure 4.5a). In order to address this problem, we present a complimentary visualization scenario in which uncertainties are projected onto a single crisp contour. We accomplish this by computing the minimum distance between the most probable contour and the most conservative one at every location of the most probable contour (distance from inner-most to outer-most contours in Figure 4.5d). A lower distance implies a higher confidence in the boundary localization as it indicates a sharper edge in the probability map. We then transform these distances into a relative color map and use it to color-code the crisp contour (Figure 4.6).

This final visualization scenario does not only provide the most probable tumor boundary localization, but also provides information about its local confidence. This visualization may prove to be effective in guiding the surgeon to quickly identify the best (most confident) place to start the resection. During the resection, the surgeon can always opt for the fourth case to see exactly how the uncertainty is diffused spatially.

Although our visualization techniques were well received by our clinical collaborators, the assumed benefits of our proposed visualizations are yet to be thoroughly evaluated in a clinical context. It is important to note that the benefits of navigation uncertainty in computer-assisted orthopedic interventions have been demonstrated [159] and there exists prior art on experimental design of user studies for evaluating the benefits of uncertainty-encoded visualizations [160].

4.4 Summary

We proposed methods that enable the extraction and registration of probabilistic data from two complimentary sources of information available in robot-assisted surgical interventions. Our approach provides the confidence in the resulting augmented information, which may prove useful to the surgeon during the demarcation of excision margins before resection.

The novel visualization we presented is a proof-of-concept. The next step is to validate our experiments on clinical data and more realistic *ex vivo* phantoms with agar-based tumors of varying intensities, shapes and sizes [77]. Our methods also stand to benefit from in-depth summative usability tests in addition to more formative usability tests to fully validate the integration of our uncertainty encoded visualization techniques into the clinical workflow. To do this, we advocate for simulated RAPN tumor identification and resection studies similar to the phantom studies recently presented by Singla et al. [166]. The aim of such user study should be to evaluate the benefits of our proposed uncertainty-encoded visualization techniques during RAPN. As such, this experiment should follow the experimental design guidelines proposed by Simpson et al. [160], while adopting the success criteria proposed by Singla et al. [166], i.e., excision time, adjusted excised specimen volume, and excision margin status.

In addition to the proposed user studies, there is a need to automate the initialization (pose estimation) steps of our methods to facilitate real-time operation. This may be done using our Nosrati et al. [124] method presented in Chapter 2. It is important to point out that the utility of our proposed methods is not contingent on automatic initialization and real-time registration. In fact, our proposed methods can be integrated into other RAPN guidance techniques [35, 153, 185]. For example, our visualizations can be used to render uncertainty-encoded tumor boundaries obtained from intraoperative CT [185] or US [153] imaging. Furthermore, the uncertainties which we compute from preoperative CT may be fused with uncertainties from additional intraoperative sources (e.g., US [153], CT [185], and tracking data [35]) to improve the confidence at the localized boundary as new data is acquired during resection.

Above all, we believe that it is imperative to extend our method to account for the uncertainties associated to the error-prone image registration stage, which has been shown to be the biggest source of uncertainty in IGT navigation [159]. In the following chapter, we present our contributions towards automatic estimation, propagation, and visualization of deformable image registration uncertainties.

Chapter 5

Encoding Deformable Image Registration Uncertainties for Scene Augmentation

"Medicine is a science of uncertainty and an art of probability."
— William Osler

This chapter continues the presentation of our contributions towards addressing our second research question by investigating the effects of deformable image registration uncertainties on intraoperative guidance. To enable such capability, we propose a mathematical framework that first estimates the registration uncertainty and subsequently propagates the effects of the computed uncertainties from the registration stage through to the visualizations, organ segmentations, and dosimetric evaluations. To ensure the practicality of our proposed framework in real world image-guided radiotherapy contexts, we implemented our technique via a computationally efficient and generalizable algorithm that is compatible with existing deformable image registration software.

5.1 Computing and Propagating Uncertainties in Deformable Registration

The details of our proposed framework are organized into four components and are presented as follows. We start by detailing our registration uncertainty (RU) estimation technique in Section 5.1.1. In Section 5.1.2, we present a parametric representation of the computed RU along with two different methods for visualizing the computed RU. In Section 5.1.3, we present our uncertainty propagation method, which encodes the parametric representation of the computed RU into associated dosimetric maps and segmentation masks representing delineation of organs at risk (OAR). Finally, in Section 5.1.4, we take the analysis of RU propagation one step further by demonstrating how RU encoded dosimetric maps can be accumulated into an RU encoded accumulated dose map, which can then be summarized into a RU encoded dose-volume histogram (DVH) representation.

5.1.1 Computing Registration Uncertainty

Given a pair of volumetric images $f_t, f_m : \mathbb{R}^3 \rightarrow \mathbb{R}$ henceforth referred to as target and moving image, the objective of parametric deformable image registration (DIR) models is to obtain an optimal set of transformation parameters θ that describe a transformation function $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. The function $\mathbf{T}(\mathbf{x};\theta)$, defined over the domain of pixels $\mathbf{x} \in \Omega_t \subset \mathbb{R}^3$ of the target image, maps a deformation field to the domain of the moving image $\Omega_m \subset \mathbb{R}^3$ in order to maximize its overall correspondence or similarity with the target image, measured with the similarity metric Ψ. This optimization is written as

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\, \Psi\left(f_t(\mathbf{x}), f_m \circ \mathbf{T}(\mathbf{x};\theta)\right). \tag{5.1}$$

If the solution to Equation 5.1 is guaranteed to be correct every time (which is not the case in many real world applications), we should then be able to perturb (geometrically deform) the target volume in a controlled manner and expect the solution to Equation 5.1 to correctly identify exactly how the target image was perturbed. If the DIR process cannot correctly predict the perturbations, there must be uncertainties in the process, and we can therefore compute the difference between the controlled perturbations and the perturbations estimated with Equation 5.1 to quantify the uncertainty in registration.

The process of applying controlled perturbations and estimating the errors in DIR is computationally expensive. Rather than applying random and unrealistic perturbations to the target image, we can leverage the initial DIR solution $\hat{\theta}$ and prior knowledge about how we expect the registration to go wrong to induce these perturbations in a more informed and computationally efficient manner [199]. As such, RU may be computed by treating θ as a normally distributed random variable $\mathcal{N}(\mu_\theta, \Sigma_\theta)$, in which μ_θ is estimated by $\hat{\mu}_\theta = \hat{\theta}$ and Σ_θ is estimated depending on the deformation model used. For the popular B-spline deformation model, Σ_θ is estimated using an *a priori* baseline Σ_o in a convex combination

$$\Sigma_\theta = (1-\rho)\Sigma_o + \rho\,\hat{\theta}\hat{\theta}^{\mathsf{T}} \tag{5.2}$$

weighted by $\rho \in [0,1)$ forming a shrinkage estimator. In context of 3D B-spline DIR, this *a priori* baseline Σ_0 is computed by taking the Kronecker product of a 3D first order autoregressive model Σ_{AR} with a 3×3 matrix $\mathbf{c}$ that encodes the covariance of the x, y, z components of the transformation parameter θ, such that $\Sigma_0 = \mathbf{c} \otimes \Sigma_{AR}$.

The 3D autoregressive kernel

$$\Sigma_{AR}(i,j) = r_x^{|x(i)-x(j)|} r_y^{|y(i)-y(j)|} r_z^{|z(i)-z(j)|}, 1 \leq i, j \leq K \tag{5.3}$$

is computed subject to smoothness parameters $\mathbf{r} = (r_x, r_y, r_z)$ and constraints $|r_x|, |r_y|, |r_z| < 1$ that control the smoothness between sampled neighboring B-spline control points, where K is the total number of control points. In this formulation, $x(i) = \text{mod}(i-1, n_x), y(i) = \text{mod}(\lfloor (i-1)/n_x \rfloor, n_y), z(i) = \lfloor (i-1)/(n_x n_y) \rfloor$ represent the mapping from index i to the corresponding x, y, z coordinate of the i^{th} control point, assuming an (n_x, n_y, n_z) grid of B-spline control points. The derivation of this prior and efficient

techniques for sampling it are detailed in Watanabe and Scott [199]. The selection of the free parameters **r** and **c** indeed depends on the chosen DIR hyperparameters, e.g., number of B-spline control points, transformation regularization, optimization strategy, extent of possible deformable motion in clinical context, etc. A good strategy for selecting **r** and **c** would be based on initial clinical phantom studies, as suggested by Brock et al. [19], or from clinical experiments as explained in Section 5.3.1.

Once the distribution of θ has been estimated, the next steps for computing RU are to: (i) draw samples $\theta_i \sim \mathcal{N}(\hat{\mu}_\theta, \hat{\Sigma}_\theta)$ from the distribution of parameters, (ii) apply the sampled transformation to the target image to generate synthetic images $f^{(i)} = f_t \circ \mathbf{T}^{-1}(\mathbf{x}; \hat{\theta}) \circ \mathbf{T}(\mathbf{x}; \theta_i)$, (iii) register the moving image f_m to each and every synthetic image to obtain

$$\hat{\theta}_i = \underset{\theta}{\operatorname{argmax}}\ \Psi\left(f_t \circ \mathbf{T}^{-1}(\mathbf{x}; \hat{\theta}) \circ \mathbf{T}(\mathbf{x}; \theta_i), f_m \circ \mathbf{T}(\mathbf{x}; \theta)\right), \tag{5.4}$$

and finally (iv) compute the errors in the simulated transformations

$$\mathbf{e_i}(\mathbf{x}) = (e_{ix}(\mathbf{x}), e_{iy}(\mathbf{x}), e_{iz}(\mathbf{x})) = \mathbf{T}(\mathbf{x}; \hat{\theta}_i) - \mathbf{T}(\mathbf{x}; \theta_i). \tag{5.5}$$

These errors represent the performance of the DIR method, in the spatial domain of the target image, subject to a set of random but structured perturbations (governed by prior Σ_o) applied to the target image. These errors implicitly account for overall end-to-end performance of a given a pair of images and a DIR method with fixed hyperparameters. The overall end-to-end performance accounts for both the contextual geometric uncertainties that exist in the image pair (e.g., lack of salient image information) and the behavior of the DIR method (e.g., correctness of the deformation model and performance of the optimization strategy used) subject to the contextual geometric uncertainties. The spatial distribution of these errors are therefore the RU associated to the DIR process. The estimation of RU is improved by drawing more samples from θ_i, however this process is expensive as computing an additional $\mathbf{e_i}$ requires a computationally expensive 3D registration. In the next section, we describe how this expensive process can be simplified by assuming a parametric model over the distribution of the RU computed for DIR parameter θ, and by extension, at every voxel **x**.

5.1.2 Parametric Representation and Visualization of Registration Uncertainty

The benefits of assuming a parametric model for the distribution of RU are three-fold. This parametric representation reduces memory requirements for storing $\{\mathbf{e_i}(\mathbf{x})\}$, reduces the number of $\mathbf{e_i}$ samples required to estimate RU, and simplifies the process of propagating the RU onto the next processing steps, i.e., using the spatial confidence to diffuse segmentations or dose maps. Assuming that $\mathbf{e}(\mathbf{x}) \sim \mathcal{N}(\mu_e(\mathbf{x}), \Sigma_e(\mathbf{x}))$, spatial confidence region $\Phi(\mathbf{x})$ is defined by

$$\Phi(\mathbf{x}) = \{\mathbf{x}' : (\mathbf{x} - \mu_e(\mathbf{x}))^\intercal \Sigma_\mathbf{e}^{-1} (\mathbf{x} - \mu_e(\mathbf{x})) < \chi_3^2(1-\gamma)\}, \tag{5.6}$$

where γ is the confidence level and $\{\mu_e(\mathbf{x}), \Sigma_e(\mathbf{x})\}$ are the sample mean and covariance of $\{\mathbf{e_i}(\mathbf{x})\}$. Sample covariance calculation is sensitive to outliers and becomes a problem if the synthetic samples

are few. Student's t-distribution is a better fit but it is computationally complex as its parameters have to be solved numerically. We propose a two-step estimation of the sample covariance by omitting the outliers that fall outside of the $\gamma = 0.95$ level set of the first estimate of $\Phi(\mathbf{x})$. It should be noted that even if θ is sampled normally, there is no guarantee that the errors $\{\mathbf{e_i}(\mathbf{x})\}$ will follow a normal distribution. This matter is further discussed in the Results section.

The confidence regions $\Phi(x)$ are effectively structural tensors parameterized by $\{\mu_e(\mathbf{x}), \Sigma_e(\mathbf{x})\}$ at every voxel $\mathbf{x}$. To visualize the quality of DIR at each voxel, these structural tensors may be rendered as ellipsoids in 3D using tensor rendering toolkits available for medical images. An alternative visualization may also be produced by transforming these structural tensors to scalar quantities that represent RU by computing the differential entropy of each tensor

$$E(\mathbf{x}) = \frac{3}{2}\left(1 + \ln(2\pi)\right) + \frac{1}{2}\ln|\Sigma_e(\mathbf{x})|. \tag{5.7}$$

Note that differential entropy E is a logarithmic measure, which depends on the log-determinant of the covariance matrix $\Sigma_e(\mathbf{x})$ and as a result $E \in (-\infty, \infty)$.

In the next section, we show how this parametric representation is used within our framework to propagate RU beyond the registration step.

5.1.3 Diffusion of Image Information using Registration Uncertainty

The parametric representation of RU, in essence, presents the likelihood of possible locations where a voxel from the domain of the moving image would be mapped to in the domain of the target image. As such, RU also presents the likelihood that a voxel in the target image domain may have the values of its neighbors due to potential errors in registration. We can propagate the effects of potential DIR errors onto associated mappings, e.g., segmentation labels and dose maps, using a weighted averaging scheme that encodes the likely contributions of neighboring voxels.

