Intra-operative Ultrasound-based Augmented Reality for Laparoscopic Surgical Guidance

by

Rohit Kumar Singla

B.A.Sc, The University of British Columbia, 2015

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

Master of Applied Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Biomedical Engineering)

The University of British Columbia
(Vancouver)

July 2017

Abstract

Laparoscopic partial nephrectomy involves the complete resection of a kidney tumour, while minimizing healthy tissue excised, and under a time constraint before irreparable kidney damage occurs. The surgeon must complete this operation in a reduced sensory environment with poor depth perception, limited field of view, and little or no haptic feedback. For endophytic tumours (grows inwards), this is particularly difficult. In order to assist the surgeon, augmented reality can provide intra-operative guidance. Intra-operative ultrasound is low cost, non-ionising, and real-time. This has tremendous potential to guide the surgeon. This thesis details the development of three intra-operative augmented reality systems from a single framework, with augmentations all based on intra-operative ultrasound. The systems were all developed on the da Vinci Surgical System®, using it as a development and testing platform. All systems leverage a single fiducial marker called the Dynamic Augmented Reality Tracker which can track the local surface and create a tumour-centric paradigm. A 3D ultrasound volume is reconstructed using a tracked ultrasound transducer. A tumour model is then extracted via manual segmentation of the volume. The three systems were developed and evaluated in simulated robot-assisted partial nephrectomies. The first system shows the feasibility of providing continuous ultrasound-based guidance during excision and achieves a system error of 5.1 mm RMS. Improving on this, the second system demonstrates clinically acceptable system error of 2.5 $\pm$ 0.5 mm. The second system significantly reduced healthy tissue excised from an average of 30.6 $\pm$ 5.5 cm^3 to 17.5 $\pm$ 2.4 cm^3 ($p < 0.05$) and reduced the depth from the tumor underside to cut from an average of 10.2 $\pm$ 4.1 mm to 3.3 $\pm$ 2.3 mm ($p < 0.05$). The third system is a novel intra-corporeal projector-based system that assists in determining the initial angle

of resection. This system is evaluated in a surgeon study with a total of 32 simulated operations and addresses the limitations of conventional augmentations from the laparoscope's point of view. All three systems show their potential benefits in improving laparoscopic surgery with minimal additional hardware. With such image-guidance systems, the widespread adoption of laparoscopic surgery can be facilitated, improving patient care.

Lay Summary

Minimally invasive surgery is rapidly becoming the standard of care for many diseases, including partial nephrectomy (the excision of only the tumour in kidney cancer operations). However, the nature of these surgeries requires the surgeon to operate with the limitations such as reduced field of view, poor depth perception and little or no sense of touch. To overcome these challenges, this thesis proposes the development of three systems based on ultrasound imaging and augmented reality. Each system presents a unique set of augmented reality overlays from a tracked ultrasound scan using computer vision. Each system is evaluated in mock robot-assisted partial nephrectomies performed by an expert surgeon. The results indicate the systems have clinically acceptable error and can significantly reduce the amount of healthy tissue excised. This work can improve on and facilitate the widespread adoption of laparoscopic surgery, broadly benefiting patients in numerous surgeries.

Preface

This thesis is primarily based on three manuscripts, one of which has been published and the other two which are pending. The manuscripts have been modified and integrated for coherency. This work is been the result of an inter-disciplinary and inter-institutional collaboration between the University of British Columbia's Department of Electrical and Computer Engineering, the University of British Columbia's Department of Urological Sciences, Imperial College London's Department of Surgery and Cancer, and Northern Digital Inc.

A modified version of Chapter 3 has been published, where the author is joint first author (denoted by asterisk), as follows:

- Philip Edgcumbe*, Rohit Singla*, Philip Pratt, Caitlin Schneider, Christopher Nguan, and Robert Rohling. "Augmented Reality Imaging for Robot-Assisted Partial Nephrectomy Surgery". In International Conference on Medical Imaging and Virtual Reality, pp. 139-150. Springer International Publishing, 2016.

A modified version of Chapter 3 and Chapter 4 has been submitted to and was accepted in a Special Issue on Augmented Environments for Computer-Assisted Interventions in IET's Healthcare Technology Letters. A modified abstract was submitted to the 11^{th} Annual Lorne D. Sullivan Lectureship and Research Day and was accepted as a podium presentation. It received the Best Clinical Sciences Research Award. The author list and title are as follows:

- Rohit Singla, Philip Edgcumbe, Philip Pratt, Christopher Nguan, and Robert Rohling. Intuitive Intra-operative Ultrasound-based Augmented Reality Guidance for Robot-Assisted Laparoscopic Surgery.

A modified abstract of Chapter 5 was submitted to the 11th Annual Lorne D. Sullivan Lectureship and Research Day and was accepted as a poster presentation. The author list and title are as follows (presenting author denoted by asterisk):

- Philip Edgcumbe, Rohit Singla*, Philip Pratt, Christopher Nguan, and Robert Rohling. Follow the Light: Intra-corporeal Projector-based Augmented Reality for Laparoscopic Surgery.

The author's technical contributions include the design and implementation of the software components for the systems. The author collaborated with Dr. Philip Pratt to create a plug-in framework for Dr. Pratt's software. From there, the author implemented modules to do the following: interface with the Analogic ultrasound machines; track one or multiple fiducial markers; track the projector in real-time and evaluate its projection accuracy; render the virtual viewpoints; render the augmented surgical instruments; render using perspective and orthographic projections; render models as point clouds, convex hulls and more; display augmentations in the projector point-of-view and laparoscope point-of-view; display augmentations via a monitor or projector. Furthermore, the author added to the mathematical framework to provide tumour-centric tracking over time and continuously render models as seen by the virtual cameras. The author developed the qualitative metrics and analyzed the results. Other contributions of the author include: characterization of the latency; reconfiguration of the fiducial marker of the projector; and manual segmentation. Finally, the author led writing of manuscripts in Chapter 4; contributed to editing of the manuscripts in Chapter 3 and Chapter 5; and created supplemental videos for all chapters.

Philip Edgcumbe developed and tested the Dynamic Augmented Reality Tracker. This included the initial design, any modifications and the finite element modeling (FEM) simulations. He performed the geometric ultrasound calibration, robot to camera calibration for Chapter 3, and made the phantom tumour models used in all experiments. He further generated the tumour models used for Chapter 3, and had contributed the idea of orthogonal views. For the systems in Chapter 3 and Chapter 4, Philip Edgcumbe's contributions to the theoretical design include the transformation equations for tracking surgical instruments relative to the DART, and determining the pose of the virtual cameras. For Chapter 5, Philip Edgcumbe

created and developed the prototype for the Pico Lantern. He evaluated surface reconstruction accuracy, and performed verification and validation experiments.

Dr. Andrew Wiles provided support in the development of the projector-based work. His research team completed the surface reconstruction accuracy and speed testing. Dr. Philip Pratt provided technical support and guidance, expanded the interface of his software as needed, and contributed to the manuscripts. Prof. Robert Rohling and Dr. Christopher Nguan provided technical and clinical guidance, and contributed to the manuscripts.

Table of Contents

List of Tables

List of Figures

Glossary

1D one dimensional

2D two dimensional

3D three dimensional

6-DOF 6 degrees-of-freedom

13-DOF 13 degrees-of-freedom

API application programming interface

B-MODE brightness mode

BM block matching on the CPU

BMGPU block matching on the GPU

BPGPU belief propagation on GPU

CSBPGPU constant space belief propagation on GPU

CT computed tomography

DART Dynamic Augmented Reality Tracker

EM electro-magnetic

FEM finite element modeling

FRE fiducial registration error

GFR glomerular filtration rate

GPU graphics processing unit

HD high-definition

IGS image-guided surgery

LPOV laparoscope point-of-view

MIS minimally invasive surgery

MRI magnetic resonance imaging

NGUAN Nephrectomy Guidance using Ultrasound-Augmented Navigation

NGUAN+ Nephrectomy Guidance Using Ultrasound-Augmented Navigation 2.0

PARIS Projector-based Augmented Reality Intra-corporeal System

PNP Perspective-N-Point Problem

POV point of view

PPOV projector point of view

PVC polyvinyl chloride

RAPN robot-assisted partial nephrectomy

RCC renal cell carcinoma

RMS root mean square

SDI serial digital interface

SGBM semi-global block matching on CPU

TRE target registration error

US ultrasound

VGH Vancouver General Hospital

Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my supervisor, Prof. Robert Rohling, for letting me explore the fields of augmented reality, medical imaging, surgical robotics, and cancer surgery — all in the same project. Prof. Rohling has allowed me to use one of the most technologically advanced tools in the da Vinci Surgical System, improve it using one of (if not the) greatest medical imaging modalities in ultrasound, and develop my own contribution of the next major computing modality in augmented reality. Prof. Rohling has tremendous patience and his expertise was a major factor of how this work was even possible. Thank you for your outstanding technical and emotional intelligence. Thank you for being an outstanding role model who continues to inspire me. Thank you for always having my back and letting me roam — whether it be to London, or Switzerland, or back to my desk.

Thank you to Dr. Christopher Nguan, the clinical champion, for being beyond open-minded to some hair-brained schemes and even pitching some of your own. The passion you have for improving surgery through technology is infectious. Thank you for your occasional 1:00AM emails. Thank you for your patience and belief. Thank you for being the namesake of this work.

Thank you to Prof. Purang Abolmaesumi for his never-ending support. I am continuously humbled by your brilliance and delighted by your uplifting laughter!

I am grateful for the generous funding from the Natural Sciences and Engineering Research Council (NSERC) through the Canada Graduate Scholarship (Master's Level) and the Collaborative Research and Training Experience (CREATE) program; the Engineers in Scrubs Travel Grant Award; and the UBC Graduate Student Initiative.

I would also like to thank Andrew Wiles and the Advanced Research Team at Northern Digital Inc. for their continued support and feedback. I would further like to thank Prof. Tim Salcudean for infrastructure and support.

Thank you to my lab mates, past and present, who made my time in the lab fun and productive: Alexander Seitel, Angelica Ruskowksi, Corey Kelly, Gregory Allan, Irene Tong, Jeffrey Abeysekra, Jorden Hetherington, Julie Hemily, Julio Lobo, Mehran Petsie, Mohammad Honaravar, Mohammad Najafi, Nathan van Woudeberg, Omid Mohaeriri, Qi Zeng, Saman Nourarian, Samira Soujoudi, Shekoofeh Azizi, Tom Curran, and Yasmin Halwani. Lab mates like you make our lab one of the best in the world.

Thank you to the friends I have met in my graduate school journey for their clever insights and cleverer jokes: Brendan Gribbons, Cameron Stuart, Charlene Leung, Claire da Roza, Jason Spiedel, Juan Pablo Gomez Arrunategui, Liam Sharkey, Luke Haliburton, Michele Touchette, Olivia Paserin, Prashant Pandey, Sampath Satti, Dr. Su Lin Lee, and Tiffany Ngo.

Thank you to the friends beyond academia that kept me balanced whenever I needed it: Amit Anand, Christopher Tan, Elizabeth Wicks, Emily Woehrle, Gorden Larson, James Mackenzie, Jeremy Lord, Jessica Stewart, Lauren Fung, Lucas and Andrea Cahill, Marc Lejay, Michael Ip, Sagar Malhi, Sarah Holdijk, and Vanessa Russell.

To three of the greatest mentors I've ever had: Dr. Caitlin Schneider, Dr. Philip Pratt, and Dr. Philip Edgcumbe - it is beyond my wildest dreams that I ever could have worked with experts in the field such as yourselves. I've been given the privilege to learn from the best, the brightest - and most importantly - the humblest trio. I am honoured to work with all of you each and every day.

Finally, endless thanks to my mom and dad, Renu and Krishan Singla, and two sisters, Rubina and Krishma. It is with you that I get to take a step away from my world and go back to being a kid again. Thank you for calling when you haven't heard from me in a while. Thank you for picking me up and taking me home. Thank you for making me endless cups of chai. Thank you for your love, sacrifices, unwavering support, jokes, and everything - big and small - that you've ever done so that I can live in the city I love, do the work I do, and have the life that I have.

Chapter 1

Introduction

Thus (through perspective) every sort of confusion is revealed within us; and this is that weakness of the human mind on which the art of conjuring and of deceiving by light and shadow and other ingenious devices imposes, having an effect upon us like magic... And the arts of measuring and numbering and weighing come to the rescue of the human understanding there is the beauty of them and the apparent greater or less, or more or heavier, no longer have the mastery over us, but give way before calculation and measure and weight?
— Plato (380 B.C.)

Little did he know, when Wilhelm Röentgen took the first X-ray image of his wife's hand in 1895, he would start a revolution in medical imaging and surgery. Since the creation of radiology, imaging modalities such as X-rays, computed tomography (CT), and ultrasound (US) have drastically improved the manner in which surgery is performed. Röentgen's would lead to development of interventional radiology; combined with the development of endoscopic imaging it would change the cut-and-see approach of the past to the see-then-cut approach seen today [54]. With the use of imaging prior to the operation (pre-operative) and later during the operation (intra-operative), surgeons gained unprecedented abilities to diagnose and understand the human anatomy and the pathologies that their patients faced [54]. With additional advances in hardware and mathematics, the field of image-guided therapy was born. However, despite this enhanced skill set, surgeons face numerous challenges in performing operations, particularly when operating on soft

tissues in the abdomen.

This thesis focuses on the creation of new guidance systems with applications to laparoscopic surgery, specifically the robot-assisted partial nephrectomy (RAPN). In order to assist the surgeon further, with no additional harm to the patient, this thesis presents three systems based on ultrasound (US) and computer vision. Each system provides the surgeon unique augmentations that enhance what he or she can see and how he or she can operate, all with the motivation of improving patient care. While this thesis is applied to the RAPN, its potential is large and can be extended to other organs like the liver, prostate, pancreas and more.

Chapter 1 is organised as follows:

1. Section 1.1 outlines the evolution of minimally invasive surgery,
2. Section 1.2 discusses the development of robot-assisted minimally invasive surgery,
3. Section 1.3 provides a brief overview of key concepts in image guided surgery,
4. Section 1.4 breaks down the objectives of this thesis, and finally
5. Section 1.5 gives a chapter-by-chapter overview of the thesis as a whole.

1.1 Minimally Invasive Surgery

The field of minimally invasive surgery (MIS) has become a relatively standard approach for various abdominal surgeries. Here, rather than one long morbid incision in the patient, as in Figure 1.1, a set of small incisions are made instead. These small incisions of 0.5 - 1.5 cm in length, give this type of surgery the nickname, ”keyhole surgery”. The surgeon inserts a camera, known as the laparoscope, and surgical instruments through the various port sites into the abdomen where a working space has been generated through the insufflation of carbon dioxide gas. The surgeon operates using long, rigid instruments like in Figure 1.2. With MIS, the patient benefits from shorter recovery time, less post-operative pain, less intra-operative blood loss, and better cosmesis.

The cost for MIS is at the expense of the surgeon who must now operate in a reduced sensory environment. In regards to visualizing the surgical scene, a

Figure 1.1: Example of the incision required in open surgery, as seen on a porcine model.

surgical assistant must hold a monocular laparoscope inserted in the patient. The video feed is displayed on a monitor. This combination presents a reduction in the field of view and poor depth perception, with added instability from the manual manipulation of the laparoscope. Stereoscopic laparoscopes, which offer partial recovery of depth perception, are commercially available from vendors including Olympus and Intuitive Surgical, but are not yet ubiquitous. The laparoscope's lack of depth resolution influences the surgeon's spatial understanding of the scene and the relative positioning of structures to one another.

With regards to the laparoscopic tools, the surgeon's haptic sense is impaired and his or her movement becomes unintuitive. The latter comes because of the fulcrum effect, where inversion is required in moving an instrument by its handle, rather than by its tip. Essentially, when the surgeon moves his or her hand right, the instrument moves left instead of right. As well, these instruments are

Figure 1.2: Example of the long and rigid laparoscopic instruments used in minimally invasive surgery.

not articulated (wristed), producing dexterity challenges. All of these constraints on how the surgeon sees, feels and thinks lead to increased operation time and increased surgical errors. Tasks requiring fine motor skills or complex manipulation are more difficult. To mitigate these challenges, the field of robot-assisted MIS would emerge.

1.2 Robot-Assisted Minimally Invasive Surgery

Over 30 years ago, research into robotic-assisted surgery began, yielding novel systems like the ROBODOC released by Integrated Surgical Systems in 1992. Since that time, the da Vinci Surgical System® (Intuitive Surgical Inc., Sunnyvale, CA, USA) has become one of the most successful robot-assisted systems used worldwide. In 2015 alone, there were an estimated 3600 systems completing 650,000 procedures including gynaecology and urology. In this thesis, the term "robot" refers to the da Vinci Surgical System®, which is treated as the exemplar system herein. With robot-assisted surgery, the surgeon is in control at all times, in a master-slave configuration. The robot extends but does not replace the surgeon's

Figure 1.3: The da Vinci S Surgical System®. The surgeon's console (left), the patient-side cart (middle), and vision cart (right). ©2017 Intuitive Surgical, Inc.

abilities and role and has no autonomous ability. The surgeon, through the use of a robot, regains some of his or her senses which were reduced when the field moved to laparoscopic surgery.

With the da Vinci® (Figure 1.3), the operator sits at a surgeons console with two "master"-side manipulators, four foot pedals and a console viewer with a three dimensional (3D) screen instead of a traditional two dimensional (2D) one. The da Vinci's laparoscope has a pair of stereo high-definition (HD) cameras, allowing for the 3D visualization. The console also permits integration of additional digital data directly into the surgeons console through the use of the TilePro® function. These could be in the form of preoperative and intra-operative imaging modalities such as computed tomography (CT), US, or in the form of navigation and guidance tools.

The console is connected to a vision cart and patient cart. The patient cart is located at the patient bedside which has a center column with four robotic patient-

side manipulators hanging from it. The manipulators are docked to specialised ports inserted into the same incisions in the insufflated abdomen as in traditional MIS. Each manipulator can operate a unique instrument such as small 8 mm diameter scissors or electrocautery tools. The master-side manipulators' motion is translated to the patient-side manipulators. The da Vinci® is referred to as a teleoperated robotic system because of this master to patient manipulator mapping. In addition to real-time teleoperation, the movement is filtered to minimize natural hand tremor and the motion can be variably scaled to allow for fine movements. The articulated tools themselves give the surgeon back a degree of freedom he or she lost with in conventional laparoscopic surgery. The da Vinci® allows for intuitive movement of the tools, removing the fulcrum effect from the list of challenges the surgeon must face. By design, the robotic manipulators can match the surgeons range of motion in open surgery. These improvements allow for more complex minimally invasive procedures and also simplify routine laparoscopic operations. Additionally, the da Vinci's® ability to localise points in space has been quantified to be 1 mm [37]. This means the da Vinci® is suitable for use in image-guidance systems, with high precision of instrument tracking [37]. It has further been associated with a reduction in mental effort and workload in comparison to traditional laparoscopic tasks [49]. Studies indicate there is likely a benefit to using the robot-assisted approach over conventional laparoscopy to achieve better margin sizes and post-operative function in partial nephrectomies [11, 55, 64].

Nevertheless, there are disadvantages to robot-assisted surgery. With the da Vinci®, a prominent example is the fact haptic feedback is entirely absent, not just limited as in conventional laparoscopy. This lack of tactile feedback increases the risk of excessive force applied or the clashing of instruments. As well, the da Vinci® requires significant investment upfront, so robot-assisted procedures often cost more to perform. Surgeons and staff must undergo specific training for use of the robot. Overall, the da Vinci® remains a promising avenue for surgery. From a research perspective, it is an excellent development and testing platform. Integrating image guidance can expand the existing benefits of the da Vinci®, especially for the surgeon.

1.3 Image Guided Surgery

With the advent of computers, the field of medicine was revolutionised. This onset of computational power has led to improvements to various parts of medicine, but arguably none more than radiology and surgery. Due to the advances in the imaging modalities available, clinical decision making has been drastically improved. By being able to see inside the patient, with no excision, and understand the underlying anatomy, the choice to operate or not is better informed. When surgery is chosen, pre-operative imaging is vital in developing a patient-specific plan [54]. Through the use of X-ray, CT, magnetic resonance imaging (MRI), or US, the surgeon can understand what the internal structure looks like in each patient. Intra-operatively, through the use of fluoroscopy, cone-beam CT, US and others, the surgeon can see within and understand the nuances of anatomy in real-time without needing to see the target with his or her own eyes. Medical imaging enhances the surgeon's sight and the visualization enhances the surgeon's reasoning [84]. The field of using such powerful imaging to assist and navigate during surgery is called image-guided surgery (IGS) [53].

Image-guidance can tackle some of the challenges in MIS and to a degree, those in open surgery as well. For example, it can mitigate for the loss of haptic feedback to subsurface patient anatomy [53]. The standard video feed can be complemented by the use of intra-operative imaging, pre-operative imaging, or a combination of both [53]. Tracking surgical tools or anatomical structures, and aligning them in a common coordinate system with the imaging can be beneficial [53]. Through this, the surgeon can understand where he or she (i.e. the instruments) are spatially in relation to the target. IGS can then lead to better clinical decisions, which in turn may lead to fewer complications, less blood loss, less tissue excised, and prevent disorientation [6]. IGS can also reduce the cognitive load on the surgeon [6, 54]. Traditionally, the surgeon must observe and mentally connect what he or she has seen pre-operatively to what he or she view intra-operatively; a challenging task given the existing environmental stressors such as time [54, 84].

The importance and value of image-guidance is reflected in the development of the Advanced Multi-modality Image Guided Operating (AMIGO) suite [34]. AMIGO is a state-of-the-art integrated three room design dedicated to allow the

use of MRI, CT, US and an array of additional imaging modalities intra-operatively. However, IGS's benefits are limited if not correctly implemented. IGS itself encompasses several technical components, including but not limited to imaging, tracking, registration, and display [53]. These are briefly discussed as follows:

- **Imaging**: The choice of imaging modality is guided by the desired application. MRI, for instance, provides excellent and detailed 3D tomographic images, with excellent contrast of various soft tissue. However, it is not real-time, is subject to motion artefacts and is difficult to use intra-operatively. CT and X-rays are beneficial because they can be used intra-operatively and potentially in real-time, but both introduce ionising radiation. US is a possible option both intra-operatively and pre-operatively. It is both real-time and non-ionising but has variable contrast, noise and resolution, and the image quality is heavily dependent on the user. US can also not image beyond areas of high acoustic impedance, and has relatively low penetration. Additionally, there is the endoscopic image itself [6]. As this provides an intra-operative view of the scene, other modalities may be registered to it, or the view itself can provide guidance information.

  No single modality is superior to another. However, the choice of modality will influence the accuracy and use of the IGS system. Take, for example, an intra-operative US of a subsurface target. While relatively safe for the patient, the image is often difficult to interpret. The segmentation of the target in a given image may be particularly difficult, and a poor segmentation will limit the entire system's accuracy. One must be careful in the choice of modality as no one imaging modality is best for all phases of a procedure [54].

- **Tracking**: In diagnostic and therapeutic procedures, the surgeon uses instruments to observe and manipulate the scene. To integrate these instruments with the imaging information, they must be first tracked. In doing so, they can be brought into a common coordinate system. Several tracking methods exist including optical tracking, electro-magnetic (EM) tracking, and computer-vision based tracking, described below.

  - *Optical*: Optical tracking refers to the use of infrared light to illuminate

reflective or active markers, analyse the illuminated image with a camera, and then localise the markers relative to the tracker [53]. This can be done by controlling when the markers are illuminated (active) or not illuminated (passive). While this is has been shown to be highly accurate and precise, it requires a direct line of sight with the markers [53]. In the case of laparoscopic surgery, these markers are frequently outside of the patient and located on the proximal ends of the instruments. Placing the markers far from the distal instrument tip may introduce additional errors in tracking [53]. Common systems include the Polaris (Northern Digital Inc., Waterloo, ON, CA) and Certus OptoTrak (Northern Digital Inc., Waterloo, ON, CA).