Let $\mathbf{I}_m = (D_m, S_m^1, S_m^2, ..., S_m^N) : \mathbb{R}^3 \rightarrow \mathbb{R}^{N+1}$ represent a vector valued volume that contains N binary segmentation labels $S_m \in \{0, 1\}$ of different OAR and radiation dose map D_m in the domain of the moving image f_m. This image can be transformed into the domain of the target image using the parameters $\hat{\theta}$ such that $\mathbf{I}'_m(\mathbf{x}) = \mathbf{I}_m \circ \mathbf{T}(\mathbf{x}; \hat{\theta}), \mathbf{x} \in \Omega_t$. To compute the posterior distribution $\mathbf{P}'_m \sim \mathcal{N}(\mu_P, {\sigma_P}^2)$ of $\mathbf{I}'_m$, RU is encoded using a convex combination characterized by the weighted mean and variance

$$\begin{aligned} \mu_P(\mathbf{x}) &= \frac{\sum_{\forall \mathbf{x}'} \mathbf{W}(\mathbf{x}|\mathbf{x}')\mathbf{I}'_m(\mathbf{x}')}{\sum_{\forall \mathbf{x}'} \mathbf{W}(\mathbf{x}|\mathbf{x}')}, \\ \sigma_P(\mathbf{x})^2 &= \frac{\sum_{\forall \mathbf{x}'} \mathbf{W}(\mathbf{x}|\mathbf{x}')(\mathbf{I}'_m(\mathbf{x}') - \mu_P(\mathbf{x}))^2}{\sum_{\forall \mathbf{x}'} \mathbf{W}(\mathbf{x}|\mathbf{x}')}, \end{aligned} \tag{5.8}$$

where the weights

$$\mathbf{W}(\mathbf{x}|\mathbf{x}') = (2\pi)^{-\frac{3}{2}}|\Sigma_e(\mathbf{x}')|^{-\frac{1}{2}} e^{\left(-\frac{1}{2}(\mathbf{x}-\mu_e(\mathbf{x}'))^\intercal \Sigma_e(\mathbf{x}')^{-1}(\mathbf{x}-\mu_e(\mathbf{x}'))\right)} \tag{5.9}$$

are obtained by evaluating the probability density of the trivariate normal distribution $\mathcal{N}(\mu_e(\mathbf{x}'), \Sigma_e(\mathbf{x}'))$

at $\mathbf{x}$. The computational performance of this diffusion process can be increased by assuming that the contribution of $\mathbf{I}'_m(\mathbf{x}')$ to $\mathbf{P}_m(\mathbf{x})$ is negligible, i.e. $\mathbf{W}(\mathbf{x}|\mathbf{x}') \sim 0$, when $\mathbf{x}$ and $\mathbf{x}'$ are more than a certain distance apart; effectively reducing the space of $\mathbf{x}' \in \Omega_t$, over which Equation 5.8 is evaluated.

It is worth noting that a possible simplification of Equation 5.8, which we do not perform, is to assume that $\mathbf{W}(\mathbf{x}|\mathbf{x}') = \mathbf{W}(\mathbf{x}'|\mathbf{x})$, then the computation of the mean simplifies to $\mu_P = \mathbf{W} * \mathbf{I}'_m$. This allows for the mean of the posterior to be calculated by convolution or, in other words, by blurring $\mathbf{I}'_m$ with an adaptive Gaussian kernel parameterized by $\{\mu_e(\mathbf{x}), \Sigma_e(\mathbf{x})\}$. This approach was used in Simpson et al. [163] to diffuse segmentation labels in a study involving registration of inter-subject brain MRI scans.

Having computed a parametric distribution of associated dose maps $\{\mu_{D_m}, \sigma_{D_m}\}$ as part of $\mathbf{P}'_m$, we can now extend the propagation of RU though to the accumulated dose map and resulting DVH representation, which is detailed in the following section.

5.1.4 Propagation of Registration Uncertainty in Accumulated Dose Analysis

DVH parameters have been used in numerous published toxicity analyses, which are the basis for recommended dose constraints used in modern-day treatment planning. It is important to compute and visualize the effects of RU on the DVH curves as the RU information may be used to (i) simulate potential errors in how the DVH was computed, and to (ii) provide uncertainty metrics for the DVH-derived parameters that are used in practice for evaluating the treatment plans. We propose to use the propagation technique presented in Section 5.1.3 to compute the distribution of accumulated dose at each pixel after dose information from different fractions have been probabilistically aligned onto the target (reference) image volume. An RU encoded probabilistic DVH curve is then computed from the generated probabilistic accumulated dose map.

The DVH for each $n \in N$ OAR, denoted $H_n : d \in \mathbb{R} \rightarrow \mathbb{R}$, is defined as the cumulative distribution function of the number of pixels within an OAR that receive an accumulated radiation dose above a certain accumulated dose value d. Let $\mathbf{x}_n \subset \mathbf{x}$ define the subset of pixels that fall within a certain OAR indexed as n, as indicated by the corresponding segmentation label, such that $S^n(\mathbf{x}_n) = 1, \forall \mathbf{x}_n$. The DVH is formally defined as

$$H^n(d) = \sum_{\forall \mathbf{x}_i \in \mathbf{x}_n} \mathbb{I}(D_{acc}(\mathbf{x}_i) > d), \tag{5.10}$$

where $\mathbb{I}(.)$ is the indicator function. $D_{acc}(\mathbf{x})$ represents dose map containing the summation of the total amount of radiation received throughout all fractions accumulated onto one fraction (selected as the target volume) such that

$$D_{acc}(\mathbf{x}) = D_t(\mathbf{x}) + \sum_{\forall m} D_m \circ \mathbf{T}(\mathbf{x}; \hat{\theta}_m), \tag{5.11}$$

where D_t is the dose map of the fraction selected as the target and D_m are the dose maps of all $m \in M$ moving fractions. The DIR alignment of each moving fraction onto the target fraction is parameterized by $\hat{\theta}_m$ obtained by registering each f_m to f_t.

Having defined how the DVH is computed using Equation 5.10, we can now extend the formulation

to account for RU by leveraging the parametric representation of probabilistic dose maps $\{\mu_{D_m}, \sigma_{D_m}\}$ computed in the previous section. Assuming that $\{\mu_{D_m}, \sigma_{D_m}\}$ computed for each moving fraction $m \in M$ are independent random variables, the sum of them is also normal, with its mean being the sum of the means, and its variance being the sum of the variances. Summing the aligned RU encoded dose volumes results in the parametric representation of the accumulated dose map

$$\begin{aligned} \mu_{D_{acc}}(\mathbf{x}) &= D_t(\mathbf{x}) + \sum_{\forall m} \mu_{D_m}, \\ \sigma_{D_{acc}}(\mathbf{x})^2 &= \sum_{\forall m} \sigma_{D_m}^2. \end{aligned} \tag{5.12}$$

The DVH may now be computed at different α standard deviations from the mean accumulated dose value such that

$$H^n(d;\alpha) = \sum_{\forall \mathbf{x}_i \in \mathbf{x}_n} \mathbb{I}\left((\mu_{D_{acc}}(\mathbf{x}_i) + \alpha\sigma_{D_{acc}}(\mathbf{x}_i)) > d\right). \tag{5.13}$$

In summary of the methodologies presented in this section, we have proposed a framework that is capable of: (i) estimating the quality of a B-spline DIR method via RU, (ii) visualizing the RU, and (iii) propagating the effects of RU onto associated segmentation labels and dose maps using a weighted averaging technique. We then went a step further to propagate the effects of RU on the accumulated dose from multiple fractions. The corresponding DVH representation, which may be used for interventional planning and quality assessment purposes, was also computed. We discuss the potential benefits of RU encoded DVH analysis in further detail in Section 5.3.4. In the following section, we demonstrate the utility of our framework on a retrospective study of fractionated brachytherapy interventions.

5.2 Encoding Uncertainties in Multi-Fraction Cervical Cancer Brachytherapy

Our application choice for demonstrating the effects of RU on DIR and therapy delivery is image-guided multi-fraction cervical cancer brachytherapy (MFCCB). MFCCB is a form of radiotherapy that relies heavily on 3D imaging for treatment planning, delivery, and quality control. A patient undergoing MFCCB is exposed to multiple sessions (fractions) of radiation separated by one or more days to maximize the effectiveness of radiation on cancerous cells and to minimize the damage to healthy cells. In MFCCB, an updated set of 3D patient-specific planning images is acquired and segmented prior to the delivery of radiation at each fraction to identify changes in the clinical target volume and in the OAR.

Similar to other image-guided radiotherapy applications, many of the challenges faced in MFCCB are rooted in the fact that hollow elastic OAR, such as the urinary bladder, may have a significantly different size, shape and orientation across treatment fractions mainly due to differences in fill volumes and brachytherapy applicator positioning [70]. In other similar applications, the use of RU quantification in DIR-based assessment has been presented as a potential solution to the following challenges: (i) computing the accumulation of radiation doses across multiple fractions for quality control and post-treatment analysis [119, 148, 186]; (ii) transferring the delineation of organs from one fraction to the

next to expedite the planning stage [48, 69, 79, 163]; and finally (iii) visualizing localized RU to assist clinicians when decisions are made based on registered image data [147].

The application of MFCCB also stands to benefit from multimodal registration methods, such as our proposed techniques, even though DIR is being performed on different CT scans of the same patient. We believe that this is due to occasional observable differences in how the contrast agent presents within the bladder across different fractions. Although the protocol dictates that the amount of contrast agent must not be variable among different fractions of the same patient, noticeable variations occur due to both human error, which results in possible variations in the water-contrast mixture used for each fraction, as well as visible variation of contrast within the same image due to settling of the contrast material inside the bladder [88]. This settling effect is further complicated by the fact that the bladder fill level and balloon catheter position are intentionally altered across fractions in some treatment centers to reduce the risk of urinary toxicity [209]. A multimodal RU estimation method may therefore be better suited to MFCCB and other similar interventional applications.

In this clinical context, we present evidence that our proposed framework is beneficial to the three challenges associated with image guided MFCCB. Specifically, we show that (i) the effect of RU on dose accumulation provide useful insights for quality control and post-treatment analysis; (ii) RU propagation improves the transfer of delineations from one fraction to the next; and (iii) RU can be used to generate intuitive visualizations that reflect the quality of DIR and thus can assist physicians in making decisions based on registered image data. We emphasize that our proposed methodology, though presented here to address the challenges of MFCCB, can be generalized to other similar clinical application as it is applicable to different DIR software, is faster than current Bayesian approaches, and has been designed to handle multimodal registration tasks.

5.3 Experiments

5.3.1 Data and Experimental Setup

We applied the methods outlined in the previous section to a dataset of 37 MFCCB patients collected at BC Cancer Agency, Vancouver, Canada, following applicable Research Ethics Board review. Each of the 37 instances within the dataset is comprised of five fractions and the set of planning images acquired at each fraction contains a computed tomography (CT) scan f, detailed segmentations of the interior S^i and exterior S^e surfaces of the bladder wall, and a dosimetric map D representing the amount of radiation absorbed at every voxel. The segmentation and dose maps in each instance are in the same coordinate space as the CT volume. Also, three landmarks on the urinary bladder that make up the Trigone (bladder neck and ureteral orifices) were identified in all CT scans to enable the calculation of target registration error (TRE). Landmarks were manually localized with respect to the origin of the corresponding CT scan. Landmark locations were not used to facilitate the alignment of the image volumes and no further processing was done to the image volumes prior to registration.

The clinical goal in MFCCB analysis is to account for changes in patient anatomy between fractions, thus we need to perform intrapatient registration. The CT volume of the first fraction is selected as the

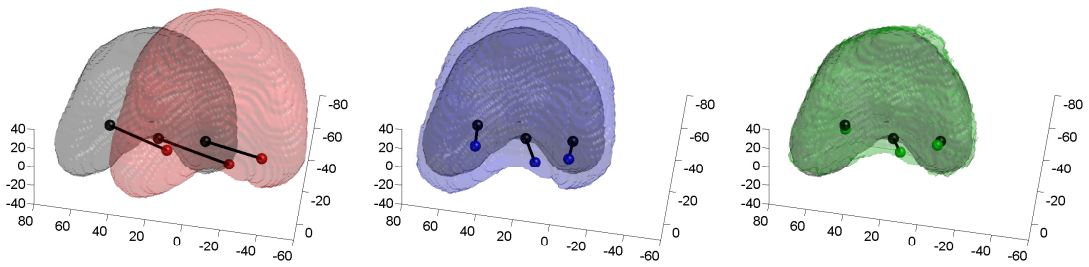


Figure 5.1: Exemplar case depicting the effects of registration on bladder and landmark alignment. 3D contours of the bladder presented in this figure are extracted from the target image volume (black) and the moving image at different stages of alignment: before registration (red), after affine registration (blue), and after affine+deformable image registration (green). Landmark locations (two ureter orifices and the bladder neck) are plotted as spheres and are colored according to their corresponding 3D bladder contour. Correspondence between landmark locations are indicated with a black line. In this example, the mean landmark alignment error is reduced from 36.4 mm before registration to a mean target registration error of 12.9 mm after affine and 5.4 mm after affine+deformable image registration.

target volume f_t to which the CT volumes of all other fractions are registered. Automatic registration is performed in two stages: (i) a simple affine image registration to roughly align the volumes followed by (ii) a DIR to account for elastic deformations. The `elastix` toolbox [91], which is based on the popular Insight Segmentation and Registration Toolkit (ITK, Kitware Inc.), was used to perform the registrations and the same pair of affine and DIR parameter files were used for all registrations. To speed up computation, each image volume was downsampled by a factor of two; to 256×256 axial resolution.

The error in landmark alignment is computed before and after every stage of registration by computing the Euclidean distance between the Trigone landmarks of the target and moving image volumes (Figure 5.1). The landmark alignment error is averaged across the four fractionated moving volumes of each patient and is presented in Figure 5.3. The automatic affine image registration step of our framework reduced the mean landmark alignment errors across four moving fractions and 37 patients from 29.57 mm to a mean TRE of 10.67 mm. This mean TRE is further reduced from 10.67 mm to 7.71 mm following the initial DIR. All resulting reductions in landmark alignment errors following each registration step were found to be statistically significant subject to a Wilcoxon signed-rank test, $p < 0.001$. The independent and joint histograms of the displacements between corresponding landmarks are presented in Figure 5.2.

RU is computed only for the DIR step of the registration. To compute the RU after the initial DIR, our method requires an *a priori* baseline Σ_0 in order to realize 40 samples of the transformation parameters. We parametrize the free parameters of the first order 3D autoregressive model Σ_{AR} by setting the value of $\mathbf{c}$ to the covariance of absolute x, y, z error components of the TRE computed post-DIR. We set the smoothness parameter $\mathbf{r}$ to 0.9 for all x, y, z directions, and ρ to 0.1 as recommended in the original paper [199].