- *Electromagnetic*: EM tracking uses a field generator to create an electromagnetic field, in which sensor coils on the instruments are tracked [53]. This eliminates the line-of-sight issue of optical tracking with similar accuracy to optical tracking [53]. However, the presence of ferromagnetic material within the operating room will cause distortions to the EM field, resulting in non-uniform accuracy and noise [53]. The presence of such material is quite likely. Common systems used are the Aurora (Northern Digital Inc., Waterloo, ON, CA) and the Ascension Bird (Ascension Technology Corp, Shelburne, VT, USA).
- *Computer-vision*: Computer-vision based tracking (also called image based) analyses the laparoscopic image to track organs and tools. This may be through the use of a single 2D laparoscopic image, or a stereoscopic 3D pair. It may even be through analysis of medical images themselves (ex. in US-guided needle interventions). This has potential in the tracking of soft organs and the instruments with no exogenous hardware added, but its robustness and accuracy for use in clinical practice remains to be proven. Challenges include deformable objects, foreshortening, occlusions, and the need for concurrency.

Similar to the choice of imaging modality, there is no single superior tracking method. The best approach is likely a combination of each to balance benefits and drawbacks.

- **Registration**: in order to be of use, imaging data and tracking data must be combined together. The process of bringing these two data sets together in a common coordinate system, such that a point in one set and its equivalent in another set is known, is referred to as registration [53]. A rigid registration only requires a rotation and translation between coordinate systems, while a non-rigid registration requires additional parameters [53]. Registration may occur between a 3D dataset to another 3D dataset, 2D to 2D, or 2D to 3D and vice versa. Registration may also involve two different or imaging modalities (pre-operative/intra-operative or intra-operative/intra-operative). Regardless of the method used, the end outcome should be an accurate alignment that is confirmed using a validated evaluation method [53].

- **Visualization**: the display of registered tracked tools and imaging data is perhaps one of the most significant barriers to broad adoption of IGS[53]. Regardless of all the complexities involved in the other aspects, if the end result of an IGS system cannot be easily understood, then it is hard to envision any benefit. Questions to consider in designing visualizations include whether it includes one or more registered sets and modes of data; whether it augments the view or supplements it; whether the imaging is in 2D or 3D; and whether it introduces additional mental load. Visualization is a significant challenge, but achieving an intuitive visualization is how IGS can be accepted and utilised at scale. In this aspect, augmented and virtual reality displays may play a role [53]. Visualization which mimics the real world, showing virtual images with correct context, is the key to effective IGS[53]. Common methods of rendering include surface rendering (a set of polygons, often triangles, which form a mesh) or volume rendering (visualization of the entire dataset, visualized through ray casting) [6].

Furthermore, the integration of all these aspects into one unified IGS system presents a significant challenge in itself. One must consider costs, practical implementation, physical requirements, accuracy, usability, and clinical utility.

1.4 Thesis Objectives

The primary goal of this thesis is to develop novel intra-operative image guidance systems using US and augmented reality. It does so by presenting three systems that share a common framework and principle of operation. The systems are called NGUAN, NGUAN+, and PARIS. These systems undergo evaluation of their feasibility in improving RAPN with the hypothesis that measurable quality metrics of the surgery will be improved. As part of this goal, the following objectives must be met:

- the integration of a US machine, the da Vinci®, the laparoscopic video feed, and additional components as necessary into a unified framework such that image guidance is possible.
- a method to register US information to a surgical scene despite tissue deformation and organ movement.
- a method to provide continuous and real-time guidance with as few constraints on the surgeon as possible.
- the development of augmented and virtual reality visualizations to address specific surgical challenges of the RAPN.
- the thorough evaluation of the overall system and its components for clinically acceptable accuracy.
- the design of a user study or studies to evaluate the utility of the developed systems in a clinical context.

Achieving these will illustrate the feasibility of creating an IGS system using components that are low in cost and can be broadly distributed. Such a system would address existing gaps in present options, and reduce the barrier to performing laparoscopic surgeries.

1.5 Thesis Overview

This thesis is structured as follows:

- Chapter 1 provides an overview of minimally invasive surgery and image-guided surgery, describes the motivation of this thesis, and presents the thesis objectives.
- Chapter 2 presents an overview of the kidney, renal cell carcinoma, and the nephrectomy procedures; describes the use of intra-operative US; and reviews the prior work in the field of image-guided surgery.
- Chapter 3 presents the overall framework used to develop the systems in this thesis; outlines NGUAN and the evaluation and limitations of it.
- Chapter 4 presents NGUAN+, with improved accuracy that is clinically acceptable and has intuitive visualizations; and evaluates the new system for its utility and limitations.
- Chapter 5 presents and evaluates PARIS, the third augmented reality system, which uses a novel projection-based intra-corporeal approach, addressing an unmet challenge from the first two systems.
- Chapter 6 concludes the thesis with a summary of the work done, contributions, highlights limitations, and discusses potential avenues for future work.

Chapter 2

Background and Related Work

2.1 The Kidney

2.1.1 Anatomy and Physiology

The kidney is a vital organ in the human body. It is a bean shaped organ that, when fully developed in an adult, is approximately 13 cm $\times$ 5 cm $\times$ 2 cm in size, or approximately the size of a fist [59]. The normal human has a pair of kidneys which are located in the posterior of the abdominal cavity, and caudad to the diaphragm and the liver (the upper back side of the abdomen) [59].

A kidney has a fairly complex structure. The kidney itself is encapsulated in a layer of fascia, perirenal fat, and the renal capsule [59]. This covers the renal cortex, the outermost part of the kidney itself. The cortex is smooth and appears red in colour [59]. The cortex goes from the renal capsule to the base of the renal pyramids. The cortex and renal pyramid bases together make up the kidney's parenchyma. Beneath the cortex lies the renal medulla layer, which appears red-brown in colour [59]. The medulla contains the renal pyramids themselves, which are oriented with an apex inwards to the center of the kidney. The renal pyramids are formed by an aggregation of nephrons and tubules. Within the medulla lies the collecting duct system, composed of minor and major calyxes [59]. Urine passes through from the collecting duct into the renal pelvis and finally into the ureter which leads to the bladder [59]. This layer consists of millions of the kidney's

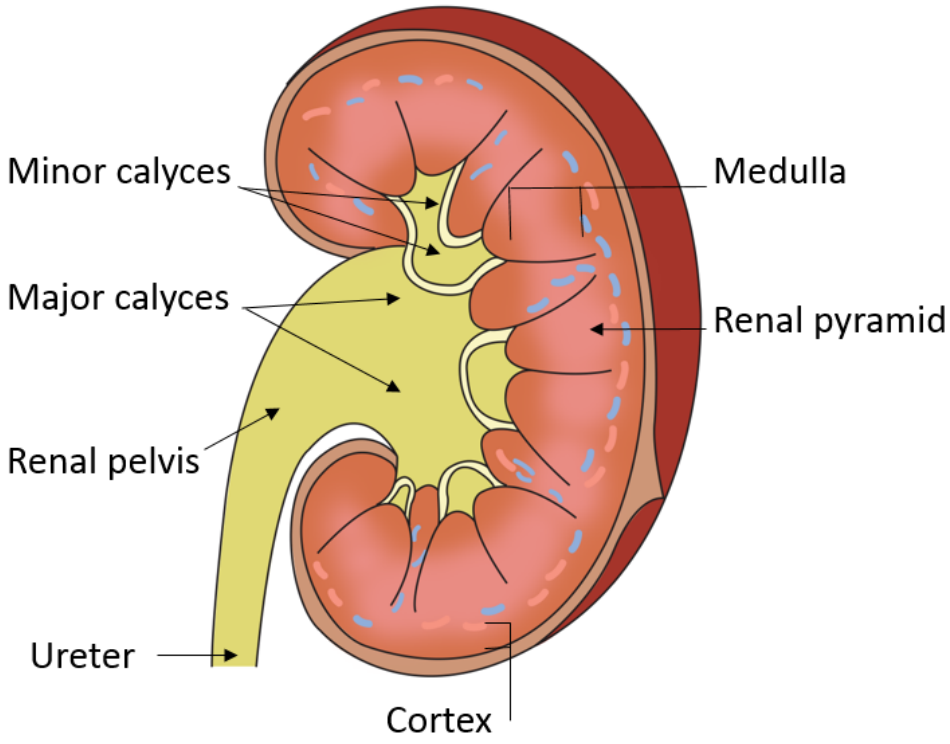

Figure 2.1: Illustration of the kidney.

functional unit, the nephron, which are microscopic tubes [59]. Blood supply to the kidney comes through the renal hilum, composed of the renal artery, vein, and pelvis. Through the renal artery and vein, the renal hilum is connected to the aorta and vena cava [59]. A structural example of the kidney can be seen in Figure 2.1.

The nephron is the functional unit of the kidney and performs blood filtration. Filtration starts in the renal corpuscle, which is composed of the glomerulus (a bundle of capillaries) and the Bowman's capsule which contains the glomerulus [59]. The afferent arteriole brings blood glomerulus, while the efferent arteriole takes blood away [59]. The glomerular filtrate then travels through the proximal convoluted tubules, the Loop of Henle, and the distal convoluted tubule where further filtration occurs [59]. The distal convoluted tubule ends in a single collecting duct, leading to the renal pelvis. In order to evaluate kidney function and health,

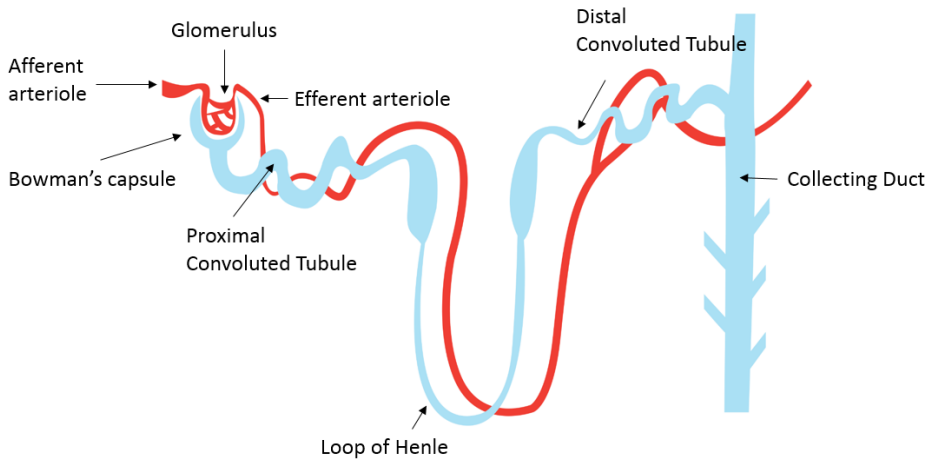

Figure 2.2: Illustration of a nephron, the functional unit of the kidney.

the glomerular filtration rate (GFR) is used [59]. The structure of the nephron is illustrated in Figure 2.2.

The kidney has a multitude of roles to play in maintaining homeostasis [59]. The kidney's roles include:

- **Waste Excretion**: filtration and execretion of toxins such as urea is excreted by the kidney
- **Urine Regulation**: regulation of the urine volume, and additional ions such as sodium and potassium.
- **Blood Pressure Regulation**: maintenance of the blood pressure through renin production, vessel constriction as well as the concentration of salts and water in the body.
- **pH Regulation**: maintenance of a balance of hydrogen ions in the blood itself
- **Hormonal Secretion**: production of erythropoietin which causes the creation of blood cells in bone marrow and activates vitamin D which causes absorption of calcium.

Due to the crucial role of the kidney plays, renal failure becomes a significant issue. Renal failure can be treated by either hemodialysis or peritoneal dialysis [59]. With hemodialysis, a machine is used to filter blood, acting as an artificial kidney equivalent. Hemodialysis requires minor surgery to access blood vessels. These treatments take multiple hours and occur multiple times a week. With peritoneal dialysis, a catheter is inserted into the abdomen which is then filled with a dialysate [59]. The dialysate itself causes waste removal. This approach permits the blood to stay within the vessels themselves. Any form of damage or disease that impedes renal function may cause renal failure [59]. One such example is kidney cancer of which the most common type is renal cell carcinoma.

2.1.2 Renal Cell Carcinoma

In North America, kidney cancer is estimated to be the sixth most common cancer in men, and eighth in women. In the United States alone, an estimated 62,700 cases of kidney cancer were diagnosed in 2016, causing 14,420 deaths [65]. Despite a relatively high survival rate, kidney cancer has an increasing incidence rate commonly due to incidental discovery from medical imaging. [65]. Of this cancer, renal cell carcinoma (RCC) is the most common type making up 85% of all cases.

RCC occurs because of the uncontrolled growth of cells within the lining of kidney tubules [83]. The cause is currently unknown [83]. Warning signs of the onset of kidney cancer include urine containing blood, an unexpected abdominal mass or lump, appetite loss, unexpected weight loss, and pain [83]. RCC can be diagnosed through blood tests, CT or US imaging, or renal mass biopsy. RCC tumours vary significantly. Tumour descriptors include maximal diameter, exophytic and endophytic properties, nearness of the tumour to the collecting system, anterior/posterior location, and location relative to the kidney's polar lines. These descriptors are used to score the kidney using the RENAL nephrometry scoring system [36]. This RENAL score quantifies tumour properties and is used to informs clinical decision making, amongst other factors such as co-morbidities [36]. This nephrometry measure provide meaningful comparisons of RCC from case to case. Moreover, a high RENAL score has been found to be predictive of complications and increased warm ischemia time [25]. The use of pre-operative CT can

also be used to inform the choice of treatment.

Treatment options to RCC include surveillance, ablation, surgery, and radiation. Of these methods, only surgery is only the known curative option for RCC. Active surveillance is a reasonable alternative in cases where the patient is unfit for surgery or suffers from co-morbidities [7], or if the tumor is very small at time of detection. However, the risk of cancer progression remains. If not appropriately observed, may not be eligible for certain surgical procedures at a later date. Radiofrequency-ablation and cyro-ablation are additional modes of therapy which are being developed.. Radiation is considered a palliative option. Laparoscopic partial nephrectomy is the surgical treatment of choice for tumours less than 4 cm in diameter. Of these tumours, endophytic tumours (those with significant volume of tumor subsurface) cause a high rate of complications [78]. The different type of nephrectomies are described in the subsequent section.

2.2 The Nephrectomy

The nephrectomy is the name for the surgical removal of a part or the entirety of a kidney from a patient. It is performed in order to treat kidneys that are injured, or are diseased such as in patients suffering from RCC. The procedure has several variants. It can be completed as open or laparoscopic surgery and as a complete or partial procedure.

In the open approach, the surgeon will make a single large incision into the patient's abdomen in order to access the affected renal unit. This incision causes significant post-operative pain and requires lengthy recovery times. However, the surgeon is able to use their tactile senses, see the entirety of the working space, and can perform the surgery with full dexterity. The workspace is completely exposed to the surgeon, and he or she can directly access whatever area is needed. In contrast, the laparoscopic approach has the surgeon operate with rigid surgical instruments in the body. From here, the surgeon's sensory experience is reduced, but the patient benefits from reduced pain and a shorter recovery time.

The complete (or radical) nephrectomy involves the removal of the entire kidney from the patient. This reduces post-operative renal function at the trade-off of completely removing the diseased organ. In recent years, the partial approach

has gained popularity. With the partial approach, the surgeon aims to minimise the amount of healthy kidney tissue excised while performing a complete resection of the cancerous tumour. Doing so improves post-operative total renal function, as the remaining nephrons can still function independently of what is excised. Because of this, the partial nephrectomy approach is often called kidney-sparing or nephron-sparing surgery. Originally, this was indicated for patients in whom radical nephrectomy of the affected renal unit would result in an anephric state. This would result in the need for renal replacement via dialysis. In contemporary times, partial nephrectomy is leveraged so to maximize kidney function for all patients as it has been shown in numerous studies that global reduction in GFR is associated with poorer quality and quantity of life in all patients [40, 73, 74].

Finally, for completeness, there is also the donor nephrectomy. In this variant, a healthy kidney is completely removed from a patient in order to facilitate an organ transplant to a recipient in need. This is in contrast to the radical nephrectomy where a diseased kidney is removed. The recipient and donor in this approach are assessed for fitness in this surgery. The donor must be healthy enough to undergo the surgery, with no pre-existing renal disease. In addition, the donor must not present significant risk factors for future disease that would impact renal function while having only a single kidney. The donor must be able to consent to the surgery, and have a compatible blood type with the recipient. The recipient must be fit for surgery, and have had their co-morbidities diagnosed, treated, and stabilised.

It is worth noting that tumour enculeation is a new alternative to partial nephrectomy. As RCC compresses the the parenchyma, it creates a pseudocapsule around the tumour [38]. The enucleation of the tumour is then possible using this pseudocapsule, achieving similar cancer survival rates as the partial nephrectomy [38]. However, tumour enculeation still requires further clinical studies and is not the method performed at the local hospital site. While this thesis focuses on improving the partial nephrectomy, the systems and principles addressed here are also applicable in the context of RCC enucleation and indeed, any mode of mechanical intervention to RCC resection.

2.2.1 Procedure Overview

The generalised steps to completing a laparoscopic partial nephrectomy at the local institution of Vancouver General Hospital (VGH) are as follows:

- **Tumour Exposure**: The peritoneum and the Gerota's fascia must be dissected and mobilised to expose the kidney itself. The Gerota's fascia wraps around and compresses the perinephric fat that surrounds the kidney. Upon mobilization of Gerota's, this fat must be dissected in order to expose the kidney surface. The gonadal vein, ureter and hilum must be exposed. Finally, any additional fat is dissected if needed to identify the tumour of interest.

- **Boundary Identification**: Upon exposure, the surgeon must identify the bounds of the tumour and will commonly mark the bounds on the surface using electrocautery. To guide their demarcation, the surgeon will frequently use an US transducer. This use of US is described further in Section 2.3. The entirety of this step is referred to as the planning stage in this thesis. This stage is not considered to be under time constraints as the renal hilum is not clamped, and the kidney perfusion is nominal.

- **Kidney Clamping**: After the bounds of the tumor have been identified, the surgeon will cut blood off from the kidney by clamping the renal artery and/or vein at the hilum. The interval from clamping of the hilum through the duration of the remainder of the surgery, until hilar unclamping, is known as warm ischemia time. This is the time in which an organ is cut off from its blood supply but remains at body temperature. The length of the warm ischemia time has the potential to negatively impact the patient, as described in Section 2.2.2. The accepted threshold is 25 minutes.

- **Tumour Resection**: The tumour resection itself is performed with the surgeon incising into the healthy parenchyma surface of the kidney near the tumor. The surgeon must interpret their marked boundaries, and remember the tumour's subsurface shape and pose from the (now removed) US and preoperative imaging, to make the initial incision.. The surgeon must continue without imaging and complete the resection. This is referred to as the excision stage in this thesis. The excision stage is particularly challenging for

endophytic tumours because the ideal approach is to start as close as possible to the tumour and excise straight down from the organ surface along the orthographic projection of the tumour. For spherical tumours, the ideal excision specimen would fit within a cylinder with diameter equal to the tumour. The surgeon commonly takes a surgical margin of healthy tissue around the entirety of the tumour. This is further discussed in Section 2.2.2.

- **Kidney Reconstruction**: Finally, the surgeon must reconstruct the kidney due to the large defect now created from tumour resection. This involves the time consuming and meticulous action of sewing vessels, and often performing renorrhaphy (suturing of the kidney). After this reconstruction, the kidney is unclamped and blood flow to the kidney restored.

Further details, including patient positioning, port placement, and post-operative management can be found in Zhao et al. [89].

Instead of the conventional approach described above, there is the growing movement in performing RAPN. Commonly performed with the da Vinci surgical system®, the RAPN has the surgeon operate in an enhanced environment compared to conventional laparoscopy. The da Vinci® facilitates improved dexterity and precision with its improved ergonomics, filtration of tremors, and articulated instrumentation. It has an additional robotic instrument which the surgeon can use. It does, however, completely remove the haptic feedback of the surgeon, making it difficult to localise subsurface structures like RCC difficult. In a recent study of 65 patients with completely endophytic tumours, it was shown that the use of robotics could result in the safe excision such tumours [4]. Generally, the RAPN has been shown to be effective for both cystic and solid tumours, favorable in improved renal function, shorter warm ischemia time, and less blood loss and learning time [1, 11, 55]. Because facilitating the adoption of the robot-assisted approach can improve the frequency of successful partial nephrectomies [64], the RAPN is used as the exemplary surgery in this thesis.

2.2.2 Operation Benefits and Challenges

The laparoscopic partial nephrectomy yields several advantages over its open and radical counterparts. There is a benefit to preserving kidney tissue, as more tissue

is likely to reduce the chance of requiring dialysis. Dialysis worsens the patient's lifestyle, limiting their ability to work and increasing the risk of infections and other diseases like cardiovascular disease [44]. It further increases the mortality rate [18]. In comparison to the radical approach, the partial approach's preservation of kidney tissue is directly attributed to improved health outcomes. In a comparison of patients receiving the partial nephrectomy against the radical nephrectomy, results shows those receiving the partial nephrectomy have equivalent long-term oncological outcomes and even an improved overall survival by as much as 10% [40, 51, 71, 73]. Further, the patient is at a reduced risk of developing renal insufficiency and proteinuria (excessive amounts of protein in the urine) [40]. According to the American Urological Association, radical nephrectomy has potential of increasing the risk of kidney disease itself. This is because the removal of one kidney reduces the overall global kidney function, while increasing the filtration requirements of the single kidney left behind, potentially resulting in kidney insufficiency.

As Zhao et al. succintly note, "renal function following [laparoscopic partial nephrectomy] is dependent on quality, quantity, and quickness" [89]. Quality refers to the kidney reconstruction, status of the excision and handling of complications. Quantity refers to the amount of healthy parenchymal tissue remaining in the kidney post-operation. Quickness refers to the length of the warm ischemia time experienced. In comparison to the open approach, the laparoscopic approach is associated with shorter operative time, less blood loss, and reduced hospital stay. However, the laparoscopic partial nephrectomy is also associated with longer ischemia time and more urological postoperative complications such as hemorrhage and urine leakage.

The surgeon must also operate with the constraints imposed by MIS: a reduced field of view, poor depth perception, and reduced haptic feedback (or in the robot-assisted approach, no haptic feedback at all). Such constraints may cause the surgeon to deviate from the ideal excision plan during operation, as it impacts their ability to localise structures like blood vessels, nerves and tumours. In the case of endophytic tumours, this challenge can be reflected in the fact they have a 47% complication rate, nearly five times that of exophytic tumours [78]. Part of this can be attributed to their depth within the kidney, increasing the risk of the surgeon cutting into the collecting system.

While it is rapidly being adopted, the surgery itself is complex. All surgical steps are fairly involved or time-consuming. A component of this includes the identification and localization of the renal hilum, and correctly clamping it. Should the hilum, or the contained artery and vein, be damaged, severe blood loss occurs. This forces a slow approach as the surgeon "feels" their way around the anatomy. The warm ischemia time of 25 minute from the clamping of the hilum is also a significant factor [74]. Thompson et al. showed, with statistical significance, that the kidney is damaged with an ischemia time above this threshold.

Finally, there is consideration of the surgical margin taken. A positive margin is defined as either microscopic (a slight tumour exposure in the specimen) or gross (portions of tumour remaining in the kidney). A negative margin is where the tumour is completely encapsulated in tissue. While small, the laparoscopic partial nephrectomy does have a positive margin occurrence rate of 2.9%, roughly equal to the open partial nephrectomy's rate of 3.3% [38]. To achieve a cancer-negative margin, the surgeon traditionally excises a margin of 10 mm. This means there should be a 10 mm thick layer of parenchyma encapsulating the tumour completely.