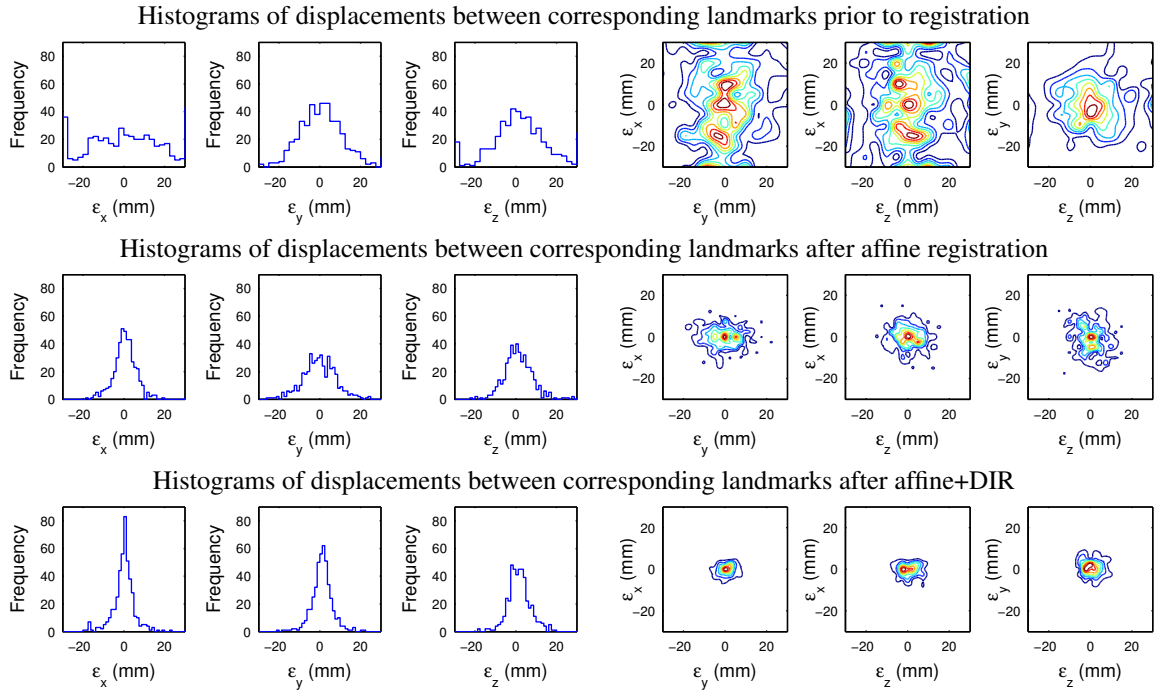


Figure 5.2: Independent and joint histograms of the x, y, z components of displacements $\mathscr{E}$ between corresponding landmarks. Displacements are computed for all three corresponding pairs of landmarks (two ureter orifices and the bladder neck) with respect to their locations in the target image volume and the moving image volume before and after affine and affine+deformable image registration (DIR) stage of image registration.

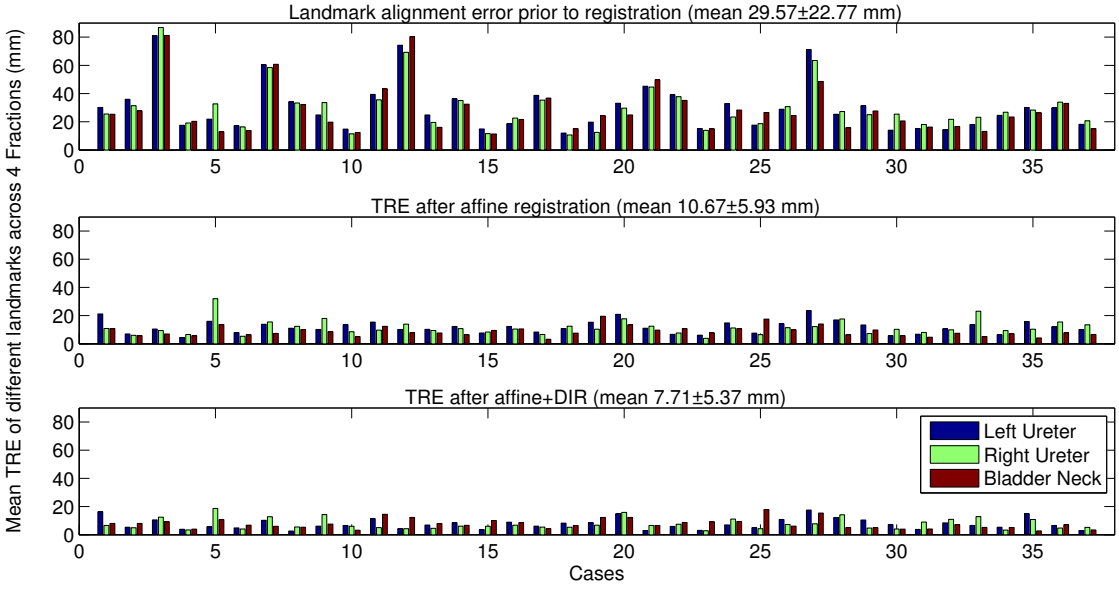


Figure 5.3: Target registration errors (TREs) computed with respect to the target image volume, averaged across the four moving volumes, acquired before and after every stage of image registration. Known landmark locations were not used during registration.

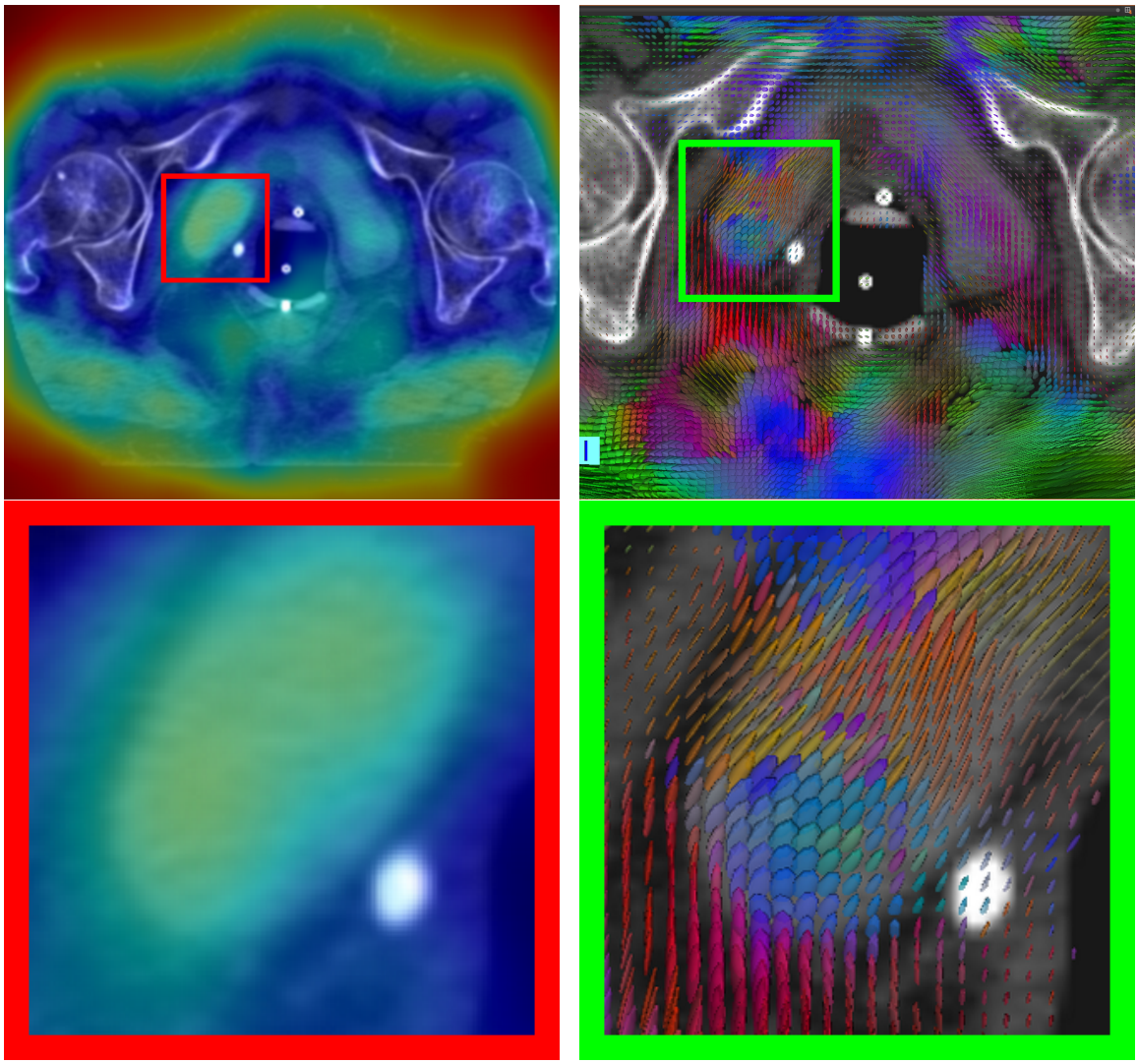


Figure 5.4: Scalar (left) and tensor (right) representations of registration uncertainty overlaid on CT. In the scalar visualization, the colors of the overlay represent uncertainty ranging from most certain (blue) to most uncertain (red). This uncertainty corresponds to the volume of the ellipsoids depicted in the tensor visualization. In the tensor representation, the colors indicate the orientation of the major axis of the ellipsoid or direction of maximum uncertainty. Note that this exemplar registration result is more certain at rigid structures (bones) compared to deformable ones (bladder, bowels, and air) and the tensor orientations follow the edge information in the image.

To increase the computational performance of diffusion process, we restricted the space over which Equation 5.8 is evaluated to an $80 \times 80 \times 80$ mm neighborhood about $\mathbf{x}$. Effectively, we assume that $\mathbf{W}(\mathbf{x}|\mathbf{x}') \sim 0$, when the distance between $\mathbf{x}$ and $\mathbf{x}'$ is more than 40 mm, i.e., twice the maximum measured TRE post-DIR.

5.3.2 Registration Results and Uncertainty-Encoded Visualizations

Applying the transformations obtained with the initial DIR to align the segmentations of the entire bladder, up to the external bladder wall $S_m^{e\prime}$, with the target volume results in an average Dice similarity coefficient (DSC) of 0.860 with a standard deviation of ± 0.092. This high similarity and low TRE (presented in Figure 5.3) indicates that the initial DIR provides a good prediction of the location of the bladder.

In the methodology, we presented a parametric representation for RU $\Phi(\mathbf{x})$, which can be interpreted as structure tensors. After computing RU, these tensors can be rendered as 3D ellipsoids using existing tensor rendering toolkits available for medical images [188] or, alternatively, transformed into scalar quantities using Equation 5.7. An exemplar visualization of the computed RU as scalar and tensor overlays is provided in Figure 5.4. To generate the color mapping for the scalar visualizations, we first normalized the computed differential entropy values to $E_N \in [0,1]$ such that $E_N = (E - \min(E))/(\max(E) - \min(E))$ and mapped color values to the normalized E_N values to represent relative uncertainty in DIR. In Figure 5.4, $\min(E) = -1.64$ and $\max(E) = 7.91$. Qualitatively, the proposed visualizations of RU are as expected; RU is more certain at rigid structures (bones) compared to deformable ones (bladder, bowels, and air) and the tensors follow the edge information in the image. Such visualizations may be useful in many similar image guided applications but its utility to clinicians needs to be thoroughly validated.

5.3.3 Propagation of Registration Uncertainties onto Segmentation Labels

We applied the proposed RU estimation method (presented in Section 5.1.1) and the proposed RU propagation method (presented in Section 5.1.3) to segmentation labels of the entire bladder to test if segmentation may be improved by encoding RU information. Specifically, we tested the following hypotheses:

1. The diffusion of segmentation labels using our proposed RU estimation method improves binary segmentation performance after non-rigid registration.

2. The diffusion of segmentation labels using the Watanabe and Scott [199] RU estimation method improves binary segmentation performance after non-rigid registration.

3. Our proposed RU estimation method is better than Watanabe and Scott [199] for improving binary segmentation labels.

4. The probabilistic segmentation labels produced using our proposed RU estimation method is better than a trivial isotropic diffusion of binary segmentation.

5. The probabilistic segmentation labels produced using the Watanabe and Scott [199] RU estimation method is better than a trivial isotropic diffusion of binary segmentation labels.

6. Our proposed RU estimation method is better than Watanabe and Scott [199] for generating probabilistic segmentation labels.

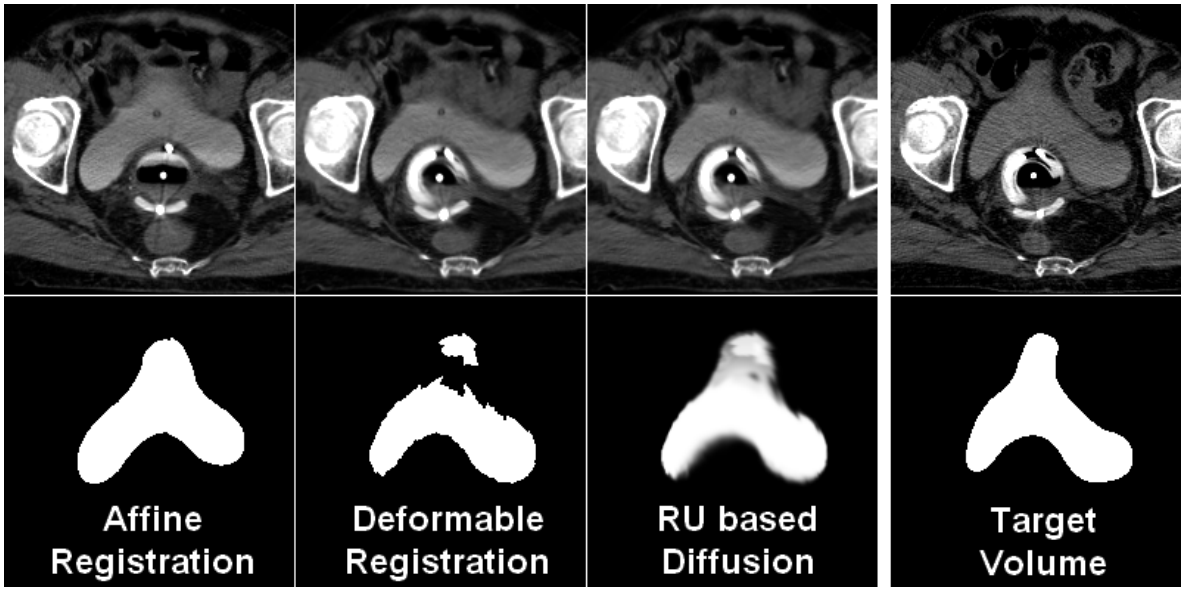


Figure 5.5: Qualitative effects of DIR and RU based diffusion on segmentation labels of the entire bladder. Exemplar slice of corresponding moving and target CT volumes (top) and segmented bladder (bottom) depicting the progressive stages of the intra-patient registration of moving to target image.

To test the first three hypotheses, we encode RU by diffusing the aligned binary segmentation labels $S_m^{e\prime}$ using Equation 5.8 and observed improvements in the alignment of the bladder labels following the diffusion (Figure 5.5). The resulting mean of the diffused labels μ_{S^e} are fuzzy and in order to quantify the improvement over the initial crisp alignment, we first threshold μ_{S^e} at 0.5 and compute the DSC. In this scenario, the average DSC increases to 0.862 ± 0.093. The improvement may seem marginal but this is due to the fact that precious information is discarded during thresholding.

The fuzziness of μ_{S^e} itself conveys useful information as well. The diffusion of μ_{S^e} is based on local RU; if μ_{S^e} is thresholded at values other than 0.5, it follows a trade-off between true positive and true negative rates. This behavior is evident in the receiver operating characteristics (ROC) space [39] shown in Figure 5.6. In the ROC space, the correctness of the alignment can be measured by the area under the receiver operating characteristics curve (AUC). To test the last three hypotheses, we interpreted the crisp alignment results as a discrete pixel classifier and μ_{S^e} as a probabilistic one, and compared their performance in classifying the ground truth bladder pixels in the target image. With the μ_{S^e} classifier, the average AUC (across 4 fractions and 37 patients) was increased from 0.942 to 0.991. Note that these reported numbers are high due to the fact that the relative number of true negative responses outnumber the misclassified ones in the CT volume. By comparison, the Watanabe and Scott [199] method for computing RU, resulted in an average AUC of 0.987. Furthermore, to demonstrate that diffusion based on our RU estimation is more representative of true registration error than an isotropic diffusion, and that the improvements in segmentation are not due to under- or over-segmentation, rather than using an adaptive RU diffusing kernel, we eroded and dilated the crisp segmentations with a 3x3x3-voxel structuring element and observed an average AUC of 0.984.

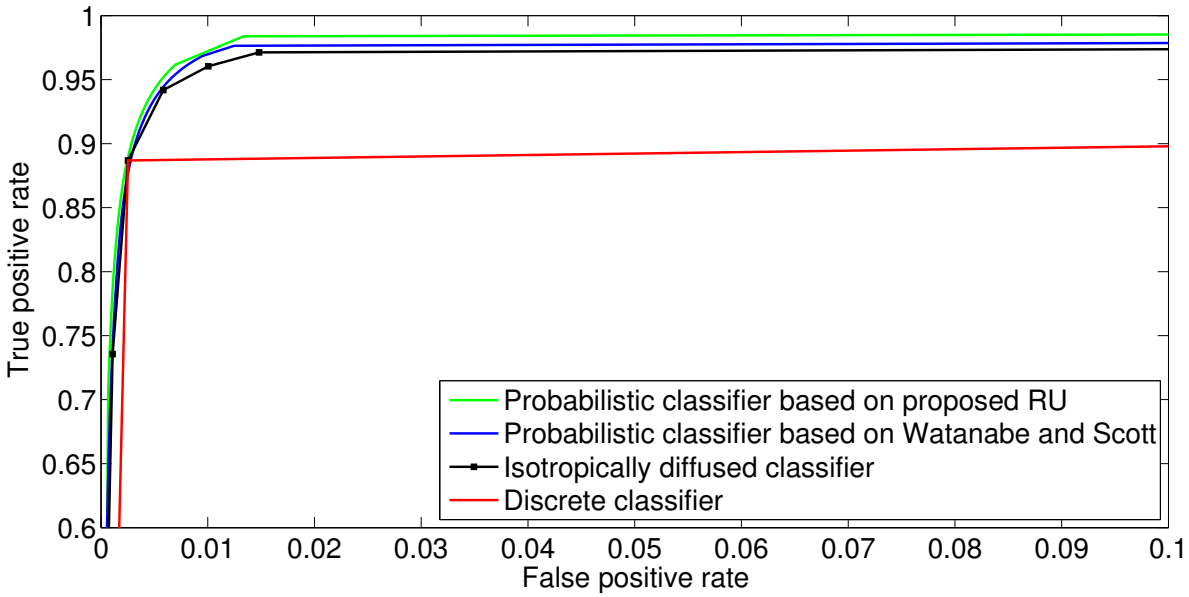


Figure 5.6: Quantitative effects of deformable image registration and registration uncertainty (RU) based diffusion on segmentation labels of the entire bladder summarized using mean receiver operating characteristics (ROC) curves. Presented ROC curves compare the mean performance of bladder alignment using different RU diffusion methods. The results of a discrete classifier (red) and an isotropically diffused version of it (black) are used as baselines for comparison. The isotropic diffusion was generated by performing 3D morphological dilation and erosion operations on the discrete segmentation. Our diffusion technique based on the proposed RU method (green) outperforms diffusion using Watanabe and Scott [199] (blue) and the baselines. The associated area under the mean ROC curves are 0.944 (red), 0.984 (black), 0.987 (blue), and 0.991 (green).

A summary of the quantitative improvements in segmentation accuracy are provided in Figure 5.7 and Table 5.1. To test our hypotheses, we tested the statistical significance of the reported differences in DSC and AUC comparisons using the Wilcoxon signed-rank test with Bonferroni adjusted alpha levels of 0.0017 per test $(0.01/6)$. We observed that the improvements using our method over both the isotropic diffusion and Watanabe and Scott [199] were tested to be statistically significant, $p < 0.001$. On the other hand, the improvement in AUC between the Watanabe and Scott [199] method and isotropic diffusion were not statistically significant, $p = 0.69$. This insignificant improvement over an isotropic diffusion implies that the standard unimodal Watanabe and Scott [199] method may not be suitable for RU estimation in context of MFCCB, which is further corroborated by the associated drop in mean DSC when the Watanabe and Scott [199] method is used. All differences in DSC were tested to be significant $p < 0.001$. Note that the crisp and isotropic diffusion approach are identical in terms of DSC performance as isotropically diffused segmentations were created from the crisp segmentations. From these observations we conclude that, as expected, our proposed multimodal RU estimation approach can better estimate the DIR errors in our context as it is capable of improving segmentation performance through the propagation of RU.

Regarding the outliers present in the boxplots in Figure 5.7, we attribute this in part to the fact that

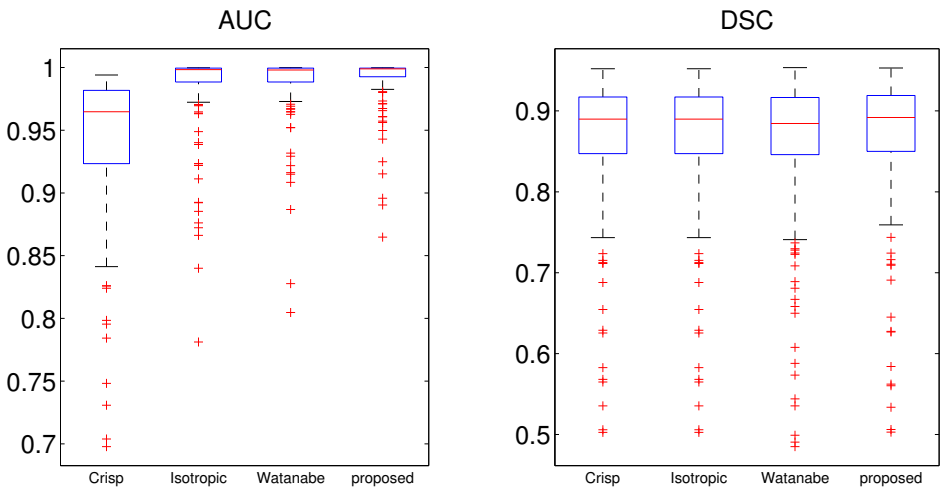


Figure 5.7: Boxplots summarizing labelling performance following the proposed and Watanabe and Scott [199] registration uncertainty based diffusion of registered segmentation labels compared to the crisp and isotropically diffused baselines. Performance results reflect performance across all 148 registrations and are measured using the Dice similarity coefficient (DSC) and area under the receiver operating characteristics curve (AUC). Note: DSC for crisp and isotropically diffused baselines are identical as the isotropic diffusion is derived from the crisp results.

the resulting DSC and AUC performance metrics are not normally distributed, and in part to our choice to use the first fraction acquired for each patient as the reference image. For two patients (cases 5 and 20), the bladder fill level in the first fraction was drastically different than the following four factions, causing the registrations be less successful compared to other 35 cases. This is also observed via the relatively higher TRE measured after affine+DIR stage, as reported in Figure 5.3; though the TRE of the Trigone landmarks may not necessarily predict DSC and AUC performance as the landmark locations do not cover the entire surface of the bladder. It is important to note the increased AUC performance on these outliers using our methods. This suggests that the fuzzy labels generated with our proposed method is effective in situations were the errors in DIR are high and, by extension, that our estimated RU is perhaps more representative of DIR errors.

A critically important, yet subtle, point that merits repeating and further discussion at this point is in regards to the direction of the RU diffusion. This behavior can only be seen in the tensor (vector) visualization presented in Figure 5.4. From this figure, one can easily observe that the diffusion of RU is greater in homogeneous regions (e.g., inside of the bladder) compared to regions with salient image texture such as observable tissue boundaries. More importantly, it can be seen that the diffusion is greater along the direction of a tissue boundary than across it. In other words, the major axis of the ellipsoidal tensors are aligned with tissue boundaries, e.g., the bladder wall. This is logical as the DIR algorithm can align two edges with relative ease, however perfect correspondence between two points along these edges is so far impossible to obtain, without the aid of perfectly accurate tissue deformation

Method	AUC	DSC
Crisp baseline	0.942 ± 0.060	0.860 ± 0.092
Isotropically diffused	$0.984 \pm 0.034^{\dagger}$	0.860 ± 0.092
Watanabe and Scott [199]	$0.987 \pm 0.028^{\dagger}$	0.855 ± 0.100
Proposed method	0.991 ± 0.021	0.862 ± 0.092

Table 5.1: Summary of the labelling performance across all 148 registrations measured using the Dice similarity coefficient (DSC) and the area under the receiver operating characteristics curve (AUC). Note that the DSC results for crisp and isotropically diffused baselines are identical as the isotropic diffusion is derived from the crisp results. †: The differences between the computed AUC for the isotropically diffused and Watanabe and Scott [199] approach were not found to be significant subject to a Wilcoxon signed-rank test, $p = 0.69$.

models, due to the lack of salient image information along the edge. Compared to the abdominal cavity and surrounding soft tissues, the boundaries of the bladder wall are easy to identify in most of the images within our dataset due to the use of contrast agents. It is thus of no surprise that the proposed RU diffusion results in a seemingly incremental improvement in segmentation alignment. By contrast the effects of RU on dose accumulation analysis, presented in the following section, is more noteworthy.

5.3.4 Effects of Registration Uncertainty on Accumulation of Dose Volumes

The aligned dose volumes D' acquired during the planning stages are also diffused using Equation 5.8. The resulting uncertainty in accumulated diffused dose maps are first computed using Equation 5.12 and then used to compute the DVH with Equation 5.13. A visual comparison between the crisp and probabilistic accumulated dose maps over the bladder walls (intersection of the interior and exterior bladder wall segmentation labels $S_t^i \cap S_t^e$) are provided in Figure 5.8.

The influence of RU on the resulting DVH is presented in Figure 5.8. In this exemplar case, the mean of the RU encoded accumulated-dose map (Figure 5.8b) is slightly higher than that of the crisp DIR aligned dose maps (Figure 5.8a), especially at the radiation hot spot that is formed on the bladder wall. This effect can also be observed from the small discrepancy between the blue and dotted black line in the DVH curves of Figure 5.8. Considering a widely used clinical dose metric such as D_{2cc} (the minimum dose received by the most exposed 2 cm^3 of tissue), which characterizes the intensity of hot spots that have the potential to cause significant morbidity [133], the DVH curves for this example yield a D_{2cc} value of 28.5 Gy using the crisp approach (dashed black line), whereas the probabilistic method (blue line) computes a more conservative value of 29.2 Gy. The mean of the RU encoded dose map is more conservative (in terms of total absorbed radiation) than that obtained from a crisp DIR alignment and, building on our conclusions from the experiments with RU encoded segmentation label propagation in Section 3.3, we hypothesize that our RU encoded accumulated-dose map and corresponding DVH curves are more representative of the true accumulated dose compared to the naïve crisp DIR approach.

Furthermore, it is evident that there exists a large uncertainty in the accumulated-dose DVH curves. More importantly, this uncertainty is greatest at high accumulated dose thresholds that constitute a smaller percentage of the total volume of the bladder. In the above example, the probabilistic DVH

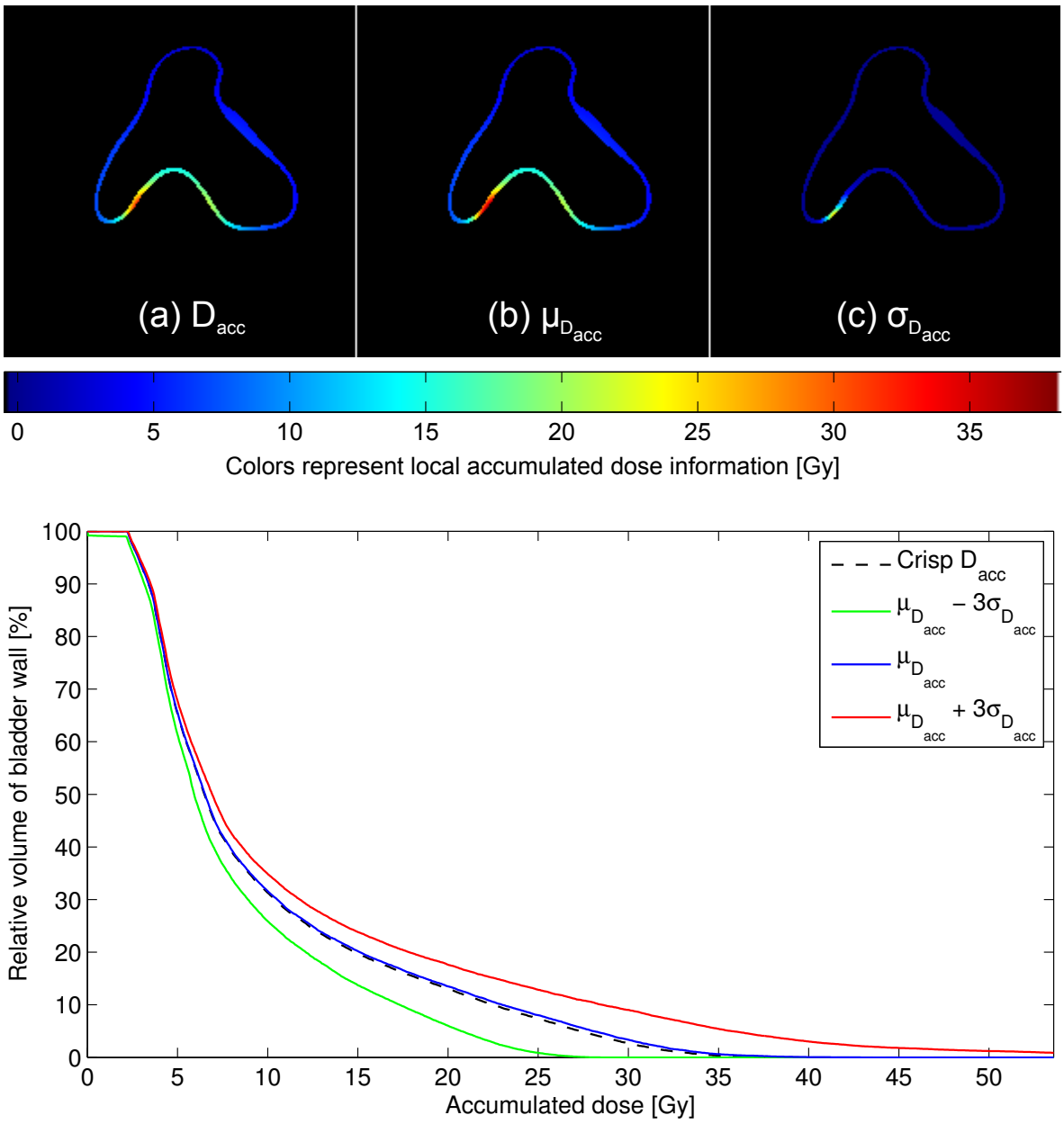


Figure 5.8: Effects of registration uncertainty (RU) on accumulated dose map for one patient. Left: comparison between accumulation of crisp aligned D_{acc} dose (a) versus the mean $\mu_{D_{acc}}$ (b) and standard deviation $\sigma_{D_{acc}}$ (c) of the accumulated RU-based diffusion of D_{acc}, which encodes the effects of RU at every voxel. Right: the effects of RU on the dose-volume histogram. The volume of the bladder wall presented in this example is 49 cm^3.

curves indicate that within six standard deviations (99.7% confidence interval) around the mean of the computed probabilistic dose values $\mu_{D_{acc}}$, the true D_{2cc} may lie anywhere within the range of 21.6 Gy (green line) to 37.8 Gy (red line). In Figure 5.8c, this region of uncertainty is shown to be spatially localized in the proximity of the small-volume/high-dose region, i.e., the radiation hot spot, which is of more radiobiological interest. The implications of this phenomena are crucial as we believe that this

observed localized variability in accumulated dose casts doubts on the usefulness of planning parameters that are derived from the DVH alone, and consequently leans in favor of methods that also account for the spatial distribution of the accumulated dose such as the method proposed by Zakariaee et al. [209]. The demonstrated effects of RU on the accumulated dose corroborates the surveyed prior art in other applications, yet their impact on MFCCB analyses is to be explored.