Recent analysis shows that that margin size is independent of local tumour recurrence, and not all positive margins produce recurrent cancer [38, 43, 69]. Instead, a normal renal parenchyma margin of 5 mm or less [69] is recommended. Thus, as margin size does not influence this risk, but does influence post-operative renal function, then one should minimise the margin as best as possible while maintaining a negative margin. Achieving this all around the tumour, particularly beneath it, is a difficult task. The surgeon should not simply try to avoid a positive margin, but instead optimise post-operative renal function. Doing so in all cases is difficult, and enhancing the surgeon's ability to do so is the goal of this thesis.

2.2.3 Metrics of Evaluation

As the RAPN is simulated in numerous studies of this thesis, the success of the systems in these surgeries must be carefully evaluated and quantified. To that end, the clinically-relevant metrics of evaluation for the simulated surgeries used in various chapters are as follows:

- **Excision Time**: the time of completion from the start of kidney clamping

through to the end of kidney reconstruction. In the studies performed here, no kidney clamping or reconstruction is simulated, only the excision. This metric corresponds directly to the warm ischemia time.

- **Margin Status**: whether a positive or negative margin occurred. This impacts whether or not an additional surgery is required. This should be negative. While the positive margin rate is low, these systems should at least illustrate non-inferiority to the conventional method.

- **Margin Size**: in negative margins, this is the maximum distance between a point on the tumour and a point on the outline of a cross section of the specimen. This measures the excess healthy kidney tissue excised, impacting post-operative renal function.

- **Excised Tissue Volume**: the volume of the specimen excised is determined by measured weight and known density. This is an additional measure of tissue excised, and similarly impacts post-operative renal function.

- **Adjusted Tissue Volume**: a corrected version of excised tissue volume. However, to account for varying tumour depth, the top layer of parenchyma above the tumour is removed prior to weighing. The specimen's weight is then multiplied by the known density.

- **Specimen to Tumour Volume Ratio**: the ratio of the adjusted tissue volume, weighed post-operatively, to the tumour's known volume, measured during construction.

- **Depth Beyond Tumour**: the distance of the excised tissue that extends beneath the tumour. Determined by US imaging of the excised specimens, this metric corresponds to one of the most challenging components of the partial nephrectomy. It also evaluates the risk of cutting into the kidney's collecting system.

- **Cross-sectional Hausdorff Distance**: after excision, the specimen is sliced into cross sections of 5 mm thickness. In the cross section that most exposes the tumour, the tumour outline and the cross section perimeter are

segmented. The Hausdorff distance is the maximum distance between all closest points on the two contours. This metric evaluates the deviation from the ideal excision.

- **Cross-sectional Centroid Distance**: the centroids of the segmented contours are determined. The Euclidean distance between centroids indicates discrepancy in alignment from the ideal resection.

There is also the evaluation of accuracy for the system. The threshold for system accuracy should be feasible to achieve and clinically useful. A high error may cause an increase in tissue excised (the surgeon believes he or she is not safe despite actually being so) and positive margins (the surgeon believes he or she are safe despite not being so). A low error may allow the surgeon to trust the system and operate with confidence, but a certain order of magnitude may not be possible as different components are introduced and their errors accumulate. There is no widely accepted threshold set for accuracy. Because of this lack of a value, this thesis uses the recommended size for surgical margins (5 mm) as the threshold. Using this 5 mm value is not related to the mitigation of cancer. It is chosen such that if the surgeon aimed to take a 5 mm margin at a given point using the guidance system, the guidance system would not falsely indicate that there is no tumour present even with the potential error.

2.3 Ultrasound Imaging

Ultrasound imaging is a valuable medical imaging modality. It can be used diagnostically and therapeutically. With US, like other medical imaging, one can see within the patient without the need to cut. US operates on the concept of processing sound and echoes to visualise anatomy within a patient. The acoustic waves are generated from an array of piezoelectric crystals, which convert electrical energy to and from mechanical energy. The arrays can be linear or curved in structure. These arrays are housed in transducers that transmit the sound pulses. By sending electric signals into the transducer, the crystals vibrate to cause the high frequency sound transmitted into the patient. These waves, with a frequency on the order of MHz, propagate through tissue at a speed of approximately 1540 m/s. Waves are reflected

at the boundaries of different structures, creating echos which are received by the transducer. These reflections occur due to differences in acoustic impedance of tissue. The energy reflected and the energy transmitted is determined by the acoustic impedance.

The same array receives the echoes, generating electrical signals which are sent to a computer for processing and image generation. By analysing the time of flight between when a wave is transmitted, when its echo is received, and the signal intensity of the echo, the depth of a reflection can be determined. The end result is a 2D brightness mode (B-MODE) grayscale image that plots a cross section of the anatomy. This principle of operation requires no ionising radiation or contrast agent, and makes for real-time processing. Each pixel corresponds to the intensity of the echo at that region. US can also come in one dimensional (1D), 3D, and colour Doppler modes. The transducers are typically hand-held, as frequently seen in obstetric and cardiac imaging, but can also be miniaturised for use intra-operatively during laparoscopic surgery. The result is a modality that is both low cost and small in footprint, and that can complement pre-operative imaging.

US does not come without limitations. Inherently, due to the coherent nature of the pulse echo imaging technique, US images will have speckle noise. Areas of high impedance limit the ability to image through bone or lung. Tissue attenuation also limits the maximum depth of the US image. Image resolution and depth are tradeoffs determined mainly by the number of piezoelectric crystals used and wavelength. Even then, the US is strongly dependent on the user's ability to position the transducer and interpret the US image. For the kidney, this requires the US to be held perpendicular to the curved surface in order to get an accurate representation of the underlying anatomy. The interpretation requires a mental registration of the 2D images and form a 3D model from them.

The use of US in laparoscopy has been used for decades. Langø et al. provides an overview on the various use of US to navigate laparoscopic procedures in a variety of soft-tissue abdominal procedures [39]. The use of US range from providing 2D and 3D guidance, for registration of CT to the intra-operative scene, image fusion and guidance. Specifically, in the context of laparoscopic partial nephrectomy, intra-operative US is used to localise the tumour [52]. While one may argue the use of pre-operative imaging suffices, Schneider et al. showed that the kidney

may move as much as 46.5 mm and 25 degrees may occur from the time of pre-operative imaging to the procedure [61]. This is attributed to the change in patient position. Further, US is used to identify the boundaries of the tumour relative to the healthy parenchyma, and is particularly beneficial in cases where the tumour lies intraparenchymal [38, 52]. Imaging can reveal the lateral bounds, the depth of the tumour inferior to the kidney surface, and relative location to other structures like collecting duct or blood vessels [38]. This more precisely informs the site of surgical excision [52].

The use of this imaging information is limited in laparoscopy. In current laparoscopic practice, the surgeon's ability to move the transducer to the ideal pose is restricted. In the robot-assisted approach, the surgeon is often in control of the transducer, and must instruct their surgical assistant on how to move the transducer. Current practice has the surgeon view the US image, while the transducer moves (either by a surgical assistant or themselves), remember what he or she viewed in the image, and mark the tissue with electro-cautery. The transducer is then removed and the excision begins. The image information is not present during the excision, arguably one of the most vital components of the partial nephrectomy. While it informs, US does not currently guide. Displaying previously acquired US images requires that the US images be registered to the surface where the images were taken so that the surgeon sees the images moving with the organ.

That said, US is still beneficial. In fact, a survey of surgeons practicing laparoscopic surgery showed that 84% expect an increase of US in the future [77]. A separate survey showed that the majority of European urologists performing RAPN use US intra-operatively [29]. If the use of the acquired US imaging data could be used throughout the partial nephrectomy, it would likely assist the surgeon in overcoming the numerous challenges in surgery. One method of doing this is through augmented reality.

2.4 Augmented Reality in Laparoscopic Surgery

One method of providing image-guidance during laparoscopy is augmented reality. Milgram et al. describe the reality-virtuality continuum, which incorporates augmented reality and virtual reality displays [46]. On one end lies reality, the real

environment that humans perceive and live in. On the other end lies virtuality, a complete virtual environment with no component of the real world. In between lies mixed reality — a mixture of the environments together [46]. Towards reality is the class of displays known as augmented reality [46]. This refers to the augmentation of the real environment with additional computer-generated inputs. These augmentations have the potential to improve what a person perceives and understands about their world. In recent years, computer science has great advanced what is capable with augmented reality. Technology now allows for people to experience visually compelling and geographically aware augmented reality on hand-held devices, packed with computer power and connected to large computational networks - and all this at the consumer stage. With the various challenges a surgeon must face to provide optimal care, it is no surprise that augmented reality with applications in laparoscopic surgery has been an active area of research.

The ability to augment a surgeon's perception with 3D models and spatial information of critical structures can significantly reduce the complexities he or she face *in-vivo*. In a recent survey of urologists, 87% felt augmented reality had the potential to be used for navigation and is of interest to the medical community [29]. This is because the use of augmented reality has the potential to enhance the surgeon's abilities, and "see" beyond what is conventionally available in the laparoscopic view. Further, the "mental" registration required by surgeons to use pre-operative and intra-operative imaging increases mental workload and reduces accuracy. Hughes-Hallett et al. showed that the use of pre-operative imaging is subject to variability in interpretation by the surgeon for intra-operative use [31]. It is insufficient for image-guided surgery to simply present data, it must utilise and display it in meaningful ways. This section presents a brief review of the numerous systems and efforts to use augmented reality in laparoscopic surgery. This section first discusses the great efforts made towards using augmented reality in laparoscopy, and then focuses in on the use of intra-operative US to provide such guidance. Several surveys on the use of augmented reality in laparoscopic surgeries, different display devices, tracking and registration methods exist [6, 27, 66]. Select publications are described herein.

As early as two decades ago, Fuchs et al. presented an early augmented reality system with the development of a see-through head-mounted device and 3D laparo-

scope [21]. Ukimura and Gill reported one of the first clinical uses of augmented reality in urology [75]. Their augmented reality system presented 3D visualization of anatomy for both laparoscopic partial nephrectomy and radical prostatectomy. They reported that augmented reality is feasible and improved the surgeon's anatomical understanding. Teber et al. presented a novel real-time surgical guidance tool for the partial nephrectomy. Their use of cone-beam CT imaging for intra-operative imaging, and multiple radio-opaque navigation aids allowed them to track the organ in real-time. They evaluated their registered guidance in *ex-vivo* models using agar-based tumours and used manual registration *in-vivo* [72]. The accuracy between virtual and real models only had an error of 0.5 mm [72]. While a significant step forward in providing guidance for the partial nephrectomy, that work required additional ionising radiation, multiple aids to be inserted into the organ, and was used for enucleation rather than excision [72]. Furthermore, aid placement was not guided or informed, risking damage to subsurface structures. Their augmentation was also the superposition of segmented data onto the laparoscope organ [72]. It is unclear how the radio-opqaue navigation aids were excised or removed from the organ, and whether or not they introduce unnecessary tissue damage. Regardless, that system was refined and brought to clinical use [67]. Simpfendörfer et al. reported the successful use of their cone-beam CT approach for augmented reality to localise complex and endophytic tumours *in-vivo*. Fluorescent markers have since been introduced to facilitate automatic registration of pre-operative CT to the intra-operative scene [81]. While such fluorescent markers are promising due to being metabolizable and their robustness in the face of bleeding and smoke, the steps to clinical use requires the development of a clinically-acceptable marker and injection of a contrast agent into the patient [81]. The need to place multiple markers into the organ is a limitation of both radio-opaque and fluorescent markers.

Using pre-operative 3D CT, Su et al. showed it is feasible to register such imaging to the stereoscopic laparoscopic view [68]. They further showed accuracy of 1 mm for their registration [68]. However, their work required initial manual alignment and is not real-time. Later, Mohareri et al. presented a novel guidance system using real-time registered MRI-US in robot-assisted laparoscopic radical prostatectomy. That work integrates a combination of different components including MRI

to trans-rectal US biomechanical deformable registration in real time, registration of the US transducer to the da Vinci®, and semi-automatic image segmentation [47]. Further, they show the first ever use of such a system in human patients [47]. That work builds upon significant engineering, and is an excellent illustration of the great effort involved to create a useful image-guidance system.

Isotani et al. used reconstructed data for pre-operative planning to evaluate renal structures from CT imaging. They used this planning to identify the best approach for their resection in several RAPNs performed *in-vivo* [32]. Intra-operatively however, the use of this data was limited to manual manipulation by a surgical assistant and displayed via the da Vinci's TilePro® [32]. Improving on this, Volonté et al. created a software module that provided a stereoscopic rendering of pre-operative reconstructed data, allowing the surgeon to view the model in 3D [79]. The interaction with this model was improved with the installation of a joystick to allow the surgeon to autonomously manipulate the data [79]. The system was considered by surgeons to provide a perceptible benefit in their confidence.

On visualization specifically, the challenges of depth perception and convincing overlays may be considered a sub-field all on its own. Hansen et al. presented methods for intra-operative visualization that encoded the distances within the texture of the overlays themselves [23]. Despite only being applied to vascular structure, their initial results illustrated visualizations that are expressive and useful, but can unintentionally present too much information [23]. Wang et al. compare different visualizations, including the transparent overlay, a virtual window, and their own depth-aware ghosting method [80]. They noted that a single visualization method may not provide utility for both simple and complex structures, and the problem is nuanced [80]. Wang et al. also presented an interesting model of considering how the surface, when registered to the camera and tumour, impacts the visualization [80]. A related idea is explored in Chapter 5. Finally, Amir-Khalili et al. explore the value of incorporating uncertainty in the augmentations themselves as to improve on the user's trust in the guidance [3]. Using CT imaging, they used a probabilistic segmentation to then overlay onto a stereoscopic video feed, resulting in convincing augmentations [3].

In considering the display technologies, Bernhardt et al. highlight that the most common method is the static display, which is to present a second monitor next

to the traditional laparoscopic video feed [6]. Other methods include projection of augmentations onto the patient's abdomen, head-mounted devices, and silvered mirrors [6]. They note that no work currently exists on the use of a projector within the patient's abdomen [6].

2.4.1 Ultrasound-based Augmented Reality

In laparoscopic partial nephrectomies, the traditional use of augmented reality has been in intra-operative planning. The models and augmentations try to assist the surgeon in understanding the tumour location and ideal excision prior to excision.

Bajura et al. presented one of the first uses of US to augment the real patient abdomen, while the integration of US for use in robot-assisted procedures was presented over a decade ago in the development of da Vinci Canvas [5, 42]. The da Vinci Canvas integrated an US transducer, tracked it with the laparoscope's camera, and visualised the imaging. It was evaluated in target finding and US-guided biopsy tasks, but not during tumour resection. That work noted that the display of the US volume is distracting when overlaid onto the scene [42]. Several years later, the same group explored different visualizations of robot-assisted laparoscopic US [63]. They presented a split screen view of the laparoscopic and US image, a registered wire frame of the US image overlaid onto the scene with a picture-in-picture display of the US, and the registered US on the laparoscopic image itself [63]. They additionally displayed cues to orient the surgeon's transducer and indicated location of landmarks. They found the use of an integrated robot/US system was received with enthusiasm, and yielded improvements in an array of clinically-relevant tasks, even with simple user interfaces [63].

Closely related is the work of Cheung et al. which presented a visualization platform using fused video and US [10]. The platform used EM tracking of a flexible US transducer to fuse US directly onto the the laparoscopic video [10]. They investigated both 2D and 3D visualization, and performed simulated laparoscopic tumour resections on kidney models. While the system accuracy was acceptable between the tracked US and the laparoscope camera (2.38 ± 0.11 mm), they showed no statistically significant improvement in excision time; the use of 3D visualization in fact increased the excision time [10]. Although 2D visualization benefited

the planning time but this stage is untimed as the kidney is unclamped. Cheung et al. does note that the image display is an important consideration, as the direct overlay of the US image is ambiguous in its spatial location relative to the tumour [10].

Pratt et al. presented a navigation system that integrated pre-operative CT- and MRI-based models with semi-automatic registration. Pratt et al. presented models and surgical margins using virtual and augmented reality techniques like inverse realism in real-time [57]. The augmentation was validated in both offline and online analysis. While the registration error was as high as 3.16 mm, it shows the feasibility of image guidance for the RAPN procedure, and the value of providing such guidance.

Pratt et al. additionally showed the tracking of US intra-operatively, without the use of EM hardware, is practical using computer vision methods [56]. They presented the use and calibration of a checkerboard pattern attached to a micro-surgery US transducer [56]. From this marker, they tracked the US in 6 degrees-of-freedom (6-DOF) and create freehand volumes. However, complex image analysis involving the triangulation of the detected pattern, enforcement of topographical constraints, outlier removal, and parallelization was required. They achieved real-time use and provided relatively small operating range of 42 mm from the laparoscope [56]. Further, this study did not use the US to perform the surgery and only had surgeons estimate tumour thickness. In a similar vein, Jayarathne et al. and Zhang et al. extended the idea of transducer tracking to non-planar transducers. Jayarathne et al. use a Gaussian Mixture Model to estimate a laparoscopic transducer using previously acquired data and known geometry of the pattern [33]. However, it is not real-time. Zhang et al. combine the circles grid pattern in this thesis with corner features, and developed a real-time method of tracking based off this new pattern, but do not validate this method for the accuracy of its US augmentations [87].

Finally, there is Hughes-Hallett et al.'s presentation of an image-enhanced operating environment built around the RAPN. This study is the largest of its kind, reporting over 60 cases that use image-guidance in both planning and excision stages of the partial nephrectomy. Its planning stage is unique as a tablet interface is used to visualised pre-operative segmented imaging data which is not registered auto-

matically to the scene. The excision stage is similar to Pratt et al.'s work in using registered intra-operative US to create and display a reconstructed 3D volume, augmented onto the surgical scene. This was done to assist the surgeon in accounting for tissue deformation. That work does not report quantitative surgical outcomes such as reduction in tissue volume excised, but does report a subjective benefit. It uses improved marker tracking from Pratt et al. which was shown to be robust *in-vivo* [58]. Motivated by these works, there remains a need for real-time and intuitive US-based image guidance during the laparoscopic partial nephrectomy, for both its planning and excision stages.

2.5 Challenges of Guidance in Laparoscopy

Despite the significant advances developed over the last few decades, there has not been a widespread adoption of image guidance systems in laparoscopic surgery [54]. There are inherent challenges in providing beneficial augmented reality in an environment as complicated as laparoscopic surgery. There are numerous technical challenges that must be tackled including accuracy and perception [6, 54]. These challenges include:

- **Accuracy**: the most important criteria for laparoscopic augmented reality, the system must provide high accuracy in order to be useful [6]. The registrations involved, imaging modalities chosen, and the dynamic scene will impact the accuracy [6]. A proper assessment of system components and the accumulated errors should be done.

- **Organ Motion**: the organ of interest may move significantly between the time of imaging to the time of excision [6]. During the operation, mobilization of structures around the organ may shift it significantly, affecting its positioning [6]. Even normal respiratory function or blood pulsation may cause these organs to shift [6]. Schneider et al. reported significant kidney motion occurred between pre-operative and intra-operative imaging [61].

- **Deformation**: the soft organs within the abdominal cavity will deform when interrogated by instrumentation. This deformation must be accounted for

when providing guidance. Accurate modeling has been shown to be difficult, with inaccuracies of 3-4 mm from actual to modeled deformation [2, 19].

- **User Friendly**: image guidance systems should be usable and reduce cognitive load on the already stressed surgeon [54]. An intuitive interface will improve the surgeon's perception and understanding of the scene [54]. It will also minimize the number of interactions required from the surgeon [6].

- **Visualization**: the visualization is a vital aspect of the utilty of image-guidance systems. Too much or too little information may be presented [23]. Visualizations also risk being difficult to interpret, increasing cognitive load of the surgeon. Additionally, accuracy of registration may impact the visualization, and the surgeon's trust in the image guidance.

- **Validation**: the guidance should be reproducible and consistent [6]. Works that include phantom models for validation still need to consider the nuances of the *in-vivo* environment and show robustness in such [6]. There is a need for a validation method that is repeatable in all cases, including the guidance for subsurface structures [6].

- **Latency and Refresh Rate**: the guidance provided must have low latency, and a high refresh rate. Doing this for systems with complicated hardware, requiring synchronization of many components is difficult. For algorithms that are computationally intensive like dense registration or dynamic renderings, there may be a trade-off between accuracy and computation speed.

Beyond the technical barriers, other challenges include the demonstration of improved or maintained long-term patient outcomes, reduction of operation time, avoidance of additional monitors, and cost-effectiveness [54]. These all require high volume cases studies which are difficult to achieve [54]. Most of the works have only been done on small volumes of *in-vivo* cases, evaluated on *ex-vivo* models, or simulated.

2.6 Remaining Needs

To understand where the proposed systems fit in relation to this previous work, one must consider the stage at which guidance has been used, the imaging modalities, and point of view of the guidance. Peters and Linte note that the task of understanding the tool to target relation is equally as important as of identifying the target's location [54]. In the context of a partial nephrectomy, it is reasonable to relate the target location task to the intra-operative planning stage and the instrument to target task to the excision stage. Guidance is valuable in both. The majority of work in the field has focused on the planning stage of the partial nephrectomy, and so there remains a need for continous guidance during the excision. Work that has contributed to the excision stage has used intra-operative CT, adding additional radiation, while the use of US imaging has been limited to the planning stage, despite its potential value in understanding the tumour's depth.

The choice of display method is also important. Numerous works have tried to superimpose or "fuse" the imaging data acquired onto the laparoscopic scene. However, there are concerns on the impact this has to operative inattentional blindness (failure to recognize unexpected stimulus), and the risk of occluding unexpected regions of interest [14, 28, 54]. This thesis tries not to interfere with the surgeon's endoscopic view directly, and explores supplemental displays and, for the first time, projections within the abdomen.

The field to date has, naturally, explored augmentations rendered from the laparoscopic point of view. This is a consequence of using the laparoscope. This has certain limitations, and may in fact not be the ideal choice for the point of view. This thesis explores alternative viewpoints for rendering.

Chapter 3

Intra-operative Ultrasound-Augmented Reality

In order to address some of the challenges of performing laparoscopic and robot-assisted laparoscopic partial nephrectomies, this work proposes three novel augmented reality systems using intra-operative ultrasound imaging. All of these are designed with the overarching goals to reduce the volume of healthy kidney tissue removed and reduce the warm ischemia time. As mentioned in the previous chapter, maximizing of healthy parenchyma and maintaining a warm ischemia time under 25 minutes will improve post-operative renal function [74]. This chapter introduces the framework that is used throughout all systems in this thesis, and then introduces the first system called Nephrectomy Guidance using Ultrasound-Augmented Navigation (NGUAN). It additionally introduces the Dynamic Augmented Reality Tracker (DART), a surgical navigation aid that overcomes challenges in tissue deformation. It is important to note that this chapter's purpose is to present the DART and the NGUAN in the context of a RAPN and evaluate their feasibility in a laboratory setting. The main novelties here are the DART, the tumour-centric tracking paradigm, and the augmentations created from a US-based tumour model.

The structure of the chapter is as follows: in Section 3.1 introduces the hardware and software components; Section 3.2 describes the computer-vision based tracking used for pose estimation of the fiducial markers used throughout this

thesis; Section 3.3 provides an overview on how the system works to provide guidance; Section 3.4 covers the transformation theory behind the augmented reality overlays; Section 3.5 discusses the augmentations themselves; Section 3.6 discusses the calibration and accuracy testing performed; Section 3.7 describes the single-user single-phantom study performed, and finally Section 3.8 and Section 3.9 review the results and lessons learned.

3.1 Framework Overview

The augmented reality systems in this thesis are developed from a common framework. This framework was originally created by Dr. Philip Pratt of Imperial College London, and has been developed in conjunction with the author and co-authors at the University of British Columbia over the course of three years. After several iterations and the extension of the framework to support external modules, it now incorporates interfacing with Analogic US machines, tracking fiducial markers in the scene, efficiently reconstructing the 3D US volume, and then displaying the augmentations via a display device. These are all done through different C++ modules written by the author that leverage OpenGL and OpenCV which are two publicly available programming libraries for graphics and computer vision respectively.

3.1.1 Hardware Components

The framework is composed of a HP-Z820 PC (Intel Xeon E5-2670 2.6GHz CPU and 16GB of RAM). It has a NVidia Quadro 6000 graphics processing unit (GPU) (NVidia Corporation, Santa Clara, CA, USA) in it, along with the NVidia serial digital interface (SDI) Capture and Output cards (NVidia Corporation, Santa Clara, CA, USA). This hardware allows up to four video feeds into the PC for processing and up to two video feeds output. With this, the da Vinci® Surgical System and the da Vinci Si® Surgical System (Intuitive Surgical, Sunnyvale, CA, USA) have their stereo video feeds into the PC. These video feeds are 1080i HD, connected using SDI, and are not hardware synched. The PC can additionally interface with the da Vinci® systems and an US machine (Analogic, Richmond, BC, Canada), via ethernet connections. Both of these machines have their own application programming

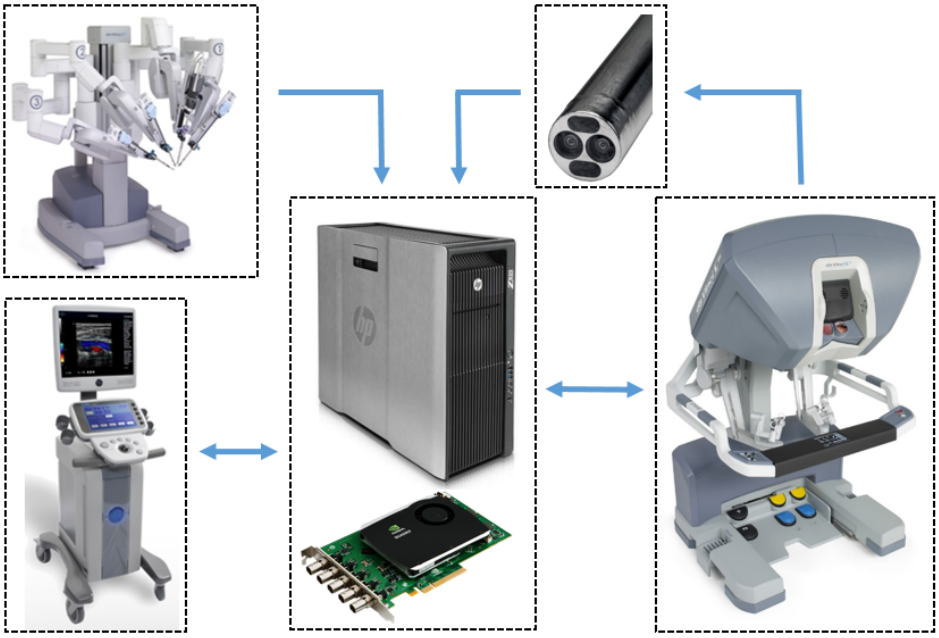

Figure 3.1: System hardware diagram. da Vinci images ©2017 Intuitive Surgical, Inc.

interface (API) allowing for the transmission of data from each machine. For the da Vinci®, each patient side manipulator's tracked pose relative to the endoscopic camera is transmitted. For the US machine, the B-mode image is transmitted. Each of these can be routed into the TilePro® functionality of the da Vinci®, which outputs them on secondary screens in the surgeon's console. While the video feeds can be routed back as the main display for the surgeon, this was not done for this work. That is because initial testing revealed a significant lag in receiving a video feed from the da Vinci®, running the feed through image processing, and returning the feed into the surgeon's console. This lag was not quantified, but was qualitatively large enough to warrant using the TilePro® function. The hardware diagram of the system is illustrated in Figure 3.1.

The US transducer used in all experiments is a custom transducer designed for robot-assisted minimally invasive surgeries by Schneider et al.[62]. It has a 28 mm linear array, contains 128 elements, operates at a centre frequency of 10 MHz, and

Figure 3.2: The custom "pick up" US transducer used in this work from Schneider et al.. Adhered is the tracked KeyDot®.

is compatible with the Ultrasonix machines. It has a unique grasp designed for the Pro-Grasp instrument. This provides autonomy to the surgeon. Schneider et al. showed this transducer has a grasping repeatability within 0.1 mm in all axes, and within 1 degree for roll, pitch and yaw [62]. The US machine is set to a depth of 35 mm, and operated at a frequency of 10 MHz for all experiments. Additionally, the US transducer has a KeyDot® optical marker (Key Surgical Inc., Eden Prairie, MN, USA) on one of its flat faces. These markers have been approved for human use and can be sterilised by auto-clave. Further, these markers can be tracked using computer vision as described in Section 3.2.1 [58]. While this framework interfaces with Analogic US machines, it is extensible to support additional manufacturers as long as a video feed of the machine can be obtained. This means that US machines that give the US image itself can be supported, even without a research API. The US transducer and the KeyDot® are seen in Figure 3.2.

As the da Vinci® has a large catalogue of supported surgical instruments, it is important to note that this work primarily supports a subset of those instruments including the Pro-Grasp, Monopolar Curved Scissors, and Black Diamond Micro Forceps. as the lengths of every tool is known from Intuitive Surgical's instrument catalogue, the framework can be extended to support any number of instruments.

Each of the augmented reality systems produced from this framework are eval-

uated in simulated RAPNs. To this end, realistic kidney phantom models are created. Cylindrical polyvinyl chloride (PVC) phantoms is created using Super Soft Plastic (M-F Manufacturing, Fort Worth, TX, USA). The phantoms additionally had a curved surface. In order to accurately represent the kidney, the phantom's elastic modulus is designed to be 15 kPA, consistent with the reported cortical elastic modulus for *in-vivo* porcine kidneys [22]. Each phantom has a 10-30 mm spherical inclusion at a depth of approximately 20 mm. By design, the tumour appears as a hyperechoic mass in a US image. For the NGUAN, the phantoms are white with transparent inclusions. For the remaining systems, the phantoms are dyed red with black inclusions for ease of post-operative analysis. Due to the depth and endophytic nature of these inclusions, each phantom achieves a RENAL score of 10x given their size, location and depth [36]. This indicates the phantoms simulate difficult cases, and provide a significant risk of cutting into the kidney's collecting duct system. The gold standard for the tumour volume is determined by weight during construction, and using the known density of the material.

3.1.2 Dynamic Augmented Reality Tracker

Tracking the kidney's surface is difficult in both real and phantom models due to its nearly feature-less texture and because of its deformable nature. To overcome these challenges, so that one can register US images to the surface, this work proposes the use of the DART, a custom surgical navigation aid. It is designed in Solidworks (Solidworks, Waltham, MA, USA) and can be 3D printed in stainless steel or plastic (Xometry, Gaithersburg, MD, USA). The DART has barbed legs of 10 mm in length and can be inserted into the patient abdomen via a 12 mm trocar. It can also be repeatably picked up in a similar manner as the custom US transducer used. It has a flat face for the placement of a KeyDot® marker. The DART and a metal version of it are seen in Figure 3.3.

The DART can be inserted in the planning stage of the RAPN and is inserted into the kidney's renal cortex, above the tumour of interest. As there is a layer of healthy parenchyma above the endophytic tumour which is excised regardless, inserting the DART into this layer allows it to be excised along with the specimen. Due to its barbed legs, it is considered to be rigidly connected relative to the tu-

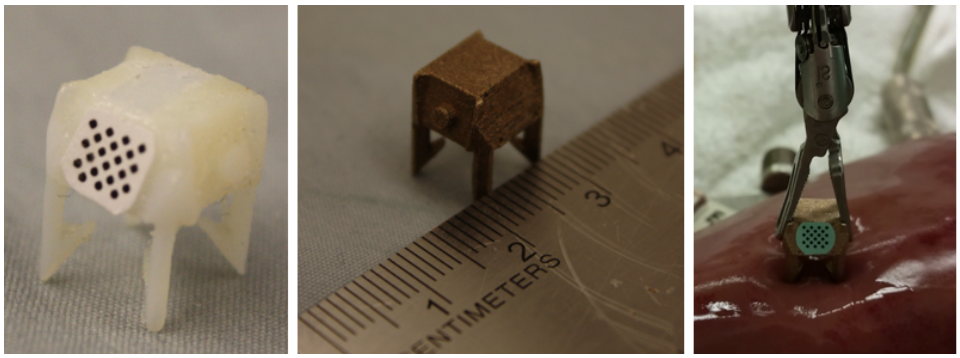

Figure 3.3: The plastic DART with a pattern adhered (left), metal version with scale reference (middle), and the DART as inserted into an *ex-vivo* porcine kidney (right).

mour. Thus, the DART can track a local region of the organ's surface relative to the tumour. With this, tracking of the DART creates a tumour-centric tracking paradigm whereby all tracking and guidance is done relative to the tumour. This is useful in order to reconstruct an accurate US volume, improve system accuracy, and provide persistent information over time, even after removal of the US transducer and without the need for surface tracking.

It is noted that the DART is a intermediate solution. While there have been some steps towards deformable tissue tracking, these are often not robust, are too sparse, or do not account for deformation throughout the entirety of the surgeon. Examples include the case of Collins et al., who present a promising solution but their evaluation does not present an accuracy measure for their tracking and the case of Mahmoud et al. whose ORBSLAM-based approach still achieves a relatively high error for tracking of 3 to 4.1 mm RMSE and is far from real-time [12, 45]. Further, such solutions have not been made publicly available for validation [45]. Until this challenge can be solved, the DART allows the exploration of augmented reality in soft-tissue surgery without the need for advanced algorithms. With this single small drop-in tracker, augmented reality and surgical guidance can be created to enable maximal nephron sparing while maintaining a negative margin. Initial experimentation of the DART inserted into *ex-vivo* porcine kidney showed that the kidney itself could be lifted by pulling on the DART. To assess the rigid relation-

ship of the DART to a simulated tumour finite element modeling (FEM) is done in ANSYS (ANSYS, Pittsburgh, PA, USA). This modeled the DARTs movement in a kidney during an US scan. The kidney is modeled as a 50 x 50 x 50 mm cube. It is modeled as being linear, elastic, homogeneous, and isotropic. The tumour is 20 mm in diameter and simulated to be 20 mm within the kidney. The DART is placed above the tumour, and the US transducer is placed 10 mm from the DART's edge. The model is treated as isotropic and homogeneous with linear elasticity. Poisson's ratio of 0.48 for both the kidney and the parenchyma is used. Input parameters to the simulation included applied US force, barbed leg length of the DART, and kidney stiffness. Using a calibrated force sensor, the average maximum downward force for three complete scans of phantoms is 0.7 ± 0.3 N. Thus the forces of 0.1, 0.5, and 1.0N are evaluated. The leg length is varied between 0, 5 and 10 mm. As Grenier et al. reports different cortical and medullary elasticities for *in-vivo* porcine kidneys (15.4 ± 2.5 kPa and 10.8 ± 2.7 kPa respectively), simulations are done using each average elasticity [22]. To evaluate the simulations, the distance between the theoretical tumour center (20 mm below the aid, regardless of pose) and the actual center is calculated.

3.2 Vision-based Tracking

As described, both the US transducer and the DART have a KeyDot® marker placed on them. These planar markers, illustrated in Figure 3.2 and Figure 3.3, are a grid of black and white circles. The grid is asymmetric in its dimensions, meaning the number of circles per row differs from the number of circles per column, providing rotational invariance. This design allows the system to estimate the pose of the planar grid from a single camera. Further, these fiducials can be tracked in full 6-DOF. In order to discuss the pose estimation however, camera calibration and centroid detection need to be discussed first.

Camera Calibration

Cameras can range in their properties which influence the imaging process itself. These include the focal length, number and types of lenses used, and even whether or not the pixels are isotropic. Using a pinhole camera model, these cameras can

be modeled mathematically. This can then compensate for lens distortion, measure the size of objects in real world units instead of pixels, or estimate the pose of the camera in the world. For this work, Zhang's calibration method is used [88]. In this chapter, this is done using the CalTech Camera Calibration toolbox (a software module for MATLAB) and in the later chapters with OpenCV [8]. Both use the same implementation.

The da Vinci's laparoscope has a pair of HD cameras which creates a 3D stereoscopic display for the surgeon to see. Each camera of the stereo endoscope can be calibrated according to a pinhole camera model with intrinsic parameters. These parameters indicate how a 3D point in space is projected onto the camera's 2D imaging plane. The intrinsic parameters include the focal lengths f_x and f_y, the length from the modeled pinhole to the imaging plane, and the coordinates of the camera's principal point c_x and c_y. The principal point is the location of the intersection of a perpendicular line from the modeled pinhole to the imaging plane. The principal point is often not in the center of the camera image. Further, a skew parameter s is included to account for image axes that are not perpendicular to one another. This is represented by the matrix K, seen in Equation 3.1. This can alternatively be considered a combination of a shear (from the skew factor), a scaling (from the focal lengths), and then a translation (from the principal point). What this means is that a point in the camera's 3D coordinate can be transformed into its 2D imaging plane by Equation 3.2.

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3.1}$$

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_{xx} & r_{yx} & r_{zx} & t_x \\ r_{xy} & r_{yy} & r_{zy} & t_y \\ r_{xz} & r_{yz} & r_{zz} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.2}$$

Where (X,Y,Z) is a 3D world point, the position and orientation of the camera is represented by the r,t parameters, (u,v) represent the pixel point, and s is a scale factor. When considering a set of points that lie on the same plane, this equation

can be simplified by treating the Z component as zero. This results in Equation 3.3, where K and the $(r,t($ parameters can be treated as a single 3 by 3 homography matrix, H. This matrix can then be found via decomposition.

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{3.3}$$

The calibration for a single camera is performed by taking pictures of a calibration target [88]. In this work, the calibration target used is a checkerboard pattern. The pattern is a 7×8 checkerboard with each square being 4×4 mm. The checkerboard pattern is moved in the camera's field of view, with images taken of it in varying poses. It is important to include rotational and translational changes in the set of images, as these will impact the quality of the calibration [88]. Each image is converted into grayscale, and run through a Harris corner detection algorithm, producing a set of 2D pixel points [88]. These points are refined for sub-pixel accuracy. This set of 2D points then corresponds to a set of 3D points, known through the geometry of the checkerboard target [88].

The 3D points are treated as lying a single plane, having a Z-coordinate of zero as in Equation 3.3 [88]. This is done to reduce the problem complexity, simplfify the problem to determining a homography and provide constraints on the intrinsic parameters [88]. These feature points can then be used to estimate the intrinsic and extrinsic parameters using a closed-form solution [88]. The estimations are refined using a Levenberg-Marquardt least-squares optimization [88]. This algorithm aims to minimise the re-projection error of the calibration parameters. Re-projection error is defined to be the sum of squared differences between the 2D points found by the corner detection and the estimated 2D points created when using the current iteration's calibration parameters. Simply, it is the error between where the calibration thinks the pixels are versus where the pixels actually are. The optimization continues until no further improvement in error is observed below a threshold. The result is an iteratively obtained set of camera calibration parameters. Throughout this work, the re-projection error is less than 0.4 pixels; sufficient for use.

However, the above model does not incorporate a lens, which is used to increase the number of rays that pass through the sensor. These lenses can cause

distortion that bend the rays near the edge of the images called radial distortion. There is also tangential distortion when the lens is not parallel to the camera sensor itself. The camera model is expanded to have a set of seven coefficients to account for these distortions.

Each camera of the stereo pair can additionally be calibrated for their extrinsic relation to one another, creating a transformation of how to go from one camera's coordinate system to another. Doing so is useful for stereo-based surface reconstruction, discussed in Chapter 5, and potentially incorporating triangulation for improved tracking accuracy. With a set of calibration parameters for even one of the laparoscope's cameras, one can accurately estimate its pose relative to a known pattern in the scene. This is a fundamental concept in the vision-based tracking used throughout this work.

3.2.1 Pose Estimation

Given a set of 2D/3D point correspondences, it is possible to estimate a calibrated camera's full 6-DOF pose relative to a known pattern by solving the classic computer vision problem called the Perspective-N-Point Problem (PNP).

In the case of the circles grid, a corner detection algorithm cannot be used. The image containing the circles grid is run through a blob detection algorithm, specified to look for regions with certain convexity. Each region is filtered based off its convexity and the known relation of the circles grid. The end result is then a set of 2D points which are the centroids of each circle. This creates the 2D/3D point correspondence as in with the checkerboard pattern. If the entirety of the circles grid is not detected, then the tracking algorithm stops processing the current frame.

The same Equation 3.2 can be used here to model PNP. However, unlike camera calibration, the camera parameters are known and the only components requiring estimation are the rotation and translation (r,t) parameters. These values can be similarly be estimated using the Levenberg-Marquardt optimization. The result is the pose of the KeyDot® (on either the DART or US transducer) relative to the camera is obtained. A limitation of this approach is that the solution to PNP may have multiple solutions as a result of the least squares algorithm. For tracking, this would result in flipped coordinate systems. To mitigate this risk, the previ-

ous found transform is used to initialise the least squares optimization algorithm. This improves the reliability of the tracking, particularly when the pattern is facing perpendicular to the camera which, from experience, often results in multiple solutions.

The process of detecting the circles and computing the pose from them is time intensive on a full HD resolution image. To improve the tracking speed, a motion estimation algorithm is used to create a region of interest for the next frame. The circle detection then first runs in the region of interest, significantly improving speed, and will resort to the full resolution image if no candidate circles grid is found. This method of tracking has been developed by Pratt et al. [58]. Initial *in-vivo* use was explored and proved to be reliable and accurate. Specifically, the tracking algorithm was found to have an operational envelope of -50.1 to 52.5 degrees in rotating about the X-axis, -52.7 to 57.6 degrees in rotating along the Y-axis, and had a working depth of 13 to 86 mm in the Z-axis. The use of motion estimation improves the pose estimation to only taking 11ms, resulting in real-time processing.

Pratt et al. demonstrated that the motion estimation with tracking circular pattern outperformed checkerboard tracking [58]. It showed to track more constantly over a large workspace different illumination levels [58]. These appealing aspects of the tracking algorithm warrants its use in the intra-operative systems presented in this work.

3.3 Principle of Operation

There are several components of NGUAN required for its use, and it is important to understand their role in the overall system. Prior to the steps below, it has been assumed the calibrations have been completed. The principle of operation is described as follows:

1. **DART Placement**: After the kidney surface has been exposed, an untracked US scan is performed. The surgeon uses this scan to estimate the tumour location and identify where he or she would like to insert the DART. In normal practice, this scanning is already done so little additional time is consumed here.

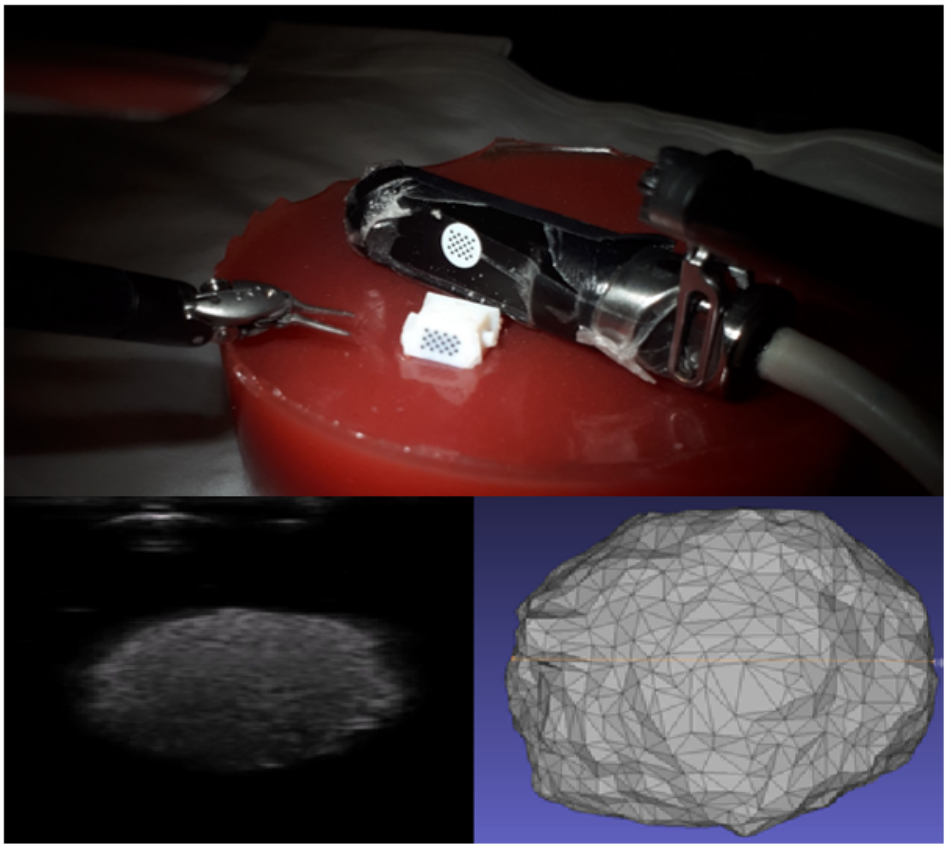

Figure 3.4: Simulated surgery set-up with the DART inserted into a phantom, and tracked US scan performed (top). US images (bottom left) are segmented to create a 3D tumour model (bottom right)

2. **Tracked US Scan**: Once inserted, a freehand tracked US scan is performed. Both the US transducer and the DART's KeyDot markers are tracked by the laparoscope, using the method previously described. The US images are recorded relative to the DART. Depending on the DART location, this scan can be performed in a entirely translational manner or include rotation in order to capture the entire tumour. This step is done during the planning stage where no clamping is done, and there is no time limit.

   The 3D volume is reconstructed from the set of tracked 2D US images. Voxel length and slice thickness are both set to 0.375, determined experimentally

to provide a good quality reconstruction. The optimised reconstruction takes less than 30 seconds. Figure 3.4 shows the DART as inserted into a phantom model, with a tracked US being performed.

3. **Tumour Model Generation**: The manual segmentation of volume's cross sections for the tumour is performed for each slice. Segmentation is done using ITK-Snap, a third-party software [86]. Again, this is performed during the untimed planning stage. From the segmentations, surface extraction is performed.

4. **Augmented Reality Overlays**: In additional to the regular surgical scene view, augmentations are provided to the surgeon in real-time. These augmentations include a virtual rendering of the tracked surgical instruments and the mesh model, as well as an augmented surgical scene. Treating the tumour as a rigid body, as the DART moves, the tumour's movement can be continuously rendered. The surgeons console view is conceptually shown in Figure 3.5, and the overlays are described in depth in subsequent sections.

5. **Guided Tumour Excision**: During the excision of the tumour, if the da Vinci® surgical instrument comes within a set threshold distance of the centroid of the tumour the viewpoints flash red to warn the surgeon he/she is approaching the tumour. Last, the DART is removed together with the tumour and surrounding tissue.

NGUAN differs from the work of Teber et al. in the following ways: only one surgical navigation marker (DART) is inserted into the kidney, the tracking provided is 6-DOF, the augmented reality is a 3D representation of the tumour generated by US, and the augmentations involve virtual camera viewpoints.

3.4 Transformation Theory

In the equations of this thesis, the notation for transformations is as a follows: ${}^{A}T_{B}$ is a 4x4 transformation matrix that takes the coordinate system B and rotates and translates it into the coordinate system A. The notation A_0 indicates the coordinate system A at time $= 0$. As coordinate systems may move over time, a coordinate

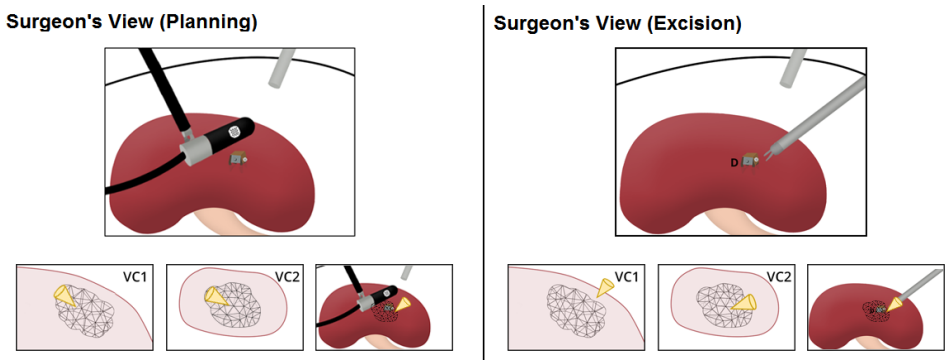

Figure 3.5: Conceptual illustrations of the surgeon's console view in both stages of the RAPN.

transform at time $= n$ would correspondingly be denoted as A_n. The notation for a point in a given coordinate system is p_A.