5.3.5 Computational Performance

Our pipeline is well suited to applications where image volumes from different modalities need to be registered due to our choice to use the established multimodal mutual information similarity metric readily implemented in `elastix`. Furthermore, our algorithm is fast; our unoptimized MATLAB plus `elastix` implementation takes, on average, 37 minutes to compute RU and 12 minutes to diffuse both the segmentation and dose volumes per fraction on an 8-core workstation with two Intel 3GHz Xeon x5472 processors and 8GB of RAM, which is faster than the Bayesian approach of Risholm et al. [148] that takes 11 hours per fraction on an 8-core machine. In our RU computation for every fraction, the initial registration and sample generation takes less than one minute and the following 40 registrations are independent and may be performed in parallel. Similarly, the diffusion equations Equation 5.8 may be evaluated at each voxel location independently, and therefore higher performance gains are expected with a larger computing cluster. It should be noted that the computational performance of Risholm et al. [148] can be improved with the help of modern sampling techniques, i.e., as was proposed in Simpson et al. [164] and Simpson et al. [161] with variational techniques for performing Bayesian inference. We were however unable to quantify the performance improvements proposed in Simpson et al. [164] and Simpson et al. [161] as performance times were not reported in their publications.

5.4 Summary

In this chapter, we presented a mathematical framework for estimating RU and propagating the effects of the computed uncertainties from the registration stage through to the dosimetric evaluation and visualization stages that are critical to the physician performing radiation therapy. Specifically, we presented a novel and computationally efficient RU computation technique that enables a parametric representation of RU as structure tensors, as well as two different methods for visualizing the computed RU. We further presented an uncertainty propagation method that encodes the RU into dosimetric maps and organ segmentation masks. Finally, we demonstrated how the RU encoded dosimetric maps computed for different fractions can be accumulated into an RU encoded accumulated dose map, which can then be summarized into a RU encoded DVH representation.

Our proposed framework is designed to compute RU from existing B-spline DIR software such as the popular `elastix` software package. Our framework can therefore be easily integrated within other image-guided radiotherapy applications as it generally does not require major changes to the existing medical image analysis workflow.

In the context of our showcased clinical application, we provided preliminary evidence indicating that the computation and propagation of RU results in improvements during treatment planning and

quality control stages of MFCCB. Specifically, we demonstrated how our proposed RU estimation and propagation can be used to visualize potential errors in DIR and, further, to transfer the delineation of organs from a prior fraction to the next more accurately. We also substantiated the utility of RU propagation in evaluating the potential errors and deviations in accumulated dose at a specific voxel, as well as in the DVH. More importantly, we observed that the resulting accumulated dose uncertainty is greatest in the proximity of the small-volume/high-dose region, i.e., the radiation hot spot, which is of significant radiobiological interest. Our next objectives include assessing the effects of RU on the correlation between accumulated dose and MFCCB patient toxicity outcomes. We also plan to extend our framework to group-wise or symmetric registration frameworks to mitigate uncertainties and potential biases associated with manual selection of one image volume as the reference image.

Chapter 6

Conclusions

Throughout this thesis, we presented novel directions towards fully automatic image-based scene analysis and augmentation for the benefit of image-guided therapy (IGT) procedures. We proposed novel motion-based segmentation methods that enable fast and safe localization of vascular structures from endoscopic video and dynamic ultrasound (DUS) sequences. We then presented novel methods for computing and encoding different sources of navigation uncertainties in the context of robot-assisted minimally invasive surgery (MIS) and fractionated radiotherapy. In this chapter, we conclude with a discussion of the strengths and weaknesses of our contributions and future research opportunities in IGT navigation afforded by the novel research directions which we have established.

6.1 Motion-Based Localization of Vasculature

In Chapter 2 and Chapter 3, we investigated how motion information may be used to automatically segment vasculature during IGT where there are multiple sources of motion, the data is noisy, vessels may be visually occluded by layers of tissues, and observable vascular motions are faint.

In Chapter 2, we presented our novel automatic phase-based motion segmentation (PBMS) method for localizing visually occluded vasculature from dynamic medical image sequences [5, 7], which we evaluated on a retrospective study of fifteen clinical robot-assisted partial nephrectomy (RAPN) procedures [7] and demonstrated quantitatively promising segmentation performance, i.e., a mean area under the receiver operating characteristics curve (AUC) of 0.72. To the best of our knowledge, we were the first to attempt and demonstrate promising results for the challenging task of localizing occluded vasculature in endoscopic video without the use of additional hardware or preoperative scans. In a follow-up study, we then presented the evaluation of our high-level variational scene segmentation method [126], which incorporates our PBMS methods in addition to other sources of information, on the same fifteen clinical RAPN procedures and demonstrated a 45% increase in pixel-wise accuracy for localizing renal vasculature compared to our original PBMS method.

In Chapter 3, we presented our kinematic model-based vessel segmentation (KMVS) method [7, 8], an extension of PBMS, that is designed to localize vasculature from dynamic medical image sequences by leveraging: (i) the complete estimation of local motion vectors and (ii) a novel pulsatile radial mo-

tion model (PRMM) that enables the modelling of divergent motion patterns. To enable fast and complete estimation of motion, we implemented a parallelizable technique, which estimates motion via the changes in the monogenic representation of image information. In total, we presented four alternative implementations of our KMVS method using different motion computation techniques and discussed the advantages of the different implementations. In our evaluations on a synthetic dataset and two real DUS datasets of the common carotid artery (CCA), we demonstrated that, compared to PBMS, our fast OF2 implementation increases the mean AUC from 0.82 to 0.99 on the UBC dataset and from 0.83 to 0.98 on the SPL dataset.

We emphasize that our PBMS and KMVS methods were not designed for any specific imaging modality, and that our proposed techniques may therefore be employed within many different image-guided interventions that stand to benefit from automatic blood vessel localization, e.g., robot-assisted prostate cancer surgery [109, 110] and stereotactic neurosurgery [138]. To promote the adoption of our methodologies for real-time guidance applications, we discussed real-time alternatives to the temporal filtering techniques which we utilized in our applications. Finally, we have made the MATLAB executables[1] of our PBMS and KMVS methods publicly available for download. We hope that, in doing so, we encourage fellow researchers to incorporate our low-level motion features into their specific segmentation pipelines. In addition to the code, we provided our phantom dataset to allow others to compare the performance of their methods to ours.

Though we conducted a preliminary user study of our proposed PBMS methodologies, additional clinical evaluations are required to demonstrate the quantifiable clinical benefits of motion-based vessel segmentation and augmentation in terms of improved IGT outcomes. The most pivotal and natural extension to the works presented in Chapter 2 and Chapter 3 would be to integrate the proposed methodologies directly into a prototypical image-guided navigation system. Such an implementation would shed more insight into the clinical utility of our methodologies and would allow for further investigation into the most effective strategy for optimizing the trade-off between localization accuracy and lag-times that are inherent to our methods.

Another compelling research direction, which presents itself as an extension to our works, is to investigate whether motion-based cues may be leveraged to distinguish between venous and arterial structures. In our preliminary studies [5, 7], we observed that the pulsations of the renal artery and the renal vein occur out-of-phase with one another. Though this phenomenon was observed in almost all of the fifteen cases in our study, this effect may not present itself at other locations in the human body. Nonetheless, flow rates and vascular pressures are drastically different between venous and arterial structures [198], and these differences may perhaps manifest as distinguishing pulsatile motion signatures that can in turn be used as features to automatically discriminate between arteries and veins. Furthermore, a motion-based discriminative model may also be used to automatically detect pathological vasculature or aneurysms based solely on apparent motions. It is important to note that kinematic differences between veins versus arteries or healthy versus pathological vessels are likely to be very complex and may not be simple to model mathematically. It is therefore probable that future investiga-

[1] MATLAB executables are available for download from https://bisicl.ece.ubc.ca/software/radialDistension.html

tion into such discriminative frameworks will involve data-driven machine learning techniques designed to exploit the capabilities of our proposed motion-based features.

6.2 Computation, Propagation, and Visualization of Navigation Uncertainties

In Chapter 4 and Chapter 5, we investigated ways in which navigation uncertainties can be computed, propagated, and visualized, specifically in the context of computer-assisted IGT navigation systems that target deformable soft-tissues.

In Chapter 4, we presented a proof-of-concept endoscopic scene augmentation method for facilitating the registration of probabilistic preoperative computed tomography (CT) segmentations with stereo endoscopic video data [4]. We proposed an uncertainty-encoded computational stereopsis technique for extracting probabilistic surface information from stereo endoscopic data and used this probabilistic surface to register probabilistic preoperative CT segmentations with stereo endoscopic scene. We applied our framework to an *ex vivo* lamb kidney phantom to simulate the tumor demarcation stage of RAPN interventions. We proposed novel uncertainty-encoded visualization techniques for presenting probabilistic tumor margins onto the endoscopic scene and discussed the potential advantages of our proposed visualizations compared to existing crisp (deterministic) techniques. One such potential advantage is that the ability to visualize uncertainties associated to preoperative imaging—and therefore the associated resection margins—may guide the surgeon's attention to localized regions of high uncertainty, at which point the surgeon may choose to revise the resection strategy or opt for additional intraoperative imaging, e.g., ultrasound, to improve confidence in the localization.

Our proposed visualizations are not exclusive to preoperative CT imaging or the context of RAPN. Our contributions are extendable to other IGT applications that stand to benefit from visualization of uncertainties, namely other robot-assisted interventions. Furthermore, the proposed augmentations may also be used in surgical navigation frameworks that make use of fiducial markers, instead of image registration, to align preoperative data with the intraoperative scene. In such applications, the effects of calibration and tracking errors may also be estimated and propagated onto the visualizations of the probabilistic margins [159].

Our uncertainty-encoded surgical navigation method stands to benefit from further rigorous validation and summative usability tests to fully evaluate the utility of our uncertainty-encoded visualization techniques in the clinical workflow. The most challenging component in our initial proof-of-concept framework, which we did not fully address in Amir-Khalili et al. [3] and have since tried to overcome, is the probabilistic registration piece to the uncertainty-encoded navigation puzzle. As we further elaborate below, probabilistic image registration is difficult to perform in real-time. Nonetheless, with the exception of the registration component, real-time performance is attainable for all other components of our framework.

Deformable image registration (DIR) is an important source of uncertainty in computer-assisted navigation systems and there is presently a need for fast algorithms to compute registration uncertainty (RU) in DIR and to then propagate the effects of RU onto the augmented scene. DIR methods typi-

cally entail computationally expensive numerical optimizations and are thus often unable to operate in real-time. This computational cost is further exacerbated when RU is to be computed from a sample containing multiple potential DIR solutions. Some intraoperative surgical navigation applications, such as tumor resection in RAPN, demand real-time performance and there yet remains the open problem of whether DIR and RU-encoding methodologies are ever feasible in such applications.

Towards the goal of developing fast and computationally efficient RU-encoding methodology, in Chapter 5, we implemented an end-to-end mathematical framework that estimates RU in DIR and subsequently propagates the effects of the computed uncertainties from the registration stage through to the visualizations, organ segmentations, and dosimetric evaluations that are critical in the context of multi-fraction cervical cancer brachytherapy (MFCCB) [9]. In our framework, we proposed a novel method for computing RU that is designed to: (i) interface with existing multimodal DIR software, which we deployed using `elastix`, and (ii) represent RU in a parametric manner using structure tensors. We also proposed a weighted averaging technique for propagating the effects of RU, onto volumetric segmentation and dose data, to produce a probabilistic map of aligned segmentation and dose information subject to the estimated RU. We evaluated our framework on a retrospective study consisting of 37 patients and presented preliminary evidence that our proposed framework may be advantageous to MFCCB applications. Specifically, we showed that (i) the effect of RU on dose accumulation provide useful insights for quality control and post-treatment analysis; (ii) RU propagation improves the transfer of delineations from one fraction to the next; and (iii) RU can be used to generate visualizations that reflect the quality of DIR that may prove to assist physicians in making decisions based on registered image data. Furthermore, our RU computation and propagation methods are parallelizable.

A limitation of our RU estimation method, in fact RU estimation for DIR in general, is the long computation time. In IGT applications where strict real-time constraints are not imposed, such as the showcased MFCCB application, our framework can be integrated into the clinical workflow following simple performance optimizations and an implementation on a dedicated high-performance computing cluster. On the other hand, real-time RU computation and propagation in surgical navigation contexts is currently not feasible despite the attempts towards fast RU estimation. Concurrent with our efforts, the generative Bayesian alternatives to RU estimation presented in Risholm et al. [147] and Risholm et al. [148] have moved towards faster mean-field variational Bayesian approximation techniques as presented in Simpson et al. [164] and Simpson et al. [161]. These approaches can easily incorporate other data-terms, such as the visual cues and elastically deformable shape priors that we proposed in the variation framework of Nosrati et al. [126], as regularizers in order to improve the overall performance of the registration framework. Though these computationally expensive generative frameworks are seemingly far from real-time operation, in light of the continuing advancements towards more powerful and affordable graphical processing units (GPUs), real-time performance may be attainable in the near future. Similarly, our proposed uncertainty encoding framework can also benefit from advancements in GPU-based computation techniques. Therefore, at this point in time, the best strategy for encoding the effects of RU into real-time IGT navigation systems remains, ironically, shrouded in uncertainty.