In order to provide augmented reality overlays, the system must track the surgical instruments relative to the tumour at any given time. Further, for the overlays involving virtual cameras, these virtual cameras must be placed and positioned correctly over time so that they appear fixed relative to the real camera. Each of these has its own set of transformations. The underlying transformation theory for these steps is outlined in this sections.

Illustrated in Figure 3.6, the individual coordinate systems, their origins, axes and units in NGUAN are:

- U: the 2D US image. The origin is the top left pixel. X-axis increases laterally from left to right of the image. Y-axis increases axially from top to bottom of the image. Units are pixels.

- I: the surgical instrument's coordinate system. The origin is located on the da Vinci® instrument's wrist which is tracked by the API. The Z-axis increases along the length of the tool, towards the tip of the tool. The X-axis and Y-axis are arbitrarily defined. Units in millimeters.

- C: the calibrated camera's 3D coordinate system. The origin is located in 3D space. X-axis moves left to right on the camera image. Y-axis moves top to

bottom on the camera image. The Z-axis goes into the camera image itself. Units are in millimeters.

- L: the laparoscope's coordinate system. The origin is unknown but located somewhere physically within the laparoscope. Determined experimentally, the laparoscope's axes and units follows the same conventions as C.
- V: a virtual camera coordinate system. The origin is defined arbitrarily. The axes and unit conventions at the same as C.
- K: the KeyDot® marker on the US transducer. Origin is in the top left circle. X-axis moves up along the columns. Y-axis moves right along the rows. Z-axis is into the page of the marker. Units are in millimeters.
- D: the KeyDot® marker on the DART itself. Same coordinate description as in K.

For simplicity, the differences in convention between OpenGL and OpenCV, the intermediate coordinate systems used in OpenGL when rendering, and the projection of a 3D point onto the camera's 2D imaging plane are not covered in this thesis.

3.4.1 Virtual Cameras and Time

When considering the surgeon's normal viewpoint, surgeons have difficulty interpreting two aspects of the tumour: how deep the tumour lies and how far from the laparoscope it extends. To facilitate easier view of these aspects, two orthogonal viewpoints relative to their normal view are provided. The first viewpoint is a "bird's eye" view and the second viewpoint is looking in from the side of the scene. These viewpoints can be achieved using virtual cameras placed relative to the DART's first detected pose. They are not placed relative to the real camera as it would create a confusing experience to move the camera and have two additional viewpoints move at the same time. The same is true for moving the DART over time. Placing virtual cameras relative to the first detected pose lets their positions to remain fixed despite a dynamic scene.

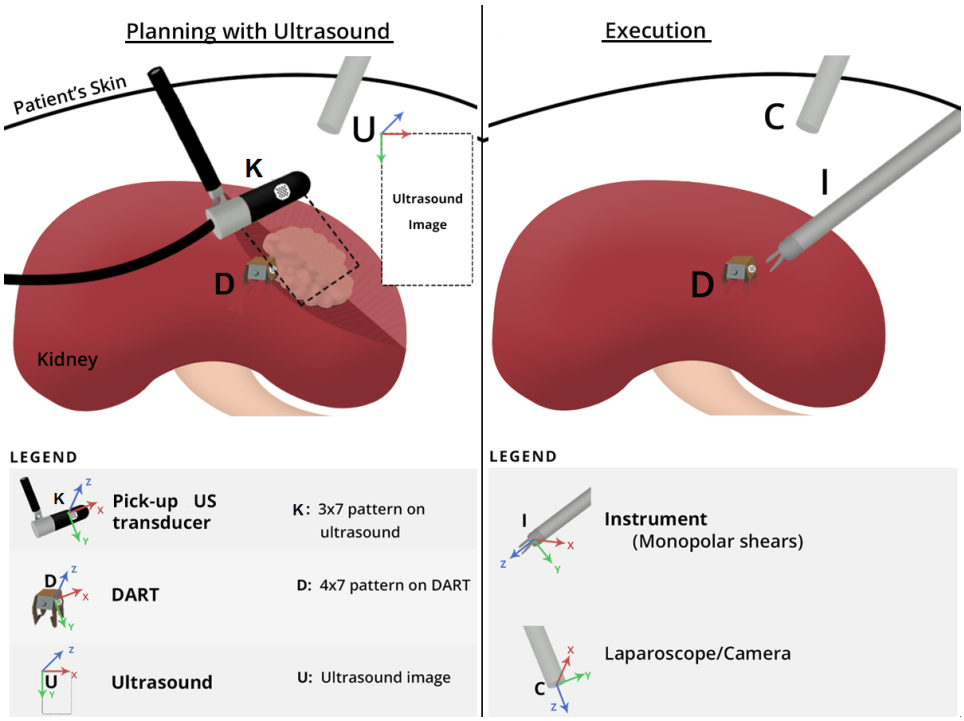

Figure 3.6: Coordinate system diagram in each stage of the RAPN using NGUAN.

Virtual cameras themselves are mathematical models of cameras that do not exist (hence "virtual"). Using the same pinhole camera model as in calibration, it is possible to simulate alternative viewpoints of a 3D scene. As the surgical instruments and tumour are modeled in full 3D, it is then evidently possible to create virtual viewpoints. Doing so requires the instruments, tumour, and virtual cameras to be registered in the same coordinate system.

These virtual cameras are tumour-centric, as their placement is focused around the tumour model. The transformation from the virtual camera coordinate systems V to the initial DART coordinate system, D_0, is decomposed into its rotational and translational components which are each computed separately. The translations are a pre-defined distance between the tumour and the virtual camera. This is described by Equation 3.4. ${}^{D}N_{C}$ can be any of the three unit column vectors of the transform ${}^{D}T_{C}$, describing an axis of the real camera in the DART coordinates. For the virtual

cameras in NGUAN, these are the X and Y axes of the real camera. The constant s is an arbitrary scalar constant that sets the distance along that real axis. p_D is the 3D location of tumour centroid in DART coordinates. The result of Equation 3.4 is the translation component, t, of ${}^{D_0}T_V$ for each virtual camera.

$$t_{D_0 T_V} = p_D + s \cdot ({}^{D}N_C) \tag{3.4}$$

The rotational component for the virtual cameras is defined to be 90 degree rotations about either the real camera's X or Y axis relative to D_0. Combined with the translation above, this creates ${}^{D_0}T_V$ for for each camera. This suffices for the first time the DART is found. However, in the description thus far, only the surgical instruments may move. If the DART or camera move, these changes in the coordinate systems cannot be accounted for as is. Therefore when either the camera or DART move, the pose of the virtual cameras must be updated. This is done using the relative transform of the DART at $t = 0$ to $t = n$ as in Equation 3.5. This can then be applied to ${}^{D_0}T_V$ to produce ${}^{D_n}T_V$. as seen in Equation 3.6.

$${}^{D_n}T_{D_0} = {}^{D_n}T_{C_n} \cdot {}^{C_n}T_{C_0} \cdot {}^{C_0}T_{D_0} \tag{3.5}$$

$${}^{D_n}T_V = {}^{D_n}T_{D_0} \cdot {}^{D_0}T_V \tag{3.6}$$

3.5 Augmented Reality Overlays

Summarizing, the surgical instruments can be continuously tracked relative to a US-based tumour model in 3D. Two virtual cameras can be placed relative to the tumour model over time, with the DART and camera moving freely. This information must be relayed to the surgeon. The surgeon's normal video feed is supplemented with two augmented feeds with TilePro®.

Seen in Figure 3.7, the first is the surgeon's endoscopic view with augmentations (referred to as "direct overlay"). The second is a split screen of the two virtual viewpoints (top-down and side). The views both face the centroid of the tumour and remain fixed relative to the real camera. The tumour and instruments are continuously rendered as the DART moves. The rendering also displays the movement of the tumour in the virtual viewpoints. For NGUAN, the surgical instruments

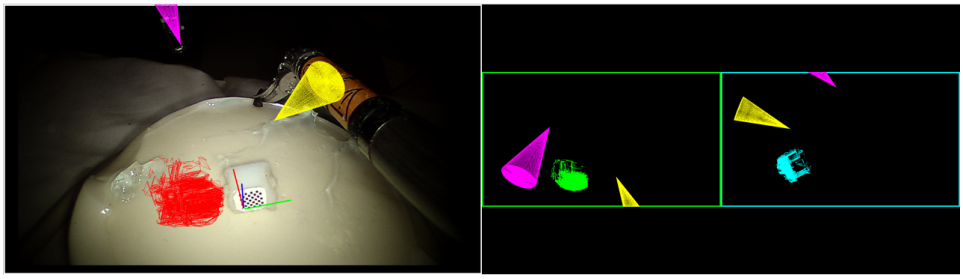

Figure 3.7: The set of visualizations as presented in TilePro®. Endoscopic view augmented (left) and virtual viewpoints (right). Pink and yellow cones are virtual renderings of the tracked surgical instruments. Red, green, and blue meshes are visualized in each view. No interpolation was performed between segmented slices of the mesh, resulting in the poor mesh visualized.

are rendered as cones. These cones are centered in I with a height equal to the current instrument's length. These lengths are obtained from Intuitive Surgical instrument catalogue. The left instrument is coloured pink, and the right instrument is coloured yellow to distinguish between them. The tumour model is coloured red in the direct overlay, green in the top down view, and blue in the side view. Furthermore, a flashing red warning is given to the surgeon based on his or her instruments' distance to the tumour center and a pre-defined threshold.

3.6 System Calibration and Accuracy

There are several components in the NGUAN system that require calibration. The calibration for the US transducer, the da Vinci® to camera, and the total system error are described in this section.

3.6.1 Ultrasound Image to KeyDot Transform

To create reconstruct a 3D volume from the 2D US image, relative to the DART coordinate system, the pixel to millimeter scale factor and US calibration is needed.

To convert a pixel value to a physical value, the pixel to millimeter scale factor is required. This is determined by imaging a block at known dimensions and observing its US image. Segmenting the block in the lateral and axial dimensions, and

dividing by the known length obtains the scale factor. The US image is assumed to have isotropic pixels, so the same scale factor is expected in both dimensions.

The purpose of the US calibration is to calculate the transformation from the 2D image from to the KeyDot® marker on the transducer face. These two coordinate systems are illustrated in Figure 3.6. The unknown transformation between the two systems is denoted as ${}^{K}T_{U}$. Given the US transducer's CAD model is available, this transform is determined geometrically.

To assess the US calibration accuracy, reconstruction accuracy of a pinhead is used. By imaging a pinhead in a water bath from 10 different poses of the tracked transducer, and manually segmenting the US images for their 2D pixel locations, each pinhead point can be transformed into 3D coordinates. The Euclidean distance from each pinhead point to the centroid of all points is calculated. The root mean square (RMS) error is reported.

With this transform, a tracked US scan relative to the DART can be done. This is captured in Equation 3.7. Here ${}^{C}T_{K}$ is the transform from the Keydot to the calibrated camera, and ${}^{D}T_{C}$ is the transform from the camera to the DART, both obtained from pose estimation described previously.

$$p_D = {}^{D}T_C \cdot {}^{C}T_K \cdot {}^{K}T_U \cdot p_U \tag{3.7}$$

3.6.2 da Vinci Laparoscope to Camera Transform

The API from Intuitive Surgical provides tracking information of the instruments' coordinate system I relative to the laparoscope coordinate system L. The da Vinci's instrument is a 12 foot long, 13 degrees-of-freedom (13-DOF) kinematic chain. This lends itself to an absolute tracking accuracy of approximately 50 mm and relative tracking accuracy of 1 mm [37]. However, the single camera of the laparoscope used for tracking has a different origin than the laparoscope as seen in Figure 3.8. Thus, for accurate tracking of the surgical instrument relative to calibrated camera, the robot needs to be registered with respect to the camera, solving for the unknown transform ${}^{C}T_{L}$. This registers the surgical instrument to the camera coordinate system. This is seen in Equation 3.8. Additionally, the da Vinci's tracking information for a manipulator comes from a combination of a high res-

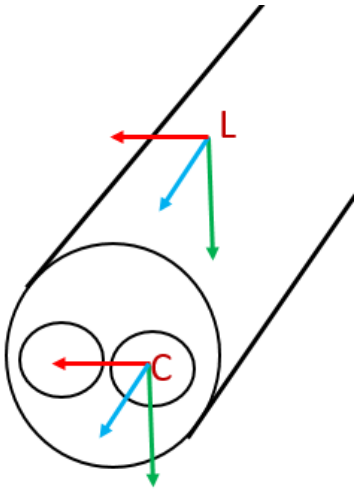

Figure 3.8: The calibrated camera coordinate system (C) differs from the laparoscope coordinate system of the da Vinci® (L). The two must be registered to one another.

olution encoder, giving relative information, and a low resolution potentiometer, giving absolute information. Because of inaccuracies in the encoder and potentiometer, these must be accounted for. To simplify this, the da Vinci's tracking error is accounted for in the same calibration as as solving for ${}^{C}T_{L}$.

$$p_D = {}^{D}T_C \cdot {}^{C}T_L \cdot {}^{L}T_I \cdot p_I \tag{3.8}$$

For Chapter 3, solving ${}^{C}T_{L}$ is achieved via registration of 14 pairs of points, one in the camera coordinate system (C) and one in the laparoscope coordinate system (L). To generate each pair of points, a KeyDot® is moved to a unique pose and at each location the surgical instrument tip touches the known origin of the KeyDot®. In turn, a leave-one-out error for each of the 14 pairs is calculated based on a registration of the other 13 pairs using Horn's method. That is, after calibrating on 13 pairs of points, the target registration error (TRE) is calculated using the remaining pair. The RMS of those 14 errors is reported.

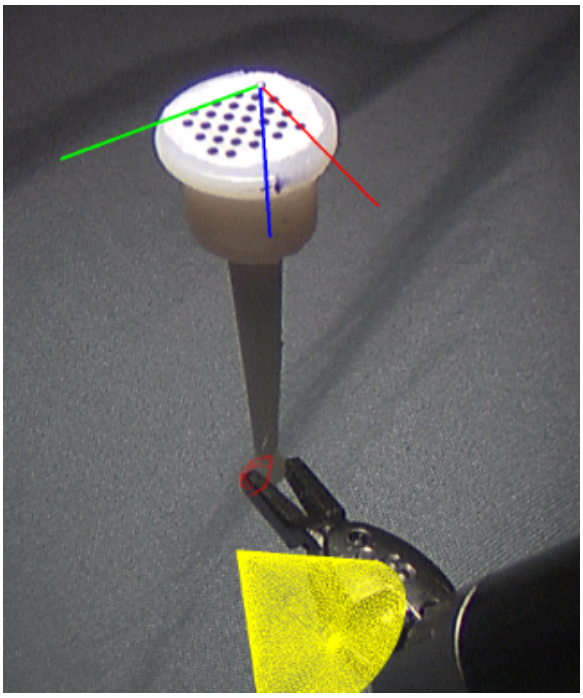

Figure 3.9: The modified DART used for error testing with instrument and pinhead overlaid.

3.6.3 Total System Error

Finally, to characterise the accuracy of the overall system, a modified DART is designed. As seen in Figure 3.9, this modified version has a a flat circular top and 2.5 mm pinhead that extends along the Z-axis of the DART. This pinhead simulates the tumour centre. This extension is 25 mm in length. By taking tracked US scan, a pinhead model can be generated in the DART coordinate system. With this, the da Vinci® surgical instrument is used to pick up the pinhead itself. The instrument's location in the DART coordinate system is recorded. The error is calculated as the distance between the pinhead centroid and the surgical instrument. The RMS error of 10 different poses is reported.

3.7 User Study

In evaluating the NGUAN, simulated RAPNs are performed. One expert urologist versed in performing RAPNs participated. The phantoms provided had inclusions that are purposefully unique in shape and location, limiting the surgeon's ability to learn from case to another. The tumour models are generated prior to the user study. In each case, the surgeon is instructed to use the US transducer to scan the phantom surface. Using a permanent marker, the surgeon simulated electrocautery and marked the tumour boundaries. This mimicked the planning stage. After this, the surgeon immediately began the excision itself. In the first case, the surgeon is only given the US transducer during the planning stage and no additional guidance thereafter. In the second case, the surgeon spent 20 minutes training and learning the NGUAN system prior to starting the operation. Then the surgeon is given NGUAN to operate with during planning and execution stage.

After both trials are completed, the surgeon answer a questionnaire in which he provided feedback about both cases and both systems. The survey included questions regarding usability and helpfulness of each system. The surgeon is interviewed for open feedback on the system. To capture quantitative benefits, the metrics of excision time, margin status, max margin size, adjusted excised tissue volume, and specimen to tumour volume are reported for the two cases performed. These metrics are previously described in Section 2.2.3. The excised specimen mass is cut into 10 mm slices to determine margin status and size.

3.8 Results

3.8.1 Finite Element Simulations

For all FEM simulations, the distance between theoretical and actual tumour centre never exceeded 1 mm. From this, the rigidity assumption for the navigation aid results in an error of kidney tumour location of no greater than 1 mm. Simulation results are summarised in Figure 3.10.

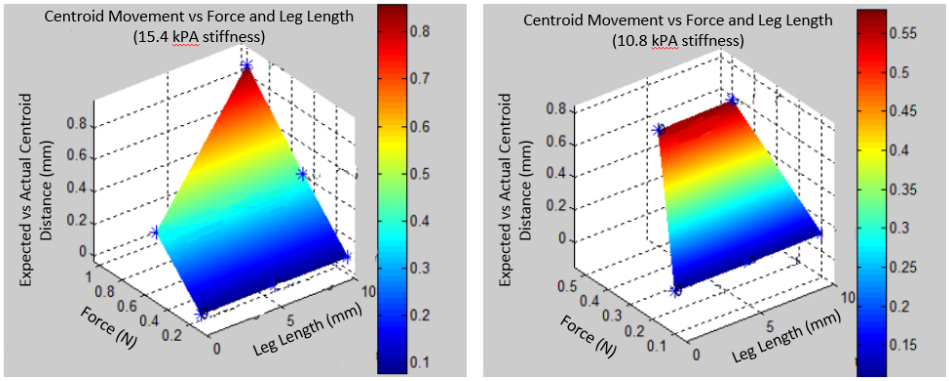

Figure 3.10: FEM simulation of tumour movement as a function of force and leg length using 15.4 kPA stiffness (left) and 10.8 kPA stiffness (right).

3.8.2 System Calibration and Accuracy

The geometric US calibration's pinhead reconstruction accuracy is 0.9 mm RMS. The US calibration result of the pinhead reconstruction relative accuracy is 0.9 mm. Over the course of capturing the 10 US images of the pinhead, the US transducer covered a range of 16 $\times$ 10 $\times$ 19 mm. The da Vinci® laparoscope to camera calibration TRE is 1.5 mm RMS overall. The single lowest TRE is 0.6 mm. The overall system TRE is 5.1 mm RMS.

3.8.3 User Study

In using only the US, the execution time is 10 minutes and 45 seconds. The tumour volume itself is 4 cm^3 and the adjusted excised tissue volume is 24 cm^3 Thus the specimen to tumour volume ratio is 6:1. The largest negative margin size is 24 mm. In using NGUAN, the execution time is 7 minutes and 30 seconds. The tumour volume itself is 5.5 cm^3 and the adjusted excised tissue volume is 16.5 cm^3 Thus, the specimen to tumour volume ratio is 3:1. The largest negative margin is 12 mm. In both cases, there is a gross and a separate microscopic positive margin.

After the user study, the surgeon preferred the use of NGUAN over the US for visualizing the tumour in the execution phase. General comments about the NGUAN system include that the most useful guidance cue is that the screen flashed red once the instruments got to within a certain distance of the tumour. The warning aided

Table 3.1: NGUAN initial feasibility study results. Results of the trials using ultrasound only (US) and the guidance system (NGUAN) are shown.

Metrics	**US (n=1)**	**NGUAN (n=1)**
Excision Time (min:secs)	10:45	07:30
Margin Status (/1)	1 gross & micro	1 gross & micro
Margin Size (mm)	24	12
Known tumour volume (cm^3)	4.0	5.5
Adjusted Tissue Volume (cm^3)	24	16.5
Specimen to Tumour Volume Ratio	6:1	3:1

the surgeon in avoiding the tumour and minimizing the healthy tissue excised. The surgeon found the top-down view easier to interpret than the side view.

3.9 Discussion

The success of image-guided surgical systems is largely dependent on their accuracy, usability and the clinical need for the image guidance. Each of those aspects of NGUAN will be addressed in the discussion. Both the US pinhead reconstruction precision error of 0.9 mm and the da Vinci® calibration of 1.5 mm are consistent with error for similar experiments in the literature of 1.2 mm and 1.0 mm respectively [15, 37]. The larger error in NGUAN may be because the gold standard used is optically tracked KeyDot® markers as opposed to an Optotrak® 3020 stylus (Northern Digital Inc., Waterloo, ON, Canada), which has a reported tip error of 0.25 mm [37]. The navigation aid was simulated in a finite element analysis and, relative to the tumour, did not deviate more than 1.0 mm than the expected distance. This is adequate for the purposes of providing guidance in the soft kidney. More advanced simulations, such as Camara et al. who simulates the kidney's deformation under an ultrasound scan using a particle-based approach [9] 1-2 mm error, may be integrated as well.

With the individual component errors being small, the measured total system error remains high at 5.1 mm especially when compared against the standard of care's recommended margin size of 5 mm. This is likely to be reduced through further system refinement and testing. The pinhead extension on the modified DART is not designed to be grasped by an instrument, and so imprecision in simply grab-

bing the tool could lead to added error. As well, it is not evaluated how accurate the manual pinhead segmentation is against the ground truth. These aspects are improved upon in the next chapter. As per the goals of the system, sparing tissue will be impacted by system error.

The single-surgeon/single-phantom study is primarily for feasibility. With it, NGUAN can be refined and improved. The DART is used to generate a tumour model, and provide guidance, without impeding the surgeon. Future studies are required with more trials of the system. This will provide more robust results than the single surgeon/single phantom study performed, as well as provide a clearer understanding on usability and preference. This is addressed in the next chapter.

In terms of usability, the NGUAN orthogonal virtual camera viewpoints are different to other image guidance systems for abdominal surgery. The advantage of the orthogonal viewpoints is that it provides the surgeon a perspective he or she would not normally have without occluding the surgeons view of the operative field. As well, because these viewpoints are displayed based off the tracking information, and not dependent on the video feed itself, there is no additional lag introduced. However, further work is required in NGUAN for the positioning of the views, as the surgeon had difficulty orienting himself relative to given views. Additional simplistic cues such as rendering the camera, showing the centre line axis of the virtual viewpoints or letting the surgeon set the pose of the virtual viewpoints could help minimise these issues. Using a colour gradient to represent the distance of the instrument to the tumour could improve the warning cue given to the surgeon as well. These augmented reality overlays are improved upon in the next chapter.

An evident critique of the DART include the line-of-sight requirement. In order to provide any guidance, the DART must be in the field of view of the laparoscope. This is acceptable during the planning stage, and early on in the execution stage. However, due to the manner in which the surgeon excises the specimen, he or she will zoom in close to the point of excision. He or she will often also lift the specimen up to try to see underneath it. During these steps, the DART may fall out of view. As well, blood may occlude the DART. This can be mitigated. For example, one could insert an additional DART into the side of the specimen, detect the new reference, and continue tracking. Alternatively, blood occlusion of the

DART pattern can alleviated through an omni-phobic coating to repel blood [41], or attempting to wash it with saline intra-operatively.