Bibliography

[1] M. Alessandrini, H. Liebgott, D. Barbosa, and O. Bernard. Monogenic phase based optical flow computation for myocardial motion analysis in 3D echocardiography. In *International Workshop on Statistical Atlases and Computational Models of the Heart*, pages 159–168. Springer, 2012.

[2] M. Alessandrini, A. Basarab, H. Liebgott, and O. Bernard. Myocardial motion estimation from medical images using the monogenic signal. *Transactions on Image Processing*, 22(3): 1084–1095, 2013.

[3] A. Amir-Khalili, A. J. Hodgson, and R. Abugharbieh. Real-time extraction of local phase features from volumetric medical image data. In *International Symposium on Biomedical Imaging*, pages 930–933. IEEE, 2013.

[4] A. Amir-Khalili, M. S. Nosrati, J.-M. Peyrat, G. Hamarneh, and R. Abugharbieh. Uncertainty-encoded augmented reality for robot-assisted partial nephrectomy: a phantom study. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 182–191. Springer, 2013.

[5] A. Amir-Khalili, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, G. Hamarneh, and R. Abugharbieh. Auto localization and segmentation of occluded vessels in robot-assisted partial nephrectomy. In *Medical Image Computing and Computer-Assisted Intervention*, pages 407–414. Springer, 2014.

[6] A. Amir-Khalili, G. Hamarneh, and R. Abugharbieh. Automatic vessel segmentation from pulsatile radial distension. In *Medical Image Computing and Computer-Assisted Intervention*, pages 403–410. Springer, 2015.

[7] A. Amir-Khalili, G. Hamarneh, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, and R. Abugharbieh. Automatic segmentation of occluded vasculature via pulsatile motion analysis in endoscopic robot-assisted partial nephrectomy video. *Medical Image Analysis*, 25(1): 103–110, 2015.

[8] A. Amir-Khalili, G. Hamarneh, and R. Abugharbieh. Modelling and extraction of pulsatile radial distension and compression motion for automatic vessel segmentation from video. *Medical Image Analysis*, 40:184–198, June 2017.

[9] A. Amir-Khalili, G. Hamarneh, R. Zakariaee, I. Spadinger, and R. Abugharbieh. Propagation of registration uncertainty during multi-fraction cervical cancer brachytherapy. *Physics in Medicine and Biology*, 62(20):8116–8135, Oct. 2017.

[10] S. Andrews and G. Hamarneh. The generalized log-ratio transformation: learning shape and adjacency priors for simultaneous thigh muscle segmentation. *Transactions on Medical Imaging*, 34(9):1773–1787, 2015.

[11] S. Andrews, C. McIntosh, and G. Hamarneh. Convex multi-region probabilistic segmentation with shape prior in the isometric log-ratio transformation space. In *International Conference on Computer Vision*, pages 2096–2103. IEEE, 2011.

[12] M. Baumhauer, T. Simpfendörfer, B. Müller-Stich, D. Teber, C. Gutt, J. Rassweiler, H. P. Meinzer, and I. Wolf. Soft tissue navigation for laparoscopic partial nephrectomy. *International Journal of Computer Assisted Radiology and Surgery*, 3(3):307–314, 2008.

[13] C. Becker, R. Rigamonti, V. Lepetit, and P. Fua. Supervised feature learning for curvilinear structure segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, pages 526–533. Springer, 2013.

[14] P. N. Belhumeur. A Bayesian approach to binocular steropsis. *International Journal of Computer Vision*, 19(3):237–260, 1996.

[15] S. Bernhardt, J. Abi-Nahed, and R. Abugharbieh. Robust dense endoscopic stereo reconstruction for minimally invasive surgery. In *MICCAI Workshop on Medical Computer Vision*, pages 198–207. Springer, 2012.

[16] S. Bernhardt, S. A. Nicolau, L. Soler, and C. Doignon. The status of augmented reality in laparoscopic surgery as of 2016. *Medical Image Analysis*, 37:66–90, 2017. ISSN 1361-8415.

[17] M. L. Bots, A. W. Hoes, P. J. Koudstaal, A. Hofman, and D. E. Grobbee. Common carotid intima-media thickness and risk of stroke and myocardial infarction the Rotterdam Study. *Circulation*, 96(5):1432–1437, 1997.

[18] D. Boukerroui, J. A. Noble, and M. Brady. On the choice of band-pass quadrature filters. *Journal of Mathematical Imaging and Vision*, 21(1-2):53–80, 2004.

[19] K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler. Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy committee task group No. 132. *Medical Physics*, 2017.

[20] A. Bruhn, J. Weickert, and C. Schnörr. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005.

[21] J. C. Byrn, S. Schluender, C. M. Divino, J. Conrad, B. Gurland, E. Shlasko, and A. Szold. Three-dimensional imaging improves surgical performance for both novice and experienced operators using the daVinci Robot System. *The American Journal of Surgery*, 193(4):519–522, 2007.

[22] J. T. Carter, C. E. Freise, R. A. McTaggart, H. D. Mahanty, S.-M. Kang, S. H. Chan, S. Feng, J. P. Roberts, and A. M. Posselt. Laparoscopic procurement of kidneys with multiple renal arteries is associated with increased ureteral complications in the recipient. *American Journal of Transplantation*, 5(6):1312–1318, 2005.

[23] J. L. Cavalcante, J. A. Lima, A. Redheuil, and M. H. Al-Mallah. Aortic stiffness: current understanding and future directions. *Journal of the American College of Cardiology*, 57(14): 1511–1522, 2011.

[24] L. E. Chambless, A. R. Folsom, L. X. Clegg, A. R. Sharrett, E. Shahar, F. J. Nieto, W. D. Rosamond, and G. Evans. Carotid wall thickness is predictive of incident clinical stroke: the

atherosclerosis risk in communities (ARIC) study. *American Journal of Epidemiology*, 151(5): 478–487, 2000.

[25] R. A. Chaudhury, V. Atlasman, G. Pathangey, N. Pracht, R. J. Adrian, and D. H. Frakes. A high performance pulsatile pump for aortic flow experiments in 3-dimensional models. *Cardiovascular Engineering and Technology*, 7(2):148–158, 2016.

[26] S. K. Chavan, N. R. Wabale, and R. S. Daimi. Unusual variation of the renal vessels — a case report. *Pravara Medical Review*, 5(3):32–34, 2010.

[27] N. Chenouard and M. Unser. 3D steerable wavelets and monogenic analysis for bioimaging. In *International Symposium on Biomedical Imaging*, pages 2132–2135. IEEE, 2011.

[28] K. Cleary and T. M. Peters. Image-guided interventions: technology review and clinical applications. *Annual Review of Biomedical Engineering*, 12:119–142, 2010.

[29] N. J. Crane, S. M. Gillern, K. Tajkarimi, I. W. Levin, P. A. Pinto, and E. A. Elster. Visual enhancement of laparoscopic partial nephrectomy with 3-charge coupled device camera: assessing intraoperative tissue perfusion and vascular anatomy by visible hemoglobin spectral response. *The Journal of Urology*, 184(4):1279–1285, 2010.

[30] L. A. Dawson and M. B. Sharpe. Image-guided radiotherapy: rationale, benefits, and limitations. *The Lancet Oncology*, 7(10):848–858, 2006.

[31] C. De Martino and L. Accinni. *Embryogenesis of the kidney*, pages 25–51. Springer US, Boston, MA, 1985. ISBN 978-1-4613-2575-8.

[32] P. De Visschere, W. Oosterlinck, G. De Meerleer, and G. Villeirs. Clinical and imaging tools in the early diagnosis of prostate cancer, a review. *Journal of the Belgian Society of Radiology*, 93 (2), 2010.

[33] B. J. Drucker. Renal cell carcinoma: current status and future prospects. *Cancer Treatment Reviews*, 31(7):536–545, 2005.

[34] R. Duivenvoorden, E. de Groot, B. M. Elsen, J. S. Laméris, R. J. van der Geest, E. S. Stroes, J. J. Kastelein, and A. J. Nederveen. In vivo quantification of carotid artery wall dimensions. *Circulation: Cardiovascular Imaging*, 2(3):235–242, 2009.

[35] P. Edgcumbe, R. Singla, P. Pratt, C. Schneider, C. Nguan, and R. Rohling. Augmented reality imaging for robot-assisted partial nephrectomy surgery. In *International Conference on Medical Imaging and Virtual Reality*, pages 139–150. Springer, 2016.

[36] G. Eggers, J. Mühling, and R. Marmulla. Image-to-patient registration techniques in head surgery. *International journal of oral and maxillofacial surgery*, 35(12):1081–1095, 2006.

[37] N. Enayati, E. De Momi, and G. Ferrigno. Haptics in robot-assisted surgery: challenges and benefits. *Reviews in Biomedical Engineering*, 9:49–65, 2016.

[38] R. S. J. Estépar and K. G. Vosburgh. multimodality guidance in endoscopic and laparoscopic abdominal procedures. In *Intraoperative Imaging and Image-Guided Therapy*, pages 767–778. Springer, 2014.

[39] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

[40] A. Fedorov, P. Risholm, and W. M. Wells. Deformable registration for IGT. In *Intraoperative Imaging and Image-Guided Therapy*, pages 211–223. Springer, 2014.

[41] J. Feliciano and M. Stifelman. Robotic retroperitoneal partial nephrectomy: a four-arm approach. *Journal of the Society of Laparoendoscopic Surgeons*, 16(2):208, 2012.

[42] M. Felsberg. Optical flow estimation from monogenic phase. In *Complex Motion*, pages 1–13. Springer, 2007.

[43] M. Felsberg and G. Sommer. The monogenic signal. *Transactions on Signal Processing*, 49(12): 3136–3144, 2001.

[44] M. Felsberg and G. Sommer. The monogenic scale-space: a unifying approach to phase-based image processing in scale-space. *Journal of Mathematical Imaging and Vision*, 21(1-2):5–26, 2004.

[45] A. Fenster, D. B. Downey, and H. N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 46(5):R67–R99, 2001.

[46] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379–2394, 1987.

[47] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, 1990.

[48] M. Foskey, B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman, and S. Joshi. Large deformation three-dimensional image registration in image-guided radiation therapy. *Physics in Medicine and Biology*, 50(24):5869, 2005.

[49] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention*, pages 130–137. Springer, 1998.

[50] R. Freelove and A. D. Walling. Pancreatic cancer: diagnosis and management. *American Family Physician*, 73(3):485–492, 2006.

[51] K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner. Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3):346–355, 1996.

[52] S. Gao, R. van't Klooster, A. Brandts, S. D. Roes, R. Alizadeh Dehnavi, A. de Roos, J. J. Westenberg, and R. J. Geest. Quantification of common carotid artery and descending aorta vessel wall thickness from MR vessel wall imaging using a fully automated processing pipeline. *Journal of Magnetic Resonance Imaging*, 45(1):215–228, 2017.

[53] A. Gastounioti, A. Sotiras, K. S. Nikita, and N. Paragios. Graph-based motion-driven segmentation of the carotid atherosclerotique plaque in 2D ultrasound sequences. In *Medical Image Computing and Computer-Assisted Intervention*, pages 551–559. Springer, 2015.

[54] T. Gautama and M. M. Van Hulle. A phase-based approach to the estimation of the optical flow field using spatial filtering. *Transactions on Neural Networks*, 13(5):1127–1136, 2002.

[55] P. Georg, S. Lang, J. C. Dimopoulos, W. Dörr, A. E. Sturdza, D. Berger, D. Georg, C. Kirisits, and R. Pötter. Dose-volume histogram parameters and late side effects in magnetic resonance image-guided adaptive cervical cancer brachytherapy. *International Journal of Radiation Oncology*Biology*Physics*, 79(2):356–362, 2011.

[56] S. Giannarou, M. Visentini-Scarzanella, and G.-Z. Yang. Probabilistic tracking of affine-invariant anisotropic regions. *Transactions on Pattern Analysis and Machine Intelligence*, 35(1):130–143, Jan. 2013. ISSN 0162-8828.

[57] I. S. Gill, M. M. Desai, J. H. Kaouk, A. M. Meraney, D. P. Murphy, G. T. Sung, and A. C. Novick. Laparoscopic partial nephrectomy for renal tumor: duplicating open surgical techniques. *The Journal of Urology*, 167(2, Part 1):469–476, 2002. ISSN 0022-5347.

[58] L. Grady. Random walks for image segmentation. *Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.

[59] W. Grossman. *Cardiac catheterization and angiography*. Lea and Febiger,Philadelphia, PA, Jan. 1986.

[60] M. E. Guerrero, R. Jacobs, M. Loubele, F. Schutyser, P. Suetens, and D. van Steenberghe. State-of-the-art on cone beam CT imaging for preoperative planning of implant placement. *Clinical oral investigations*, 10(1):1–7, 2006.

[61] R. Gupta, C. Walsh, I. S. Wang, M. Kachelrieß, J. Kuntz, and S. Bartling. CT-guided interventions: current practice and future directions. In *Intraoperative Imaging and Image-Guided Therapy*, pages 173–191. Springer, 2014.

[62] I. Hacihaliloglu, R. Abugharbieh, A. J. Hodgson, R. N. Rohling, and P. Guy. Automatic bone localization and fracture detection from volumetric ultrasound images using 3-D local phase features. *Ultrasound in Medicine & Biology*, 38(1):128–144, 2012.

[63] E. J. Halpern, D. G. Mitchell, R. J. Wechsler, E. K. Outwater, M. J. Moritz, and G. A. Wilson. Preoperative evaluation of living renal donors: comparison of CT angiography and MR angiography. *Radiology*, 216(2):434–439, 2000.

[64] G. Hamarneh, A. Amir-Khalili, M. S. Nosrati, I. Figueroa, J. Kawahara, O. Al-Alao, J.-M. Peyrat, J. Abi-Nahed, A. Al-Ansari, and R. Abugharbieh. Towards multi-modal image-guided tumour identification in robot-assisted partial nephrectomy. In *Middle East Conference on Biomedical Engineering*, pages 159–162. IEEE, 2014.

[65] K. Hameeteman, S. Rozie, C. Metz, R. Manniesing, T. van Walsum, A. van der Lugt, W. Niessen, and S. Klein. Automatic carotid artery distensibility measurements from CTA using nonrigid registration. *Medical Image Analysis*, 17(5):515–524, 2013. ISSN 1361-8415.

[66] G. K. Hansson. Inflammation, atherosclerosis, and coronary artery disease. *New England Journal of Medicine*, 352(16):1685–1695, 2005.

[67] A. Harraz, A. Shokeir, S. Soliman, A. El-Hefnawy, M. Kamal, I. Shalaby, A. Kamal, and M. Ghoneim. Fate of accessory renal arteries in grafts with multiple renal arteries during live-donor renal allo-transplantation. In *Transplantation Proceedings*, volume 45 of *3*, pages 1232–1236. Elsevier, 2013.

[68] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge Univ Press, second edition, 2004. ISBN 0521540518.