If one or multiple DARTs are used, inserting a barbed aid into the kidney yields a potential risk of seeding. According to preliminary tests, this may be prevented by using the stainless steel DART with electro-cautery. It is also feasible that a range of DART geometries be available to the surgeon depending on tumour depth: long barbs for deep tumours, short barbs for shallow tumours and adhesive fixation or tissue branding for superficial tumours. These could be determined based off pre-operative imaging.

The ultrasound scan is performed prior to the clamping of the renal artery to minimise warm ischemia time. However, with the change in perfusion pressure, the shape of the kidney is likely to change. Simulation and evaluation of the amount the kidney changes, and incorporating that into the provided guidance, will be required.

The DART and NGUAN offer many interesting avenues for future research. One novel addition would be the incorporation of surface reconstruction. This can be facilitated by structured light using, for example, laser-based or projector-based solutions. A reconstructed surface mesh could be displayed in orthogonal views to provide further depth cues. Furthermore, the surface could be used to provide the surgeon a true top-down view, as opposed to a view that is orthogonal to their camera viewpoint. Future work could also explore the use of the tumour model in intra-operative planning. These are addressed in Chapter 5.

Chapter 4

Improvements to NGUAN

The previous chapter introduced NGUAN, an augmented reality system that combines US and computer vision-based tracking and kinematics-based tracking to provide continuous real-time guidance during tissue excision. NGUAN is a largely standalone system composed of a surgical navigation aid called the DART and an US transducer, requiring no extrinsic tracking hardware. It leverages the da Vinci® as a development and testing platform. However, the initial iteration of NGUAN had significant shortcomings. Its systematic error was reported to be 5.1 mm, which is unacceptable given the standard of care for a margin size is considered to be 5 mm. The virtual viewpoints, despite making use of the 3D modeling and real-time tracking, are hard to interpret in a time constrained environment. Further, it did not provide guidance for a significant part of the surgery: identifying when to cut underneath the tumour. This challenge is a difficult one as, with the endophytic tumour and small size of the kidney, the surgeon risks cutting into the collecting duct. Therefore, while NGUAN is promising, improving its error and simplifying its augmentations would yield better utility. To that end, this chapter presents Nephrectomy Guidance Using Ultrasound-Augmented Navigation 2.0 (NGUAN+).

NGUAN+ uses the same principle of operation as NGUAN, but has been refined with four different augmentations. These include a proximity alert, an orientation cue, a simpler virtual viewpoint, and a projected path of the instruments. Further, this chapter also presents an intra-operative validation tool that can be used to assess augmentation accuracy during surgery. NGUAN+ is similarly evaluated

in simulated RAPN by an expert urologist, but has more trials to achieve statistical significance. This chapter is structured as follows: Section 4.1 discusses the use of an augmented reality validation tool for during surgery; Section 4.2 outlines the specific methodology changes to improve calibration and system accuracies; Section 4.3 discusses the new augmentations presented to the surgeon; Section 4.4 outlines the user study used to evaluate NGUAN+; Section 4.5 presents system error and study results; and Section 4.6 discusses the results and future work.

4.1 Intra-operative Validation Tool

The modified DART presented in the previous chapter is only used for system error evaluation. However, its utility can be applied intra-operatively for both calibration and validation. This requires a few refinements to the initial design. For distinction from the DART, this tool is referred to as the ballpoint stylus.

The previous design had a circular face with a 2.5 mm ballpoint that extended from the face. The KeyDot® marker is manually placed on the face. Because of this, the location of the ballpoint can only be estimated relative to the DART. Any errors in DART placement would propagate to the supposed ground truth ballpoint location. As well, the circular face is not designed to be easily grasped by the da Vinci® instrument. This limited the ability to assess the total system accuracy. To that end, the ballpoint stylus is printed entirely 3D printed (Proto3000, Vaughn, ON, CA) including the circle pattern. The pattern is printed in colour. The ballpoint stylus has known geometry up to the printing precision of the manufacturer. Proto3000 reports using the Statasys J750 printer, which has a printing resolution of 14 microns. The circular face is replaced with a portion of the DART including the repeatable grasp design. The ballpoint itself is increased to 3 mm in diameter to ease segmentation, and has slots the same size as a surgical instrument. It is more easily grasped, and when the surgical instrument grasps the ball tip, the instrument tip and ball tip are coincident. The ballpoint stylus and the completely 3D printed DART is seen in Figure 4.1.

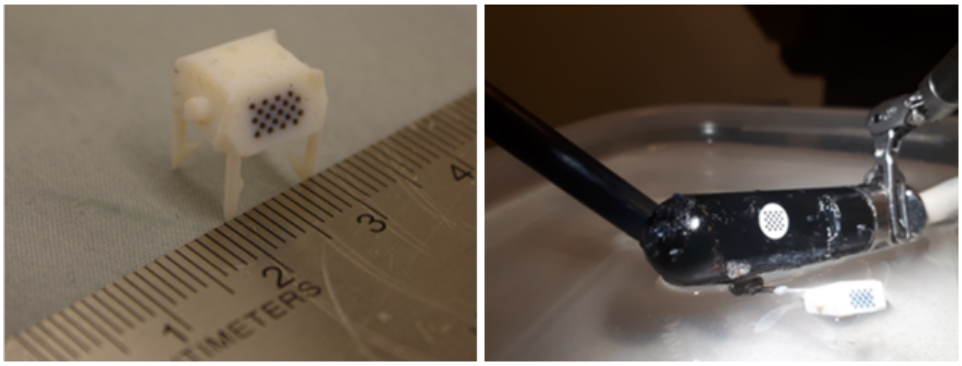

Figure 4.1: DART 3D printed in colour (left) and the ballpoint stylus being scanned (right).

4.2 Refinements to System Accuracy

4.2.1 da Vinci Laparoscope to Camera calibration

Changes in the algorithm and the design of the calibration stylus resulted in significant improvements in the laparoscope to camera calibration. Previously, Horn's algorithm is used. Horn's gives the transformation parameters between two corresponding sets of points (rotation, translation and scale) that minimises the mean squared error between the sets. However, Umeyama notes that this method may give an incorrect rotation [76]. Umeyama presents a refinement of Horn's method that, with his closed form solution, always presents the correct rotation [76]. For solving the laparoscope to camera transform, ${}^{C}T_{L}$, Umeyama's method is used.

The new ballpoint stylus is now used for calibration of the laparoscope to the camera. The same method for data collection of paired points is used. The stylus is moved to 23 unique poses. At each point, the stylus' origin in the camera is collected, and then the origin is touched with a surgical instrument. The instrument's pose is then collected. Instead of a leave-one-out approach to calculating TRE, a random set of 12 pairs are used to first calculate the ${}^{C}T_{L}$ transform. The fiducial registration error (FRE) for these 12 is reported. The resulting transform is then applied to the remaining 11 pairs, and the TRE is reported. FRE is reported to assess the registration accuracy, and TRE for guidance accuracy. While there is little to no

correlation between FRE and TRE [20], both are reported for completeness.

Finally, the previous chapter noted that ${}^{C}T_{L}$ is a combination of the laparoscope to camera transform and a correction for errors in tracking the instruments. It assumed that the single transform is valid for each of the patient-side manipulators used. In reality, there is a separate calibration required for the left and the right instrument. Both are calibrated for prior to the user study.

4.2.2 Total System Accuracy

With the improved ballpoint stylus, system error can be better assessed. This error can be determined by comparing the known center of the ballpoint, from the geometry of the CAD model, with the instrument's location when the instrument is grasping the stylus. This is captured when considering Equation 4.1 and Equation 4.2. Note that Equation 4.1 is the same as Equation 3.7 but repeated for convenience. Note that the subscript D, which previously represents the DART, represents the ballpoint stylus. The coordinate systems are interchangeable. With Equation 4.1, the segmented model's centroid can be registered to the ballpoint stylus. Comparing this centroid against the known ground truth assesses the error in vision-based tracking combined with reconstruction and segmentation. Then with Equation 4.2, the 3D location of the instrument (p_I) is transformed into the laparoscope coordinate system L, to the camera coordinate system C, and finally into the ballpoint stylus' coordinate system. By comparing the instruments location to known ground truth, it is possible to evaluate the error in the combined tracking. Moreover, when the ballpoint stylus is grasped by the instrument, and then has its ballpoint reconstructed, the two points of p_U and p_I should be equal, representing the ballpoint's centroid.

$$p_D = {}^{D}T_C \cdot {}^{C}T_K \cdot {}^{K}T_U \cdot p_U \tag{4.1}$$

$$p_D = {}^{D}T_C \cdot {}^{C}T_L \cdot {}^{L}T_I \cdot p_I \tag{4.2}$$

To evaluate these errors, the modified DART is held by an instrument in a water bath at room temperature. The ballpoint stylus is scanned, reconstructed, and segmented. The ballpoint stylus is moved to 10 poses, still held by the instrument.

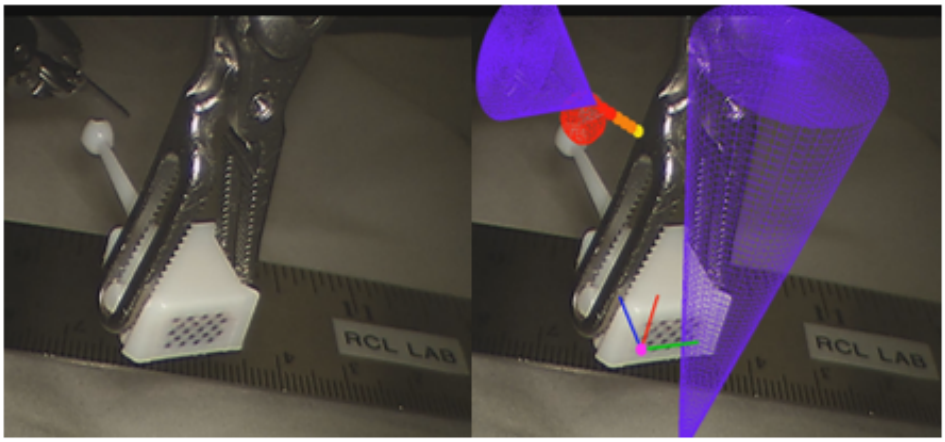

Figure 4.2: A comparison of the view without augmented reality(left) and with augmented reality (right). Red mesh model appears within 1mm of ground truth ballpoint stylus, and augmented reality overlays appear within 1mm of ground truth.

This makes $p_U = p_I$. The Euclidean distance between instruments location and the ground truth center is calculated in each pose, and the average is reported. An example of the resulting guidance is seen in Figure 4.2.

4.3 New Augmented Reality Overlays

Recall that the surgeon operates under a time constraint while trying to minimise tissue excised. Because of this, it is impractical to develop nuanced augmentations that cannot be quickly interpreted. While high fidelity overlays may be visually appealing, they are limited in utility if not intuitive and informative. Using this design consideration, four simple augmentations are proposed as seen in Figure 4.3. Augmentations are similarly provided to the surgeon using the TilePro® rather than interrupting the surgeons normal video feed or occluding the surgeons feed. They are as follows:

1. **Traffic Light**: a colour coded proximity alert of the instruments distance to the tumours surface shown as coloured blocks to the surgeon. The surgeon sets four ranges of distance of the instrument to the tumours surface. From

these ranges, the alert flashes red, yellow, orange, or green. These are provided as blocks of colour. For this thesis, the ranges are if the distance is less than 2.5mm, between 2.5mm and 3.5, between 3.5 mm and 5.0 mm, and beyond 5.0mm. A traffic light is provided for each of the two surgical instruments.

2. **Compass**: a conical overlay orienting the surgeon of his or her surgical instrument to the tumour. As the tumours in this work are endophytic, it is important to know the relative orientation of the tumour to an instrument at any given time, particularly if the instrument is past the tumour. A grey cone pointing from the instrument to the tumours center is provided, with the cones height proportional to instrument to tumour distance. The cone is occluded if the surgeon's tool is behind the tumour model.

3. **Projected Path**: a virtual needle-like extension with spheres of known diameter and spacing, also set by the surgeon. In Figure 4.3 and Figure 4.4, the spheres are all set at 1 mm apart, with 1 mm diameters. The functionality of the traffic lights are combined with the spheres, allowing the surgeon to gauge the distance of his or her instrument to the tumour should he or she continue in the current pose.

4. **Surface View**: the projected virtual scene from a virtual camera placed 50 mm away from the tracked aid, facing perpendicular to the grid of circles as seen in Figure 4.3. Treating the aid as a planar approximation of the local surface, the surgeon can then see tumour depth from the surface virtually. An example of this is seen in Figure 4.4.

The display as seen by the surgeon is captured in Figure 4.5. All four augmentations are given to the surgeon at the same time and it is up to the surgeon which augmentation to pay attention to.

Naively, determining the distance of the instrument's tip to the tumour's surface would require a comparison of the point to all points on the tumour's model. This would be a computationally expensive task. To provide real-time distance guidance, the augmentations here leverage a pre-computed signed distance field.

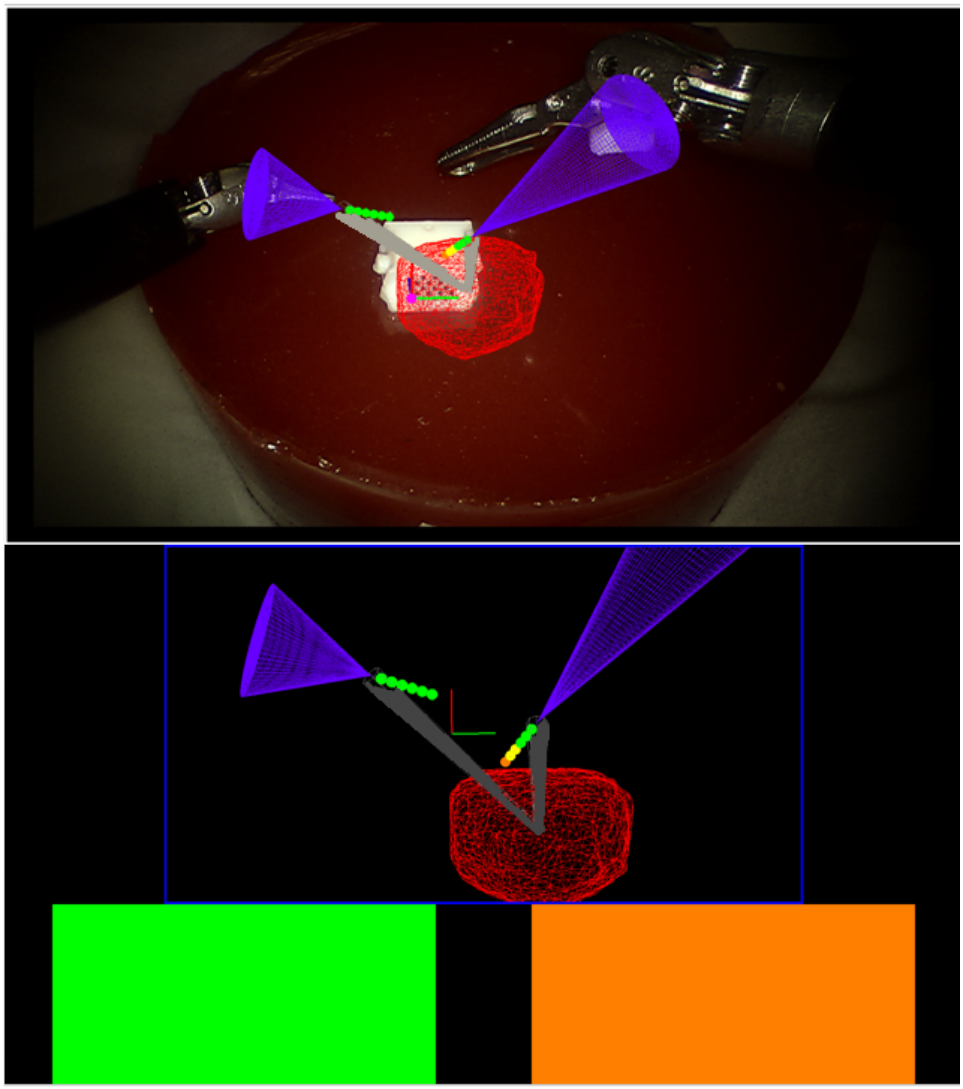

Figure 4.3: Left TilePro® with the augmented endoscopic view (top). Right TilePro® feed with virtual viewpoint and traffic lights (bottom). Compass overlay in grey, and projected path overlay for each instrument shown.

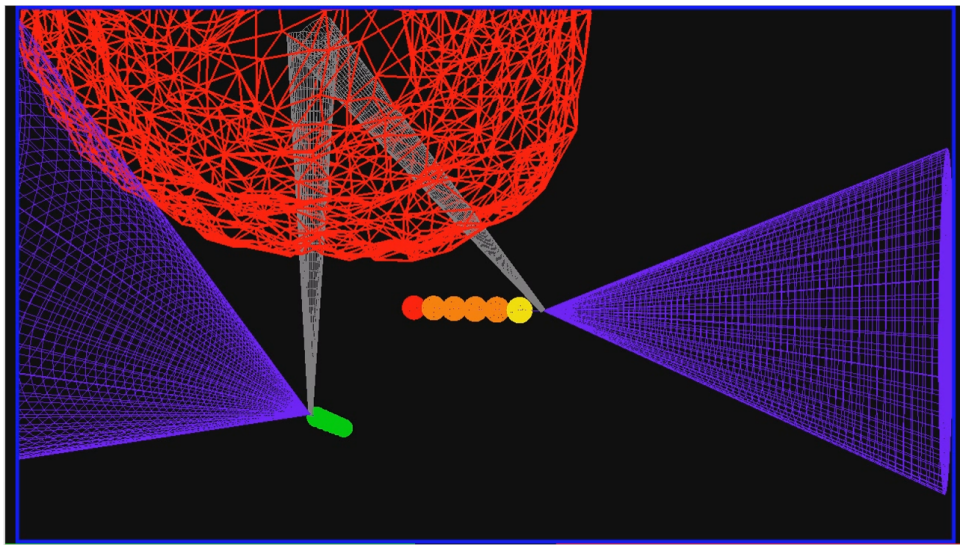

Figure 4.4: Magnified virtual viewpoint to show how the surgeon uses the guidance when close to the tumour underside. Red sphere indicates a distance within 2.5mm of tumour surface.

This field is computed after the tumour model is generated from US, and incorporates a 10 mm margin from tumour surface in each axis. Using a signed distance field reduces the complex calculation to a looking up an indexed value. It furthers captures irregularities in model topography, allowing for precise augmentations. This is particularly beneficial when the model is complex or contains additional structures.

4.4 User Study

The clinical utility of NGUAN is evaluated by having an expert perform simulated RAPNs. The participant is a practising urologist with over 10 years of experience and trained in performing RAPNs. This surgeon completes 9 nephrectomies using only laparoscopic US, as is the conventional method, and 9 with NGUAN+ for a total of 18. The simulated surgeries are performed on mock phantom models with elastic modulus of 15.4 kPA, similar to that of human kidney [22]. These models have 10 - 30 mm diameter black inclusions with graphite in them. This is to improve ultrasound contrast and to facilitate post-operative analysis by visual inspection.

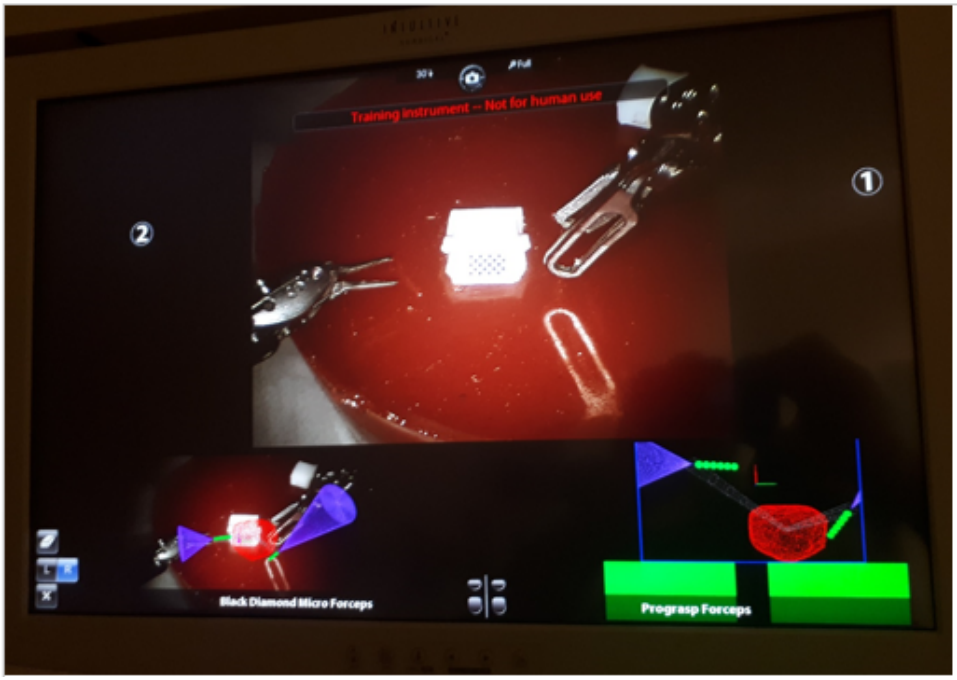

Figure 4.5: NGUAN+ as seen in the surgeon's console. Augmentations provided using TilePro®.

The location and depth of the inclusions are randomised. Prior to the study, each phantom model is scanned and the tumour models are generated. The segmented tumour model volumes and radii are compared against the ground truth from tumour construction to evaluate tumour model accuracy. The surgeon is able to train by using the ballpoint stylus. This allows them to interface with the segmented ballpoint model, understand the error in the system visually, and trust the system. The surgeon is given a practice surgery using only US as well. This training period is not timed or included in the results.

In each simulated surgery, the surgeon is instructed to scan the model's surface using the pick-up US transducer and, with a permanent marker, outline the tumour boundaries they observe. The surgeon then begins the excision stage. In the case of US, the surgeon has no additional guidance provided, while in the case of using NGUAN+ the surgeon has image-guidance throughout the excision.

For all surgeries performed, the excision times, margin status, excised and ad-

justed specimen volumes, specimen to tumour volume ratio, and the depth beyond tumour are reported. For qualitative feedback, the surgeon completed a Likert-scale questionnaire adapted from the System Usability Scale after each surgery [60]. After all the surgeries are completed, the surgeon is given open-ended questions to answer about his experience using the AR. A two-tailed paired t-test is performed for statistical significance with a power of 0.05. Holm-Bonferroni correction is used to account for multiple comparisons.

4.5 Results

The average (and standard deviation) known volume of tumours excised under guidance was 1.9 $\pm$ 0.4 cm^3, compared to the average segmented volume of 2.7 $\pm$ 0.7 cm^3. The average (and standard deviation) radius of the segmented models was 0.9 $\pm$ 0.3 mm greater than the ground truth radius. This indicates that the segmented models were slightly larger than the ground truth by only a millimeter.

In calibrating the laparoscope to the calibrated camera coordinates, the average and standard deviation FRE of using 12 points to determine the calibration transform is 0.8 $\pm$ 0.3 mm. Evaluating the determined calibration on a separate set of 11 paired points, the average and standard deviation TRE is 1.0 $\pm$ 0.4 mm. The working volume covered is 45 $\times$ 30 $\times$ 50 mm.

Total system error is defined to be the Euclidean distance between the tracked instrument's tip compared against the ground truth center of the ballpoint stylus. This requires the stylus to be scanned, reconstructed, and registered with the instrument. Over 10 poses, the average and standard deviation of the total system error is found to be 2.5 $\pm$ 0.5 mm. When comparing the instrument's tip against the segmented model, rather than the ground truth, the average and standard deviation distance between them is 1.4 $\pm$ 0.5 mm.