[69] M. P. Heinrich, I. J. Simpson, B. W. Papież, M. Brady, and J. A. Schnabel. Deformable image registration by combining uncertainty estimates from supervoxel belief propagation. *Medical Image Analysis*, 27:57–71, 2016.

[70] T. P. Hellebust, E. Dale, A. Skjonsberg, and D. R. Olsen. Inter fraction variations in rectum and bladder volumes and dose distributions during high dose rate brachytherapy treatment of the uterine cervix investigated by repetitive CT-examinations. *Radiotherapy and Oncology*, 60(3): 273–280, 2001. ISSN 0167-8140.

[71] C. Hennersperger, M. Baust, P. Waelkens, A. Karamalis, S. Ahmadi, and N. Navab. Multiscale tubular structure detection in ultrasound imaging. *Transactions on Medical Imaging*, 34(1): 13–26, 2015.

[72] R. L. Holloway. Registration error analysis for augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):413–432, 1997.

[73] B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1):185–203, 1981. ISSN 0004-3702.

[74] X. Hu and P. Mordohai. Evaluation of stereo confidence indoors and outdoors. In *Computer Vision and Pattern Recognition*, pages 1466–1473. IEEE, 2010.

[75] M. Hub and C. Karger. Estimation of the uncertainty of elastic image registration with the demons algorithm. *Physics in Medicine and Biology*, 58(9):3023, 2013.

[76] M. Hub, M. L. Kessler, and C. P. Karger. A stochastic approach to estimate the uncertainty involved in B-spline image registration. *Transactions on Medical Imaging*, 28(11):1708–1716, 2009.

[77] J. S. Huber, Q. Peng, and W. W. Moses. Multi-modality phantom development. *Transactions on Nuclear Science*, 56(5):2722–2727, 2009.

[78] J. E. Iglesias and M. R. Sabuncu. Multi-atlas segmentation of biomedical images: a survey. *Medical Image Analysis*, 24(1):205–219, 2015.

[79] J. E. Iglesias, M. R. Sabuncu, and K. Van Leemput. Improved inference in Bayesian segmentation using Monte Carlo sampling: application to hippocampal subfield volumetry. *Medical Image Analysis*, 17(7):766–778, 2013.

[80] J. A. Izatt, M. D. Kulkarni, S. Yazdanfar, J. K. Barton, and A. J. Welch. In vivo bidirectional color doppler flow imaging of picoliter blood volumes using optical coherence tomography. *Optics letters*, 22(18):1439–1441, 1997.

[81] P. Jannin, E. Krupinski, and S. Warfield. Validation in medical image processing. *Transactions on Medical Imaging*, 25(11):1405, 2006.

[82] F. Jolesz. *Intraoperative imaging and image-guided therapy*. Springer Science & Business Media, 2014.

[83] F. A. Jolesz and F. Shtern. The operating room of the future. *Investigative Radiology*, 27(4): 326–328, 1992.

[84] S. Josephson, S. Bryant, H. Mak, S. Johnston, W. Dillon, and W. Smith. Evaluation of carotid stenosis using CT angiography in the initial evaluation of stroke and TIA. *Neurology*, 63(3): 457–460, 2004.

[85] L. Joskowicz. Computer-aided surgery meets predictive, preventive, and personalized medicine. *EPMA Journal*, 8(1):1–4, 2017.

[86] I. R. Kamel, J. B. Kruskal, E. A. Pomfret, M. T. Keogan, G. Warmbrand, and V. Raptopoulos. Impact of multidetector CT on donor selection and surgical planning before living adult right lobe liver transplantation. *American Journal of Roentgenology*, 176(1):193–200, 2001.

[87] S. Kim, J. Kang, and M. J. Chung. Probabilistic voxel mapping using an adaptive confidence measure of stereo matching. *Intelligent Service Robotics*, pages 1–11, 2013.

[88] S. H. Kim and M. C. Han. Reversed contrast-urine levels in urinary bladder: CT findings. *Urologic radiology*, 13(1):249–252, 1991.

[89] C. Kirbas and F. K. Quek. Vessel extraction techniques and algorithms: a survey. In *Symposium on Bioinformatics and Bioengineering*, pages 238–245. IEEE, 2003.

[90] H. Kirişli, M. Schaap, C. Metz, A. Dharampal, W. Meijboom, S. Papadopoulou, A. Dedic, K. Nieman, M. de Graaf, M. Meijs, et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. *Medical Image Analysis*, 17(8):859–876, 2013.

[91] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim. Elastix: a toolbox for intensity based medical image registration. *Transactions on Medical Imaging*, 29:196–205, 2010.

[92] E. Konukoglu, O. Clatz, P. Bondiau, H. Delingette, and N. Ayache. Extrapolating glioma invasion margin in brain magnetic resonance images: suggesting new irradiation margins. *Medical Image Analysis*, 14(2):111–125, 2010.

[93] P. Kovesi. *Invariant measures of image features from phase information*. University of Western Australia, 1996.

[94] J. Kybic. Bootstrap resampling for image registration uncertainty estimation without ground truth. *Transactions on Image Processing*, 19(1):64–73, 2010.

[95] A. R. Lanfranco, A. E. Castellanos, J. P. Desai, and W. C. Meyers. Robotic surgery: a current perspective. *Annals of Surgery*, 239(1):14, 2004.

[96] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger. PLUS: open-source toolkit for ultrasound-guided intervention systems. *Transactions on Biomedical Engineering*, 61 (10):2527–2537, Oct. 2014. ISSN 0018-9294.

[97] S. Laurent, S. Katsahian, C. Fassot, A.-I. Tropeano, I. Gautier, B. Laloux, and P. Boutouyrie. Aortic stiffness is an independent predictor of fatal stroke in essential hypertension. *Stroke*, 34 (5):1203–1206, 2003.

[98] M. W. Law and A. C. Chung. Three dimensional curvilinear structure detection using optimally oriented flux. In *European Conference on Computer Vision*, pages 368–382. Springer, 2008.

[99] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea. A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes. *Medical Image Analysis*, 13 (6):819–845, 2009.

[100] C. Liu. *Beyond pixels: exploring new representations and applications for motion analysis*. PhD thesis, Massachusetts Institute of Technology, 2009.

[101] L. M. Lorigo, O. D. Faugeras, W. E. L. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C.-F. Westin. Curves: curve evolution for vessel segmentation. *Medical Image Analysis*, 5(3): 195–206, 2001.

[102] T. Lotfi, L. Tang, S. Andrews, and G. Hamarneh. Improving probabilistic image registration via reinforcement learning and uncertainty evaluation. In *MICCAI Workshop on Machine Learning in Medical Imaging*, pages 187–194. Springer, 2013.

[103] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, volume 81, pages 674–679, 1981.

[104] T. Maecken and T. Grau. Ultrasound imaging in vascular access. *Critical Care Medicine*, 35(5): 178–185, 2007.

[105] R. Manniesing, M. Schaap, S. Rozie, R. Hameeteman, D. Vukadinovic, A. van der Lugt, and W. Niessen. Robust CTA lumen segmentation of the atherosclerotic carotid artery bifurcation in a large patient population. *Medical Image Analysis*, 14(6):759–769, 2010. ISSN 1361-8415.

[106] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Medical image analysis*, 16(3):642–661, 2012.

[107] C. McIntosh and G. Hamarneh. Vessel crawlers: 3D physically-based deformable organisms for vasculature segmentation and analysis. In *Computer Vision and Pattern Recognition*, volume 1, pages 1084–1091. IEEE, 2006.

[108] C. McIntosh and G. Hamarneh. Optimal weights for convex functionals in medical image segmentation. In *Advances in Visual Computing*, pages 1079–1088. Springer, 2009.

[109] A. J. McLeod, J. S. Baxter, S. de Ribaupierre, and T. M. Peters. Motion magnification for endoscopic surgery. In *SPIE: Medical Imaging*, volume 9036, 2014.

[110] A. J. McLeod, J. S. Baxter, U. Jayarathne, S. Pautler, T. M. Peters, and X. Luo. Stereoscopic motion magnification in minimally-invasive robotic prostatectomy. In *MICCAI Workshop on Computer-Assisted and Robotic Endoscopy*, pages 35–45. Springer, 2015.

[111] J. A. McLeod, J. S. H. Baxter, G. Ameri, S. Ganapathy, T. M. Peters, and E. C. S. Chen. Detection and visualization of dural pulsation for spine needle interventions. *International Journal of Computer Assisted Radiology and Surgery*, 10(6):947–958, 2015. ISSN 1861-6429.

[112] R. R. Mechoor, T. Schmidt, and E. Kung. A real-time programmable pulsatile flow pump for in vitro cardiovascular experimentation. *Journal of Biomechanical Engineering*, 138(11):111002, 2016.

[113] S. A. Merritt, L. Rai, and W. E. Higgins. Real-time CT-video registration for continuous endoscopic guidance. In *SPIE: Medical Imaging*, volume 6143, page 614313, 2006.

[114] U. Mezger, C. Jendrewski, and M. Bartels. Navigation in surgery. *Langenbecks Archives of Surgery*, 398(4):501–514, 2013.

[115] D. J. Mirota, M. Ishii, and G. D. Hager. Vision-based navigation in image-guided interventions. *Annual Review of Biomedical Engineering*, 13:297–319, 2011.

[116] M. Moradi, P. Abolmaesumi, D. R. Siemens, E. E. Sauerbrei, A. H. Boag, and P. Mousavi. Augmenting detection of prostate cancer in transrectal ultrasound images using SVM and RF time series. *Transactions on Biomedical Engineering*, 56(9):2214–2224, 2009.

[117] A. Mottrie, G. De Naeyer, P. Schatteman, P. Carpentier, M. Sangalli, and V. Ficarra. Impact of the learning curve on perioperative outcomes in patients who underwent robotic partial nephrectomy for parenchymal renal tumours. *European Urology*, 58(1):127–133, 2010.

[118] K. Murari, N. Li, A. Rege, X. Jia, A. All, and N. Thakor. Contrast-enhanced imaging of cerebral vasculature with laser speckle. *Applied Optics*, 46(22):5340–5346, 2007.

[119] M. J. Murphy, F. J. Salguero, J. V. Siebers, D. Staub, and C. Vaman. A method to estimate the effect of deformable image registration uncertainties on daily dose mapping. *Medical Physics*, 39(2):573–580, 2012.

[120] S. Nag, H. Cardenes, S. Chang, I. J. Das, B. Erickson, G. S. Ibbott, J. Lowenstein, J. Roll, B. Thomadsen, and M. Varia. Proposed guidelines for image-based intracavitary brachytherapy for cervical carcinoma: report from image-guided brachytherapy working group. *International Journal of Radiation Oncology*Biology*Physics*, 60(4):1160–1172, 2004.

[121] D. P. Naidich, R. Sussman, W. L. Kutcher, C. P. Aranda, S. M. Garay, and N. A. Ettenger. Solitary pulmonary nodules: CT-bronchoscopic correlation. *Chest*, 93(3):595–598, 1988.

[122] S. Y. Nakada and S. Hedican. *Essential urologic laparoscopy: the complete clinical guide*. Springer Science & Business Media, 2009.

[123] M. S. Nosrati, S. Andrews, and G. Hamarneh. Bounded labeling function for global segmentation of multi-part objects with geometric constraints. In *International Conference on Computer Vision*, pages 2032–2039. IEEE, 2013.

[124] M. S. Nosrati, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, R. Abugharbieh, and G. Hamarneh. Efficient multi-organ segmentation in multi-view endoscopic videos using pre-operative priors. In *Medical Image Computing and Computer-Assisted Intervention*, pages 324–331. Springer, 2014.

[125] M. S. Nosrati, R. Abugharbieh, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, and G. Hamarneh. Simultaneous multi-structure segmentation and 3d nonrigid pose estimation in image-guided robotic surgery. *Transactions on Medical Imaging*, 35(1):1–12, 2016.

[126] M. S. Nosrati, A. Amir-Khalili, J.-M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, R. Abugharbieh, and G. Hamarneh. Endoscopic scene labelling and augmentation using intraoperative pulsatile motion and colour appearance cues with preoperative anatomical priors. *The International Journal for Computer Assisted Radiology and Surgery (IJCARS)*, 11(8): 1409–1418, 2016.

[127] K. Pauwels and M. M. Van Hulle. Realtime phase-based optical flow on the GPU. In *Computer Vision and Pattern Recognition Workshops*, pages 1–8. IEEE, 2008.

[128] L. Pedersen, M. Grunkin, B. Ersbøll, K. Madsen, M. Larsen, N. Christoffersen, and U. Skands. Quantitative measurement of changes in retinal vessel diameter in ocular fundus images. *Pattern Recognition Letters*, 21(13):1215–1223, 2000.

[129] P. J. Pickhardt, C. Hassan, S. Halligan, and R. Marmo. Colorectal cancer: CT colonography and colonoscopy for detection—systematic review and meta-analysis. *Radiology*, 259(2):393–405, 2011.

[130] K. M. Pohl, J. Fisher, J. J. Levitt, M. E. Shenton, R. Kikinis, W. E. L. Grimson, and W. M. Wells. A unifying approach to registration, segmentation, and intensity correction. In *Medical Image Computing and Computer-Assisted Intervention*, pages 310–318. Springer, 2005.

[131] M. Pollefeys, R. Koch, and L. Van Gool. A simple and efficient rectification method for general motion. In *International Conference on Computer Vision*, pages 496–501, 1999.

[132] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–70, 2000.

[133] R. Pötter, C. Haie-Meder, E. Van Limbergen, I. Barillot, M. De Brabandere, J. Dimopoulos, I. Dumas, B. Erickson, S. Lang, A. Nulens, P. Petrow, J. Rownd, and C. Kirisits. Recommendations from gynaecological (GYN) GEC ESTRO working group (ii): concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy—3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. *Radiotherapy and Oncology*, 78(1):67–77, 2006.

[134] P. Pratt, D. Stoyanov, M. Visentini-Scarzanella, and G.-Z. Yang. Dynamic guidance for robotic surgery using image- constrained biomechanical models. In *Medical Image Computing and Computer-Assisted Intervention*, pages 77–85. Springer, 2010.

[135] P. Pratt, E. Mayer, J. Vale, D. Cohen, E. Edwards, A. Darzi, and G.-Z. Yang. An effective visualisation and registration system for image-guided robotic partial nephrectomy. *Journal of Robotic Surgery*, pages 1–9, 2012.

[136] V. A. Prisacariu and I. D. Reid. PWP3D: real-time segmentation and tracking of 3D objects. *International Journal of Computer Vision*, 98(3):335–354, 2012.

[137] G. Puerto-Souza, J. Cadeddu, and G.-L. Mariottini. Toward long-term and accurate augmented-reality for monocular endoscopic videos. *Transactions on Biomedical Engineering*, 61(10):2609–2620, Oct. 2014. ISSN 0018-9294.

[138] A. Raabe, J. Beck, S. Rohde, J. Berkefeld, and V. Seifert. Three-dimensional rotational angiography guidance for aneurysm surgery. *Journal of Neurosurgery*, 105(3):406–411, 2006.

[139] H. Rabbani, M. Vafadust, P. Abolmaesumi, and S. Gazor. Speckle noise reduction of medical ultrasound images in complex wavelet domain using mixture priors. *Transactions on Biomedical Engineering*, 55(9):2152–2160, 2008.

[140] A. P. Ramani, M. M. Desai, A. P. Steinberg, C. S. Ng, S. C. Abreu, J. H. Kaouk, A. Finelli, A. C. Novick, and I. S. Gill. Complications of laparoscopic partial nephrectomy in 200 cases. *The Journal of Urology*, 173(1):42–47, 2005.

[141] R. Rigamonti and V. Lepetit. Accurate and efficient linear structure segmentation by leveraging ad hoc features with learned filters. In *Medical Image Computing and Computer-Assisted Intervention*, pages 189–197. Springer, 2012.

[142] K. Říha and R. Beneš. Circle detection in pulsative medical video sequence. In *International Conference on Signal Processing*, pages 674–677. IEEE, 2010.

[143] K. Řiha and I. Potúček. The sequential detection of artery sectional area using optical flow technique. In *International Conference on Circuits, Systems, Eectronics, Control & Signal Processing*, pages 222–226. WSEAS, 2009.

[144] K. Řiha, P. Chen, and D. Fu. Detection of artery section area using artificial immune system algorithm. In *International Conference on Circuits, Systems, Eectronics, Control & Signal Processing*, pages 46–52. WSEAS, 2008.

[145] K. Říha, J. Mašek, R. Burget, R. Beneš, and E. Závodná. Novel method for localization of common carotid artery transverse section in ultrasound images using modified viola-jones detector. *Ultrasound in Medicine & Biology*, 39(10):1887–1902, 2013.

[146] B. I. Rini, S. C. Campbell, and B. Escudier. Renal cell carcinoma. *The Lancet*, 373(9669): 1119–1132, 2009.

[147] P. Risholm, S. Pieper, E. Samset, and W. M. Wells III. Summarizing and visualizing uncertainty in non-rigid registration. In *Medical Image Computing and Computer-Assisted Intervention*, pages 554–561. Springer, 2010.

[148] P. Risholm, J. Balter, and W. M. Wells. Estimation of delivered dose in radiotherapy: the influence of registration uncertainty. In *Medical Image Computing and Computer-Assisted Intervention*, pages 548–555. Springer, 2011.

[149] S. Röhl, S. Bodenstedt, S. Suwelack, H. Kenngott, B. Müller-Stich, R. Dillmann, and S. Speidel. Real-time surface reconstruction from stereo endoscopic images for intraoperative registration. In *SPIE: Medical Imaging*, volume 7964, page 796414, 2011.

[150] S. Röhl, S. Bodenstedt, S. Suwelack, H. Kenngott, B. P. Müller-Stich, R. Dillmann, and S. Speidel. Dense GPU-enhanced surface reconstruction from stereo endoscopic images for intraoperative registration. *Medical Physics*, 39:1632, 2012.

[151] F. Sampaio and M. Passos. Renal arteries: anatomic study for surgical and radiological practice. *Surgical and Radiologic Anatomy*, 14(2):113–117, 1992.

[152] M. Schaap, T. van Walsum, L. Neefjes, C. Metz, E. Capuano, M. de Bruijne, and W. Niessen. Robust shape regression for supervised vessel segmentation and its application to coronary segmentation in CTA. *Transactions on Medical Imaging*, 30(11):1974–1986, 2011.

[153] C. Schneider, J. Guerrero, C. Nguan, R. Rohling, and S. Salcudean. Intra-operative pick-up ultrasound for robot assisted surgery with vessel extraction and registration: a feasibility study. In *Information Processing in Computer-Assisted Interventions*, pages 122–132. Springer, 2011.

[154] C. M. Schneider. *Ultrasound elastography for intra-operative use and renal tissue imaging*. PhD thesis, University of British Columbia, 2017.

[155] N. Sharma and L. M. Aggarwal. Automated medical image segmentation techniques. *Medical Physics*, 35(1):3–14, 2010.

[156] D. Shen, G. Wu, and H.-I. Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 2017.

[157] R. Siegel, J. Ma, Z. Zou, and A. Jemal. Cancer statistics, 2014. *CA: A Cancer Journal for Clinicians*, 64(1):9–29, 2014.

[158] R. L. Siegel, K. D. Miller, and A. Jemal. Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians*, 67(1):7–30, 2017.

[159] A. L. Simpson. *The computation and visualization of uncertainty in surgical navigation*. PhD thesis, Queen's University, 2010.

[160] A. L. Simpson, B. Ma, E. C. Chen, R. E. Ellis, and A. J. Stewart. Using registration uncertainty visualization in a user study of a simple surgical task. In *Medical Image Computing and Computer-Assisted Intervention*, pages 397–404. Springer, 2006.

[161] I. Simpson, M. Cardoso, M. Modat, D. Cash, M. Woolrich, J. Andersson, J. Schnabel, and S. Ourselin. Probabilistic non-linear registration with spatially adaptive regularisation. *Medical Image Analysis*, 26(1):203–216, 2015. ISSN 1361-8415.

[162] I. J. Simpson, M. Woolrich, A. R. Groves, and J. A. Schnabel. Longitudinal brain MRI analysis with uncertain registration. In *Medical Image Computing and Computer-Assisted Intervention*, pages 647–654. Springer, 2011.

[163] I. J. Simpson, M. W. Woolrich, and J. A. Schnabel. Probabilistic segmentation propagation from uncertainty in registration. In *Medical Image Understanding and Analysis*, 2011.

[164] I. J. Simpson, J. A. Schnabel, A. R. Groves, J. L. Andersson, and M. W. Woolrich. Probabilistic inference of regularisation in non-rigid registration. *NeuroImage*, 59(3):2438–2451, 2012.

[165] I. Singh. Robot-assisted laparoscopic partial nephrectomy: current review of the technique and literature. *Journal of Minimal Access Surgery*, 5(4):87, 2009.

[166] R. Singla, P. Edgcumbe, P. Pratt, C. Nguan, and R. Rohling. Intra-operative ultrasound-based augmented reality guidance for laparoscopic surgery. *Healthcare Technology Letters*, 4(5): 204–209, 2017.

[167] A. K. Sinop and L. Grady. A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In *International Conference on Computer Vision*, pages 1–8. Springer, 2007.

[168] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *Transactions on Medical Imaging*, 25(9):1214–1222, 2006.

[169] A. Sotiras, N. Komodakis, B. Glocker, J.-F. Deux, and N. Paragios. Graphical models and deformable diffeomorphic population registration using global and local metrics. In *Medical Image Computing and Computer-Assisted Intervention*, pages 672–679. Springer, 2009.

[170] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: a survey. *Transactions on Medical Imaging*, 32(7):1153–1190, 2013.

[171] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge-based vessel segmentation in color images of the retina. *Transactions on Medical Imaging*, 23(4): 501–509, 2004.

[172] M. Steffens, D. Aufderheide, S. Kieneke, W. Krybus, C. Kohring, and D. Morton. Probabilistic scene analysis for robust stereo correspondence. In *Image Analysis and Recognition*, pages 697–706. Springer, 2009.

[173] D. Stoianovici, K. Cleary, A. Patriciu, D. Mazilu, A. Stanimir, N. Craciunoiu, V. Watson, and L. Kavoussi. AcuBot: a robot for radiological interventions. *Transactions on Robotics and Automation*, 19(5):927–930, 2003.

[174] D. Stoyanov. Stereoscopic scene flow for robotic assisted minimally invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention*, pages 479–486. Springer, 2012.

[175] D. Stoyanov, M. V. Scarzanella, P. Pratt, and G.-Z. Yang. Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention*, pages 275–282. Springer, 2010.

[176] G. Strau, K. Koulechov, S. Röttger, J. Bahner, C. Trantakis, M. Hofer, W. Korb, O. Burgert, J. Meixensberger, D. Manzey, A. Dietz, and T. Lüth. Evaluation of a navigation system for ENT with surgical efficiency criteria. *The Laryngoscope*, 116(4):564–572, 2006.

[177] L.-M. Su, B. P. Vagvolgyi, R. Agarwal, C. E. Reiley, R. H. Taylor, and G. D. Hager. Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3D-CT to stereoscopic video registration. *Urology*, 73(4):896–900, 2009.

[178] D. Sun, S. Roth, and M. J. Black. Secrets of optical flow estimation and their principles. In *Computer Vision and Pattern Recognition*, pages 2432–2439. IEEE, 2010.

[179] J. S. Suri, K. Liu, L. Reden, and S. Laxminarayan. A review on MR vascular image processing: skeleton versus nonskeleton approaches: part II. *Transactions on Information Technology in Biomedicine*, 6(4):338–350, 2002.

[180] E. A. Tanagho. Embryology of the genitourinary system. *General Urology*, pages 18–30, 2004.

[181] L. Tang and G. Hamarneh. Medical image registration: a review. *Medical Imaging: Technology and Applications*, pages 619–660, 2013.

[182] L. Y. Tang and G. Hamarneh. Random walks with efficient search and contextually adapted image similarity for deformable registration. In *Medical Image Computing and Computer-Assisted Intervention*, pages 43–50. Springer, 2013.

[183] H. A. Taylor, T. T. Brunyé, and S. T. Taylor. Spatial mental representation: implications for navigation system design. *Reviews of Human Factors and Ergonomics*, 4(1):1–40, 2008.

[184] R. H. Taylor and D. Stoianovici. Medical robotics in computer-integrated surgery. *Robotics and Automation*, 19(5):765–781, 2003.

[185] D. Teber, S. Guven, T. Simpfendörfer, M. Baumhauer, E. O. Güven, F. Yencilek, A. S. Gözen, and J. Rassweiler. Augmented reality: a new tool to improve surgical accuracy during laparoscopic partial nephrectomy? Preliminary in vitro and in vivo results. *European Urology*, 56(2):332–338, 2009.

[186] D. Tilly, N. Tilly, and A. Ahnesjö. Dose mapping sensitivity to deformable registration uncertainties in fractionated radiotherapy–applied to prostate proton treatments. *BMC medical physics*, 13(1):2, 2013.

[187] S. Tobis, J. Knopf, C. Silvers, J. Yao, H. Rashid, G. Wu, and D. Golijanin. Near infrared fluorescence imaging with robotic assisted laparoscopic partial nephrectomy: initial clinical experience for renal cortical tumors. *The Journal of Urology*, 186(1):47–52, 2011.

[188] N. Toussaint, J.-C. Souplet, and P. Fillard. MedINRIA: medical image navigation and research tool by INRIA. In *MICCAI Workshop on Interaction in Medical Image Analysis and Visualization*, page 280. CiteSeer, 2007.

[189] S. Tovar-Arriaga, R. Tita, J. C. Pedraza-Ortega, E. Gorrostieta, and W. A. Kalender. Development of a robotic FD-CT-guided navigation system for needle placement—preliminary accuracy tests. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 7(2):225–236, 2011.

[190] W. Tsai and Ö. Savaş. Flow pumping system for physiological waveforms. *Medical & Biological Engineering & Computing*, 48(2):197–201, 2010.

[191] J. Udupa and G. Grevera. Go digital, go fuzzy. *Pattern Recognition Letters*, 23(6):743–754, 2002.

[192] B. A. Urban, L. E. Ratner, and E. K. Fishman. Three-dimensional volume-rendered CT angiography of the renal arteries and veins: normal anatomy, variants, and clinical applications. *RadioGraphics*, 21(2):373–386, 2001.

[193] K. A. Vermeer, F. M. Vos, H. Lemij, and A. M. Vossepoel. A model based method for retinal blood vessel detection. *Computers in Biology and Medicine*, 34(3):209–219, 2004.

[194] C. Wachinger, T. Klein, and N. Navab. The 2D analytic signal on RF and B-mode ultrasound images. In *Information Processing in Medical Imaging*, pages 359–370, 2011.

[195] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman. Phase-based video motion processing. *Transactions on Graphics*, 32(4):80, 2013.

[196] Q. Wang, M. Kim, Y. Shi, G. Wu, and D. Shen. Predict brain MR image registration via sparse learning of appearance and transformation. *Medical Image Analysis*, 20(1):61–75, 2015.

[197] S. Warfield, K. Zou, and W. Wells. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *Transactions on Medical Imaging*, 23(7):903–921, 2004.

[198] R. K. Warriner, K. W. Johnston, and R. S. Cobbold. A viscoelastic model of arterial wall motion in pulsatile flow: implications for Doppler ultrasound clutter assessment. *Physiological Measurement*, 29(2):157, 2008.

[199] T. Watanabe and C. Scott. Spatial confidence regions for quantifying and visualizing registration uncertainty. In *Biomedical Image Registration*, pages 120–130. Springer, 2012.

[200] L. Wietzke, G. Sommer, and O. Fleischmann. The geometry of 2D image signals. In *Computer Vision and Pattern Recognition*, pages 1690–1697. IEEE, 2009.

[201] T. Y. Wong, R. Klein, D. J. Couper, L. S. Cooper, E. Shahar, L. D. Hubbard, M. R. Wofford, and A. R. Sharrett. Retinal microvascular abnormalities and incident stroke: the atherosclerosis risk in communities study. *The Lancet*, 358(9288):1134–1140, 2001.

[202] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *Transactions on Graphics*, 31(4):65, 2012.

[203] Z. Yaniv and K. Cleary. Image-guided procedures: a review. *Computer Aided Interventions and Medical Robotics*, 3:1–63, 2006.

[204] İ. Ş. Yetik and A. Nehorai. Performance bounds on image registration. *Transactions on Signal Processing*, 54(5):1737–1749, 2006.

[205] A. Yezzi, L. Zollei, and T. Kapur. A variational framework for joint segmentation and registration. In *Mathematical Methods in Biomedical Image Analysis*, pages 44–51. IEEE, 2001.

[206] M. Yip, T. Adebar, R. Rohling, S. Salcudean, and C. Nguan. 3D ultrasound to stereoscopic camera registration through an air-tissue boundary. In *Medical Image Computing and Computer-Assisted Intervention*, pages 626–634. Springer, 2010.

[207] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*, 31(3):1116–1128, 2006.

[208] R. Zahiri-Azar and S. E. Salcudean. Motion estimation in ultrasound images using time domain cross correlation with prior estimates. *Transactions on Biomedical Engineering*, 53(10): 1990–2000, 2006.

[209] R. Zakariaee, G. Hamarneh, C. Brown, M. Gaudet, C. Aquino-Parsons, and I. Spadinger. Bladder accumulated dose in image-guided high-dose-rate brachytherapy for locally advanced cervical cancer and its relation to urinary toxicity. *Physics in Medicine and Biology*, 61: 8395–8411, 2016.

[210] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *Transactions on Medical Imaging*, 20(1):45–57, 2001.