The quantitative results of the surgeries performed are summarised in Table 4.1. These initial results show that, with no statistically significant difference in excision time, the surgeon is able to excise significantly less tissue with NGUAN+ than without. The known tumour volumes excised with US and augmented reality were not significantly different, nor is there a significant difference in positive margin rate. Note however that the positive margins with US were gross margins that left

Table 4.1: Quantitative results of simulated partial nephrectomies as average and standard deviation. Average and standard deviations (avg $\pm$ stdev) of each metric is listed. Results of the trials using ultrasound only (US) and augmented reality (NGUAN+) are shown. Bold indicates statistical significance ($p < 0.05$). Bold asterisk indicates statistical significance ($p < 0.05$) of augmented reality compared to the US only.

Metric (avg $\pm$ stdev	**US (n=9)**	**NGUAN+ (n=9)**
Excision Time (secs)	203 ± 30	257 ± 50
Margin Status (/ 9)	2 gross	1 microscopic
Known Tumour Volume (cm^3)	2.4 ± 1.0	1.9 ± 0.4
Excised Tissue Volume (cm^3)	30.6 ± 5.5	**17.5 ± 2.4***
Adjusted Tissue volume (cm^3)	22.1 ± 5.2	**10.6 ± 2.1***
Depth Beyond Tumour (mm)	10.2 ± 4.1	**3.3 ± 2.3***

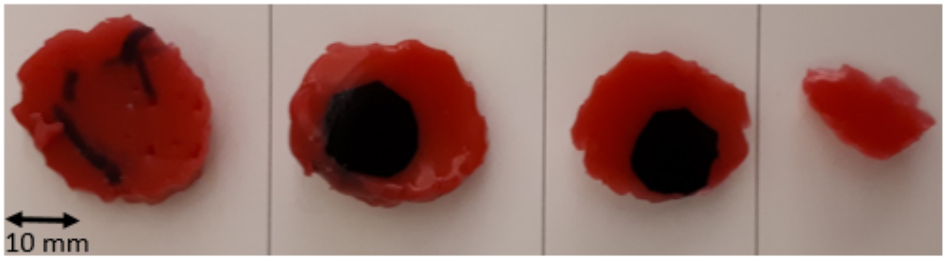

Figure 4.6: Cross section of tumour excised with augmented reality guidance. Slice closest to the surface on the left, farthest on the right.

significant amounts of tumour behind. The single positive margin achieved with augmented reality is considered microscopic, with a small amount of tumour exposed and with no visible tumour left behind. Importantly, with AR, the surgeon is able to significantly reduce the depth past the tumour from approximately 10mm to 3mm.

Figure 4.6 shows some example cross sections of specimens excised with AR, each approximately 5 mm thick. Table 4.2 summarises the qualitative metrics from the Likert-scale questionnaires. When asked to rank the augmented reality overlays from most to least preferred, the expert indicated he strongly preferred the projected path, then the traffic lights, the compass, and finally the virtual viewpoint.

Table 4.2: Qualitative metrics with the questions asked about the augmented reality system. Score reported where 1 = strongly disagree and 5 = strongly agree.

Question Asked	**Score (avg $\pm$ stdev**	**Degree of Agreement**
I found the system unnecessarily complex.	1.3 ± 0.5	Strongly Disagree
I thought the system was easy to use.	4.8 ± 0.5	Strongly Agree
I imagine most people would learn to use this system very quickly	4.8 ± 0.5	Strongly Agree
I found this system cumbersome to use.	1.3 ± 0.5	Strongly Disagree
I felt very confident using this system.	4.7 ± 0.7	Strongly Agree
I needed to learn a lot of things before I could get going with the system.	1.0 ± 0.0	Strongly Disagree
I felt I understood where my region of interest was spatially.	4.8 ± 0.7	Strongly Agree
I felt I had a good understanding of the relative distance from my tool to the tumour	4.8 ± 0.7	Strongly Agree
I felt I was not at risk of cutting into the tumour	4.6 ± 0.7	Strongly Agree
The system meets my needs.	4.6 ± 0.7	Strongly Agree

4.6 Discussion

This chapter presents an improvement to the novel intra-operative US-based augmented reality system known as NGUAN+. The total system error is significantly reduced to 2.5 $\pm$ 0.5 mm, which is acceptable with a 5 mm margin as the standard of care. This system meets the accuracy requirement to be useful in guidance. Augmented reality is beneficial in this study in resecting the lateral edges of the specimen. It is informative in determining the point to cut underneath the tumour and is considered essential in guiding the deep resection through tissue. The augmented reality is noted as being predictable when it would and would not appear (due to occlusion of the DART). This is beneficial, as the surgeon could understand why no guidance is presented at times and how to resolve it, but also frustrating. This line of sight issue could be mitigated with the use of multiple aids added during excision.

Specifically considering the virtual renderings of the instruments themselves, the surgeon noted it is useful to have them even though the small registration error is noticeable. This misalignment is in fact the approximately millimeter error of laparoscope to camera calibration, which is enlarged given the laparoscope's field of view and distance to the tools. The surgeon found it useful as he is able to mentally adjust for the error because he also understood where the physical alignment should be.

Using the projected path and its incorporated traffic light, the surgeon adopted a check and go strategy, a minor modification on his traditional approach to excision. With this strategy, he paused during cutting and checked his tools surroundings. At various points where the spheres were hard to see or his instruments were occluded, the traffic lights were used as a proxy. Counter-intuitively, this modified strategy did not significantly increase the excision time. Speculatively, this may be in part due to the overall reduced amount that needed to be excised and improved surgeon confidence. However, further experimentation and testing is required to validate this hypothesis. With respect to the virtual viewpoint, the surgeon elaborated that, although useful in concept, it is difficult to quickly interpret and mentally register to the scene while under a time constraint. In an untimed stage of the surgery, like the planning stage, a virtual viewpoint may be beneficial. The surgeon's perception of

depth is still limited with the augmentations provided. For example, the projected path, which copies the laparoscopic view and renders on top of it, is created using a single camera feed, contrary to the surgeons 3D stereoscopic video feed. This can be improved using TilePro® to provide a 3D stereo overlay.

While the study is still small, with a single user performing 18 surgeries, it does demonstrate the feasibility of using tracked US to create continuous guidance with encouraging results. The surgeon is able to use the NGUAN+ system to significantly reduce the amount of healthy tissue excised, at no increase to excision time. It is particularly valuable in reducing the risk of cutting into the collecting system, as noted by the significantly smaller depth cut under the tumour of 3.3 ± 2.3 mm. Of all four augmentations, the surgeon used the projected path the most as it mimicked his real environment more closesly than the others.

This ballpoint stylus is useful during the operation as a validation device for the augmented reality. As Bernhardt et al. notes, there is a current need to validate the accuracy of augmented reality during surgery without exposing the physical structure being modeled [6]. For endophytic tumours, validation would require exposing the tumour to then see how well the augmented reality aligns - a predicament. That said, with the grasp design, the surgeon can easily drop in and pick up the ballpoint stylus in a reliable manner. The ballpoint's circles pattern can be tracked, and the surgeon can observe his augmented reality with respect to the ballpoint stylus. With the ballpoint rendered, the surgeon can see the error in segmentation. The surgeon can interact with the model and the physical ballpoint, and gain an understanding of the system error. If there appears to be a significant error, re-calibration is warranted, which can be facilitated intra-operatively by the same tool through its trackable pattern and repeatable grasp.

The guidance presented in NGUAN+ is powerful. With it, there is the potential for a truly minimally invasive approach to partial nephrectomies. The surgeon noted that with this guidance, one could preserve the top layer of parenchyma almost entirely. This new approach would start by making a single incision above the tumour, retracting it with the da Vinci's additional arm, and inserting the DART into the gap. The model generation steps are performed as described. Then, the surgeon could leverage the guidance to core out the tumour itself. Rather than achieving an ideal resection of a cylindrical shape, the true ideal resection is a

teardrop shape. Upon tumour excision, the reconstruction phase of the RAPN would be significantly reduced. Minor reconstruction may still be required, but the large defect in the conventional approach would no longer exist. Reconstruction may become as simple as suturing itself. This approach would not only minimise the nephrons further (by preservation of a layer of tissue) but also reduce the risk of reaching the 25 minute threshold for warm ischemia time. Investigation into the feasibility of this approach is warranted.

However, while this system is excellent for the excision stage, there is a requirement for guidance during the planning stage. The surgeon may use the guidance presented in this chapter as a safety measure, ensuring he or she does not cut into the tumour, but it does not inform the surgeon's initial resection. The ideal excision approach is not easily known to the surgeon. In particular, because of the shallow angle of the laparoscope and its limited range, the surgeon cannot explore their workspace with ease to find the best angle to resect. Augmented reality during the planning stage potentially yields additional benefits. As well, since the model is generated during the planning stage, it is a natural extension to provide guidance during it. This challenge is addressed in the following chapter.

Chapter 5

Projector-based Augmented Reality Intra-corporeal System

The previous two chapters introduced the framework of a US-based augmented reality guidance system. They presented two sets of overlays, and initial evaluation of the guidance during the excision stage of a RAPN. However, these works did not assist the surgeon in the planning stage. As the tumour model is generated during this stage, and there is no time constraint, it is natural to provide guidance for intra-operative planning to improve surgical outcomes.

This chapter proposes a novel guidance system called Projector-based Augmented Reality Intra-corporeal System (PARIS) that provides guidance and addresses the issue of initial resection angle for the surgeon. PARIS uses a miniaturised projector within the patient's abdomen for both reconstruction of the kidney surface and as a method to augment the surgical scene itself. Then, using the DART, the projector can project the 3D tumour model back onto the scene in a surface corrected manner. Different visualizations can be performed with the projector, and presented from either laparoscope point-of-view (LPOV) or projector point of view (PPOV).

The chapter is structured as follows: Section 5.1 describes the surgical challenge of the initial resection angle; Section 5.2 describes the miniaturised projector used in PARIS; Section 5.3 covers the principle of operation and transformation theory behind PARIS; Section 5.4 outlines the steps involved in providing projector-

based augmented reality and the augmentations in this work; Section 5.5 describes the calibration of the system's components; Section 5.6 outlines the initial and secondary user studies performed; Section 5.7 presents the accuracy and user study results; and finally Section 5.8 recaps the work done and limitations.

5.1 The Challenge of Resection Angle

For endophytic tumours, the ideal resection approach has two components: location and direction. The ideal location is to begin where the tumour is closest to the organ surface, minimising the layer of healthy parenchyma excised above the tumour. The ideal direction is to cut straight down from this point, specifically down the normal of the surface. The surgeon would like to begin his or her resection by cutting down the orthographic projection of the tumour on the surface. For a spherical tumour, this ideal excision would be encompassed in a cylinder with diameter equal to the tumour's diameter.

With this in mind, and using the conventional US-only approach, this ideal approach is unlikely to be achieved. That is because the surgeon does not definitively know where the ideal location to begin is. US may indicate the depth of the tumour, but it still lacks information on the approach angle. Additionally, the surgeon may inadvertently hold the transducer off from being perpendicular to the kidney surface, leading to a misinterpretation of the real location to start. Finally, the US is again only temporary.

Further, when considering augmented reality and providing guidance to the surgeon, the conventional manner has been to augment the laparoscope's video. This creates augmentations in the laparoscope's point of view (LPOV). However, the laparoscope is often not positioned in the ideal manner for resection, directly above the tumour and perpendicular to the surface. The laparoscope is often at an acute angle relative to the surface. While the surgeon may be able to "see through" the surface with augmentations, this is differs from the ideal approach. Further, the surgeon's laparoscope is at a different angle than his or her instruments. When considering the set of possible angles that can be achieved by each one, the laparoscope's set is smaller than that of the instruments. This makes the mental registration between the two for the purpose of resection difficult. Simply, it is

hard to reach a target if your viewpoint is significantly different than the tool you're using to reach it. This leads to the hypothesis that the surgeon's initial resection can be improved by providing augmented reality from a different point of view. With the use of an intra-corporeal projector, the surgeon can potentially explore a wider range of angles from the PPOV. With the same device, one can then augment the scene and obtain beneficial guidance.

5.2 The Pico Lantern

Projector-based augmented reality is an appealing display modality as it augments reality itself. On the spectrum of reality-virtuality, projections are considerably closer to reality than the computer graphic-based equivalents. Using a projector for display is an alternative strategy that has yet to be explored within the patient's abdomen. Conventional computer graphics are frequently superimposed onto the laparoscope's video feed, appearing as if separate and floating on top. These superimposed rendering are not effected by changes in the scene itself, such as lighting conditions, and provide poor depth perception. The use of a projections on the other hand can provide convincing augmentations that blend naturally with the scene. To that end, the intra-operative and intra-corporeal projector used in this work is called the Pico Lantern, created by Edgcumbe et al. for surface reconstruction and augmented reality in laparoscopic surgery [16]. The terms projector and Pico Lantern are used interchangeably in this work.

This works uses updated hardware over the initial prototype, although the concept is similar. The projector is a modified PicoPro projector (Celluon Inc., Seoul, Korea) with a KeyDot® marker placed on it. Using the same pose estimation as described previously, this projector can be tracked relative to a single camera. No exogeneous tracking hardware is needed. Further, the projector requires no interposition between the laparoscope and the surgeon's console. The projector's laser raster scanning enables it to have a large range of focus [16]. This model has more than double the resolution (1920×1080) and brightness (30 lumens) compared to the original prototype with 640×480 resolution and 15 lumens [16]. It additionally has wireless capabilities, and is compatible for Android, iOS, and Windows. The Pico Lantern requires no dedicated port as it can placed through the skin in-

cision with a thin cable beside the trocar or controlled wirelessly [16]. It, like the DART and US transducer used, has a custom grasp that can be reliably picked up using the da Vinci® [16].

As an improvement to the initial prototype, this prototype has the KeyDot® marker which may be perpendicular to laparoscope. The motivation for this came from initial experimentation on accuracy and geometry constraints of having a projector in the field of view of the camera. Having the marker on the face of the projector requires the majority of the projector to be in the field of view. By moving it to an extension, the projector itself can be outside the field of view. As well, KeyDot® tracking has limited tracking accuracy when parallel to the laparoscope image. By moving the KeyDot® from a parallel to perpendicular arrangement, the tracking stability improves.

In order to leverage the projector, additions are made to the framework. First, the TilePro® feeds output from the PC are removed entirely, as the projector is independent of the surgeon's console. Then, the projector had to be additionally calibrated to model its intrinsic parameters, and the theoretical projector origin relative to the KeyDot® on it. The projector can be modeled using the pinhole camera model used for the camera calibration, as one can treat the projector as a camera in reverse. While the specifics of projector calibration are outside the scope of this work, the projector has a similar set of intrinsic parameters in (f_x, f_y, c_x, c_y) and a set of distortion parameters as a calibrated camera. Additionally, the projector calibration results in determining the transformation of the KeyDot® marker coordinate system (M) to the projector's coordinate system (P). The marker and the projector are seen in Figure 5.1, while M and P are illustrated in Figure 5.2. This is done using the Projector-Camera Calibration toolbox [17].

5.3 Projector-based Augmented Reality Intra-corporeal System

PARIS follows a similar but expanded principle of operation as NGUAN and NGUAN+. The DART is placed above the tumour using an freehand US scan. A tracked US scan is then taken relative to the DART, and manual segmentation produces a 3D tumour model. PARIS differs in creation of the augmented reality overlays, described in

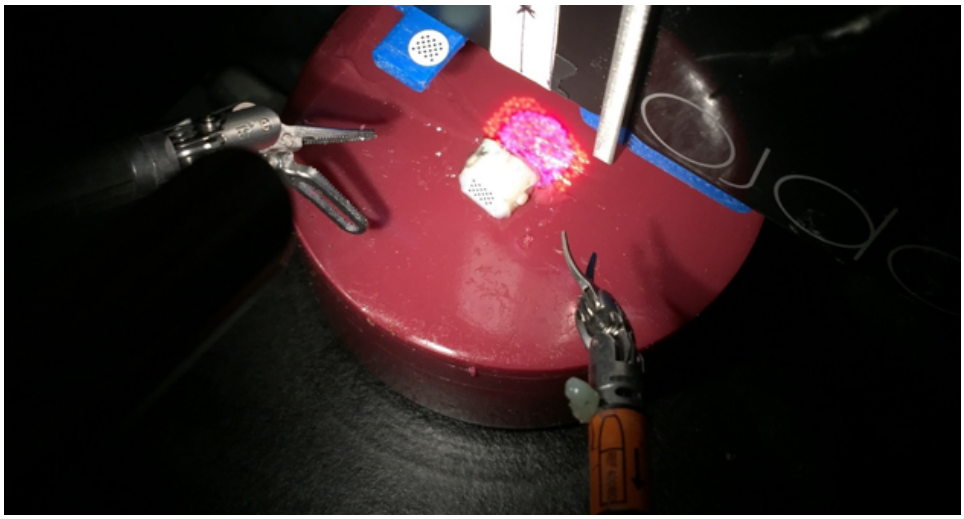

Figure 5.1: The system setup for PARIS. The projector is used to augment the tumour's surface. The scene is viewed by a stereo laparoscope.

the next section. To create the augmented reality overlays, a surface reconstruction must be performed, described in Section 5.3.1. Note that PARIS does not provide continuous guidance during the excision itself, unlike NGUAN and NGUAN+. This is because of the computational complexity required to calculate the augmentations. The projections created rely on the surface reconstruction of the scene, as discussed later, which is not real-time.

As the projector is dynamic and can be moved, it is best to move it as close to the ideal location for the resection angle. Placing the tracked projector near this is relatively easy by aligning the projection image's center (drawn as a dot) with the tumour's centroid as seen by the projector. During the process, the projector is kept approximately normal to the surface via careful manual positioning. The system set up is seen in Figure 5.1.

In order to project accurate guidance, the different components of PARIS must be registered to each other. As illustrated in Figure 5.2, PARIS uses the following coordinate systems which carry over from NGUAN, and described in Section 3.4:

- U: the 2D US image.
- K: the KeyDot® marker on the US transducer itself.

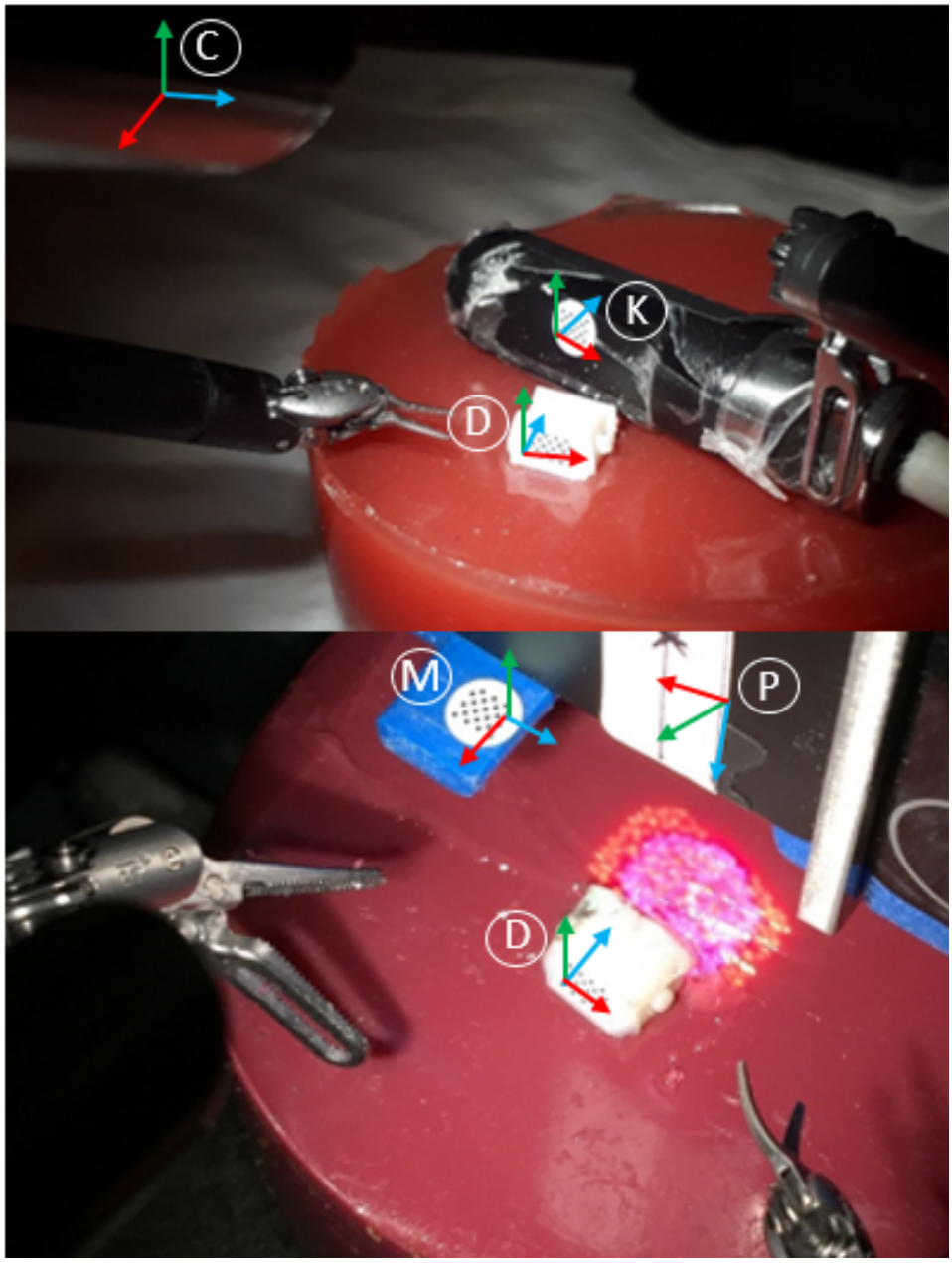

Figure 5.2: Coordinate systems used within PARIS. Tracked US scan is performed relative to the DART (top). Tracked and calibrated projector augments the scene with the tumour model (bottom).

- C: the calibrated camera's 3D coordinate system.
- D: the KeyDot® marker on the DART itself. Same coordinate description as in K.

PARIS introduces the following additional coordinate systems:

- P: the projector coordinate system. The 3D origin of the projector lies within the device. The axes are the same as the C, where X-axis goes left to right of the image, Y-axis goes top to bottom, and Z-axis goes out of the projector.
- M: the KeyDot® on the projector. It is placed on the surface of the projector, and shares the same coordinate description as K and D.

The calibration of ${}^{K}T_{U}$ is done as described previously in Section 3.6.1. The transforms of ${}^{C}T_{K}$, ${}^{C}T_{D}$, ${}^{C}T_{M}$ are found through pose estimation as described in Section 3.2.1. The transform of ${}^{P}T_{M}$ is found through projector calibration. With these transforms, it is then possible to create projection-based augmented reality.

5.3.1 Surface Reconstruction

After the tumour model is generated, a surface reconstruction must be done to provide augmented reality. The projected image, when perceived by the laparoscope, must appear accurately. To do this, the projected image must be pre-distorted to account for surface topography. Doing this will cause the projections to appear normal from the perspective of the laparoscope, and therefore the surgeon.

OpenCV has multiple implementations for stereo surface reconstruction: block matching on the CPU (BM), block matching on the GPU (BMGPU), semi-global block matching on CPU (SGBM), belief propagation on GPU (BPGPU), and constant space belief propagation on GPU (CSBPGPU). Evaluation of these algorithms is completed by the Advanced Research Team at Northern Digital Inc. Processing each algorithm on the Middlebury Stereo Datasets showed that BM is the fastest but poorest quality, CSBPGPU produced irregularities, and SGBM is the best choice considering availability, quality and speed. However, SGBM is not real time and challenging to parallelise on a GPU [26].

In order to use SGBM, stereo cameras must be used [26]. The stereo camera calibration parameters can be used to rectify the camera images. Rectification distorts

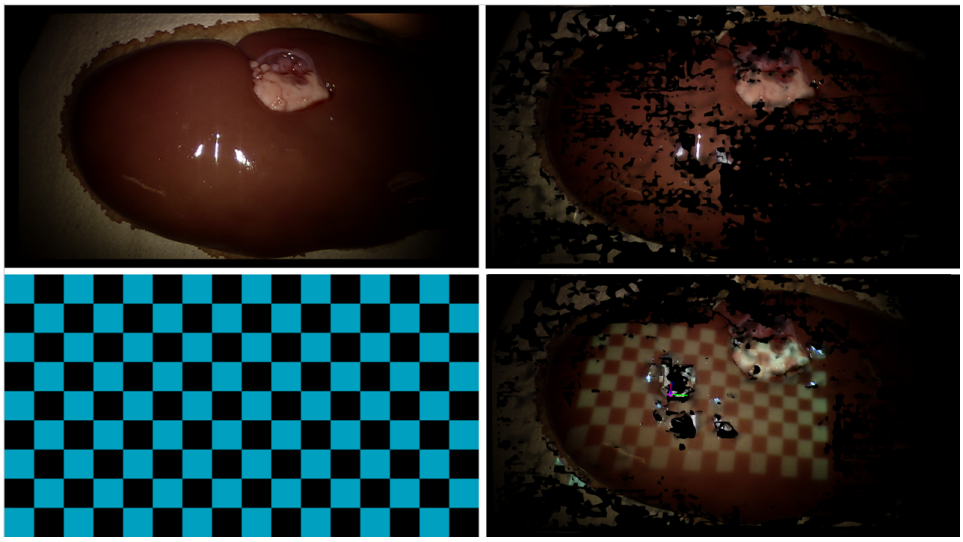

Figure 5.3: *Ex-vivo* kidney seen by the laparoscope with no projection on it with a relatively featureless surface (top left). The ideal reconstruction would match this image perfectly. A typical surface reconstruction using SGBM and no additional features (top right). Note the black spots are holes in the reconstruction. The checkerboard pattern projected onto the scene (bottom left). The additional features improve the surface reconstruction by a perceptible amount (bottom right). The two holes in the middle are due to specular reflection and the DART, which also causes reflection.

the images such that matching points in the images will lie along corresponding epipolar lines. This simplifies the search for matching points into a 1D search problem. The details of epipolar geometry is not covered in this work. Using the rectified images, SGBM becomes more efficient [26]. Normally, the disparity between matching points can be used to estimate a point's depth in a stereo image pair which tends to be noisy. SGBM improves upon this by combining global and local matching methods and matches small regions of the image [26]. The result is an algorithm that sufficiently balances speed and accuracy, with relatively good robustness against noise. However, in the case of narrow baseline stereo cameras like in the laparoscope, these surface reconstructions may not be sufficiently dense for guidance. To improve the method, the projector projects a checkerboard pat-

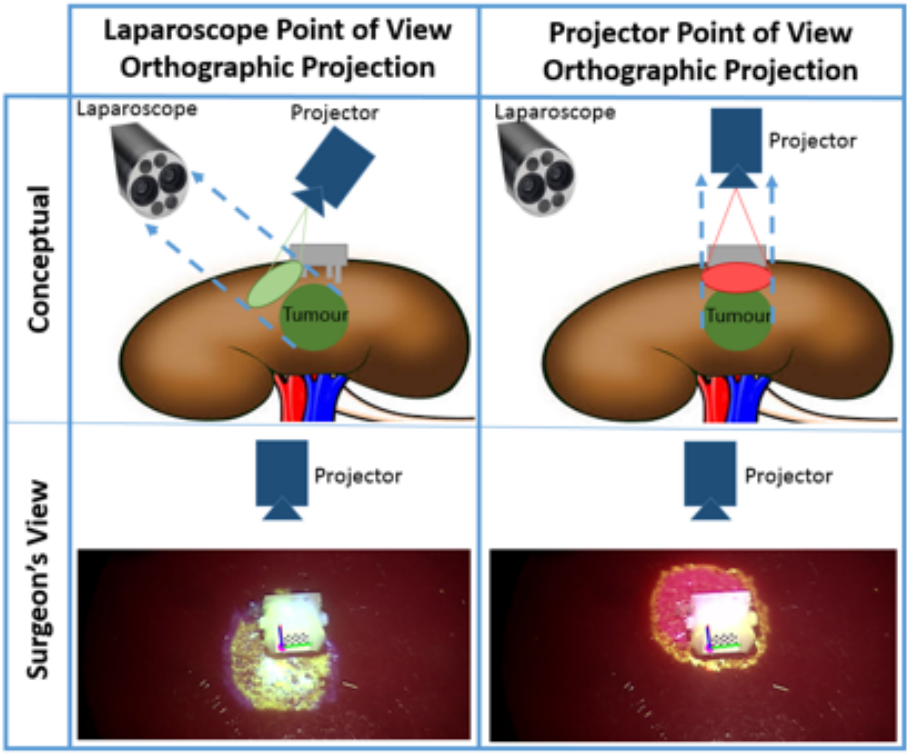

Figure 5.4: Overview of PARIS. Light green indicates orthographic projection from LPOV (left). Red indicates projection from PPOV (right).

tern to add additional features into the scene. An example comparison is seen in Figure 5.3. Surface improvement is evaluated in Section 5.5.

5.4 Augmented Reality Overlays

5.4.1 Projector and Laparoscope Point-of-Views

The augmented reality overlays can be split into two independent categories: projection type and projection point of view (POV). The two types of projection explored in this work are orthographic and perspective. The two types of POVs explored are LPOV and PPOV. All of these are displayed with the projector. Note that

the term perspective and point-of-view refer to distinctly different things, and are not interchangeable in this chapter.

As mentioned, there are limitations to the laparoscope's possible angles relative to the surface. Using a dynamic and mobile projector makes it more likely to identify the best approach angle. The two different augmentation POVs are illustrated in Figure 5.4.

Augmentations from the LPOV are straightforward. Every point in the tumour model which is in the DART coordinate system, p_D, is transformed into a point in the camera's coordinate system, p_C, as in Equation 5.1. Each p_C is then transformed onto the camera's imaging plane. The resulting image is displayed by the projector and requires no tracking of the projector. As the projector moves, the projection remains the same. As the DART or camera move, the tumour model's appearance as seen by the camera has changed. This results in an update in the projected image.

$$p_C = {}^{C}T_D \cdot p_D \tag{5.1}$$

In order to provide augmentations from PPOV, the tumour model must be transformed into the projector's coordinates. Assuming the US data has already been transformed into the DART as a result of the reconstruction and segmentation, the DART must then be registered to the projector. A point in the DART coordinate system, p_D, can be transformed into the camera coordinate system by ${}^{C}T_D$, then to the projector's marker coordinate system by ${}^{M}T_C$, and finally into the projector's coordinate system by ${}^{P}T_M$. This results in p_P as seen in Equation 5.2.

This is captured in Equation 5.2.

$$p_P = {}^{P}T_M \cdot {}^{M}T_C \cdot {}^{C}T_D \cdot p_D \tag{5.2}$$

From here, the tumour model can be displayed with the projector as if the projector is a camera. Using the intrinsic and distortion parameters, the 3D model viewed by the projector is reduced to its 2D image, and then projected onto the scene. This idea is similar to the virtual camera concept. As the camera moves, the projection remains the same. As the DART or projector move, then the tumour model's appearance as seen by the projector has changed. The projection image is updated to reflect this.

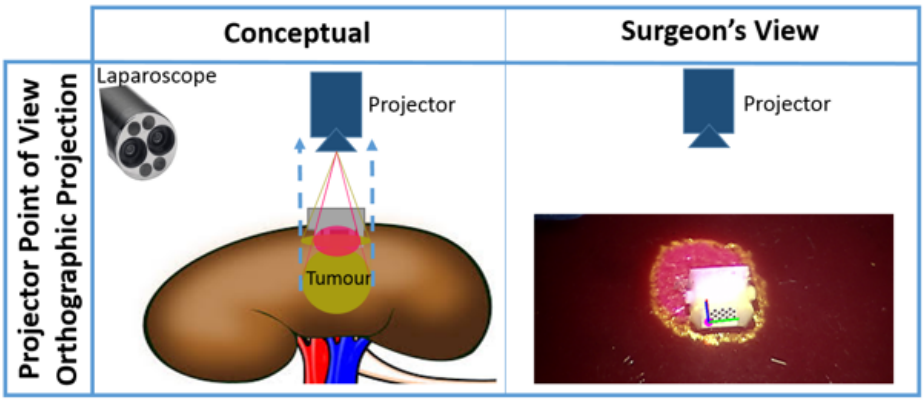

Figure 5.5: The PPOV visualization of PARIS. Red indicates perspective projection, and yellow/brown indicates orthographic projection. Both seen from the projector POV.

5.4.2 Orthographic and Perspective Projections

There are different methods in illustrating a 3D model as a 2D planar image. These primarily are organised as parallel projection (of which orthographic is a subset) and perspective projection.

With orthographic projection, parallelism between lines is maintained when projected onto the 2D imaging plane of a camera. To perform the orthographic projections, the rays are projected in parallel from the tumour towards the desired POV.

Perspective projection is akin to how humans observe the world. The rays of light converge into the eye, rather than stay parallel. This often results in one or more vanishing points, where parallel lines in 3D appear to intersect. It may also result in foreshortening, where the model appears shorter due the angle of the viewer, or uneven scaling.

With orthographic projections, the projected image is determined by the intersection of the rays with the surface. With the LPOV perspective projection, the projection image must be pre-distorted to account for the kidney's non-planar topography. If viewed without this correction, the projection image will only appear correct from the projector and warped from all other viewpoints. Both the orthographic projection and pre-distortion leverage the same ray-surface intersection

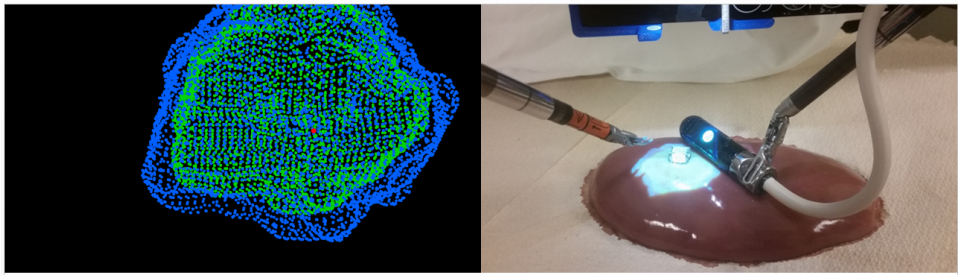

Figure 5.6: Example projection image for LPOV projections (left) and its appearance on *ex-vivo* kidney. The tumour model is pre-distorted, hence the irregular shape.

algorithm described in the next section.

5.4.3 Overview of Ray-Surface Intersection

The tumour model is first transformed from the DART coordinate system into the camera's coordinate system, registering the surface and model together.

For orthographic projections, a set of rays, R, is created. Each ray r in the set begins at one of the tumour model's vertices and all end at either the camera's origin or the projector's origin, depending on the POV. Similarly, for the surface correction for LPOV perspective projection, R is created. Here, each ray r begins at the camera's 3D origin and ends at one of the tumour model's vertices.

The surface mesh is modeled as a set of triangles, S. Each triangle, s, can be used to define a 3D plane A which has a normal vector n_A.

For all rays in R, each r is tested for intersection with each triangle s in the surface's set of triangles S. This is done by first computing if the ray intersects with the triangle or not. The existence of an intersection can be confirmed by computing the dot product of the plane's normal and the ray's direction, d_r, as in Equation 5.3. If this equation is true, then a single point of intersection exists which can be computed. If this equation is not true, then the ray either does not intersect with this triangle or it intersects completely (line intersection). Neither case is of interest here.

$$d_r \cdot n_A \neq 0 \tag{5.3}$$

If the intersection point does exist, it must be determined whether it lies within the given triangle s or not. This can be done using the geometric approach of Moller and Trumbore [48]. Here, the triangle is represented using two parameters (u, v) space and the three edges. A point in the triangle lies within the triangle's barycentric coordinates. Details can be found in [48]. This efficiently finds the points that intersect between a given ray r and a given triangle s. This process is repeated for all r in R and all s in S. Multi-threading is used to improve the computational speed to <1 second. In the future, this would be properly translated to leverage the GPU given the parallelizable nature of the problem.

5.4.4 Summary of Augmented Reality Overlays

In summary, there are four projected images. They are the perspective and orthographic projection from LPOV and the perspective and orthographic projection from PPOV. In order to create these projection images, a surface reconstruction is performed, which is supplemented by the projector. A conceptual diagram for the orthographic projections in both POVs is seen in Figure 5.4, while a conceptual view of the perspective and orthographic projection is seen in Figure 5.5. An example of pre-distortion for the LPOV is seen in Figure 5.6.

It should be noted that perspective projections is briefly evaluated and found to give less intuitive guidance. The original intent is to use them in conjunction with the orthographic projection such that the surgeon could gauge the depth as the relative difference between the two. However, given the small scale, this distance is too minute to interpret. Standalone, the perspective projection presents an incorrect view on the surface as the tumour model appears smaller than it actually is. In order to use the perspective projection alone, the surgeon would have to cut down in a conical manner. This is confusing. As such, perspective guidance is not explored further. In all augmentations, the model is projected as a dense point set.

Collins et al. presented a similar notion of presenting augmented reality from a different POV, primarily using the instrument's port [13]. This is centered on the port itself, rather than the ideal position for excision. Furthermore, they specifically use perspective projection, not orthographic and their chosen POV is fixed and determined geometrically, not variable and dynamically found. The assess-

ment of resection quality is ambiguous as well, with no quantitative measure of tissue excised.

5.5 System Calibration and Accuracy

While it is possible to perform surface reconstruction with only the stereo laparoscope cameras, a naive algorithm will produce a reconstruction that is not dense enough. To that end, the ability of the Pico Lantern to improve surface density is measured. Surface density is defined as the percentage of the surface in the laparoscope view that is reconstructed. By projecting a simple pattern, not only is the laparoscope's field of view more illuminated, there are additional features that can be matched. A simple checkerboard pattern is used here which produces an abundance of corner features. The surface density of an *ex-vivo* kidney, with and without the extra features projected, is compared for 12 unique laparoscope and projector poses. Average surface density change is reported.

Next, the accuracy of the projector's ability to display is quantified. The re-projection error of the projector here is similar to that of camera calibration. Referring to Equation 5.2, the detected origin of the DART can be transformed into the projector's coordinate system. Using the projector's intrinsic parameters, the DART origin can be projected onto the projector's imaging plane and drawn as a dot. The actual projection of this image onto the scene should result in the dot perfectly overlaid on the DART origin. The re-projection error is the Euclidean distance between the real origin and the projected origin. This captures the cumulative error in the tracking of the DART, the camera calibration, tracking the of projector's KeyDot®, and the projector's calibration. To measure this error, the projector is moved to 5 poses, and for each pose the DART is placed in 10 poses, approximately 80 mm from the laparoscope. In each pose combination, the laparoscope image is taken as it views the DART and the projected point. The two points are manually segmented. The RMS error is reported. This experiment is seen in Figure 5.7.

Next, the system's ability to localise and project a tumour model accurately must be quantified. Even if the re-projection error described above is accurate, one must evaluate the cumulative error of creating a tumour model and projecting it with PARIS. To that end, the tumour models themselves are first evaluated.

Figure 5.7: Un-augmented cross-section of phantom (left). Computer graphics overlay of tumour model (right). LPOV perpsective projection of model.

For each phantom used in the study, the segmented tumour model volumes and radii are compared against the ground truth values. These ground truth values are determined during phantom construction.

Then, a phantom model is cut in half to expose the endophytic tumour. Using the RENAL score, these tumours are among the most challenging the surgeon will face due to their depth. Exophytic tumours (those with majority protruding beyond the surface) are relatively easier to excise with their lower complication rate [78]. A tracked US is taken at the face of the exposed tumour, and is segmented. This produces a tumour model that is a single slice, rather than the complete model itself. This is done because it is imprecise to cut the phantom in half. The single slice tumour model is then reprojected by the projector onto the scene. The average RMS distance and the Hausdorff distance between the contours of the actual tumour and projected tumour, as seen by the laparoscope, for five laparoscope and projector poses is reported. The Hausdorff distance represents the maximum distance between a point in one contour to a point in the other contour.

5.6 User Studies

Two user studies were performed to evaluate PARIS and its clinical utility for intraoperative surgical planning. In the first study performed, one expert urologist completed three simulated partial nephrectomies for three different visualization modes for a total of nine simulated operations. The first mode is the conventional US only approach. The second mode is using orthographic projection from LPOV and US

imaging. The third is using orthographic projection from PPOV and US imaging. The surgeon is given US in all trials as, in clinical use, the surgeon would have it for planning. The surgeon can validate the visualization using US as well. The surgeon is given a single practice trial for each visualization, which are not included in analysis. After the practice is completed, the surgeon performs the 9 simulated surgeries in randomised order.

In the second study performed, a novice surgeon is added. The novice is a second year urological resident with no training on the da Vinci®. The novice and expert each complete a set of mock surgeries in one day long session. The surgeries are completed using the US only approach and the orthographic projection from PPOV visualizations. Each surgeon is given a practice trial with both visualizations which are, as above, not included in analysis. The novice surgeon completed 6 surgeries in each visualization for a total of 12. The expert surgeon completed 10 surgeries in each visualization for a total of 20.

For both studies, the average excision time, margin status and size, and average specimen to tumour volume ratio are reported for each visualization mode. To quantify the surgeon's deviation from the ideal excision, the excised specimen is cut into approximately 5mm thick slices. The tumour outline and the cross section contour are segmented manually. The RMS distance between the centroids and the Hausdorff distance are reported. Each surgeon also answered a Likert-scale questionnaire, which inquired about confidence and spatial understanding, and provided open feedback at the end of all surgeries.

5.7 Results

5.7.1 System Calibration and Accuracy

The surface density using a projected pattern in the scene improved by an absolute average of $15.4 \pm 8.3\%$. This increase in surface density is sufficient the calculation of projection images.

The projector's re-projection error onto the DART origin is determined to be 0.8 RMS. During the data collection, the projector is moved over a range of $32 \times 9 \times 11$ mm in the laparoscope coordinate frame.

Table 5.1: Quantitative comparison of simulated partial nephrectomies performed in the first PARIS study. Average and standard deviations (avg $\pm$ stdev) of each metric is listed. Results of the trials using ultrasound only (US), augmented reality from the laparoscope point of view (LPOV), and augmented reality from the projector point of view (PPOV) are shown.

Metric (avg $\pm$ stdev)	**US (n=3)**	**LPOV (n=3)**	**PPOV (n=3)**
Execution time (secs)	188 ± 10	180 ± 28	178 ± 27
Tumour volume (cm^3)	2.9 ± 0.2	2.4 ± 0.4	3.1 ± 0.7
Adjusted Excised Volume (cm^3)	12.3 ± 0.9	11.0 ± 3.4	11.1 ± 3.2
Positive Margins (/3)	3	3	2
Hausdorff Distance (mm)	11.7	12.6	8.4
Centroid Distance (mm)	5.3	5.3	2.4

The average tumour model and ground truth volumes are 2.8 ± 0.7 cm^3 and 4.0 $\pm$ 1.0 cm^3 respectively. The difference between measured and ground truth radii is 1.2 mm RMS. For the projection of the tumour onto the actual tumour, the average Hausdorff distance and RMS distance between the contours are 3.9 mm and 1.7 mm respectively.

5.7.2 First User Study

The quantitative results of the first study's nine simulated surgeries are summarised in Table 5.1. There are positive margins in all trials except for one instance with PPOV. The average execution time and specimen to tumour volume ratio are consistent acss the surgeries. However, the Hausdorff distance of the cross sections improved using PPOV. With PPOV, the surgeon is able to achieve 8.4 mm, indicating a more consistently tight margin compared to 11.7 mm for US only.

Qualitatively, the surgeon felt PARIS generated a clear sharp image that blended well with the phantom surface. The depiction of the tumour felt natural and intuitive. Drawbacks included the need for moderate ambient surgical light intensity to avoid poor contrast of the projection, and a need for guidance during the excision itself. The surgeon reported that he would use PARIS over the conventional US approach partly because the augmented reality provides persistent guidance. Conventionally, the surgeon observes a limited cross-section during the US scan. In comparing the two POVs, the questionnaire indicate he felt more confident using

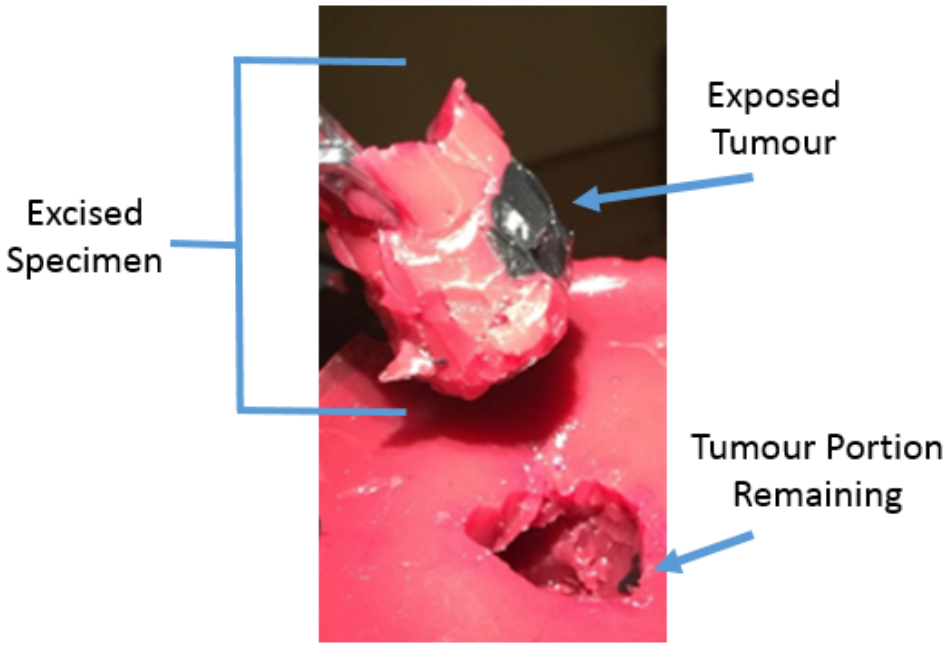

Figure 5.8: Example of a positive margin with both tumour exposed and a portion remaining in the phantom, indicated with blue arrows.

PPOV (4.3 ± 1.2 compared to 1 ± 0). He further felt had a better spatial understanding of tumour depth (3 ± 0 compared to 1.7 ± 1.2). Lastly, he felt with LPOV he needed significant additional information. The reason LPOV mode is dropped for the second user study is because the surgeon had difficulty aligning his mental model from the US scan with what he observed in LPOV. This discrepancy likely caused the gross positive margins, as seen in Figure 5.8.

5.7.3 Second User Study

The quantitative results from the 36 simulated RAPNs is summarised in Table 5.2. The novice surgeon is able to achieve 1/6 positive margins in both US and PPOV visualizations. The expert achieved 2/10 and 0/10 for US and PPOV respectively. Both surgeons excised a statistically significant less amount of healthy kidney tissue. A Wilcoxon signed-rank test, used to evaluate significance, resulted in $p < 0.05$. Fur-

Table 5.2: Quantitative comparison for second PARIS user study performed. Average and standard deviations (avg ± stdev) of each metric is listed. Results of the trials using ultrasound only (US) and augmented reality from the projector point of view (PPOV) are shown. Bold asterik indicates statistical significance ($p < 0.05$).

	Novice	**Novice**	**Expert**	**Expert**
Metric (avg ± stdev)	**US (n=6)**	**PPOV (n=6)**	**US (n=10)**	**PPOV (n=10)**
Execution Time (secs)	579 ± 155	469 ± 152	199 ± 31	207 ± 40
Positive Margins	1/6	1/6	2/10	0/10
Known Tumour Volume (cm^3)	2.8 ± 0.7	2.6 ± 0.7	2.6 ± 0.8	2.4 ± 0.9
Excised Tissue Volume (cm^3)	26 ± 3	**17 ± 3***	20 ± 4	**14 ± 4***
Hausdorff Distance (mm)	19.3 ± 0.8	13.3 ± 3.7	18.0 ± 2.2	11.0 ± 1.7
Centroid Distance (mm)	5.1 ± 1.5	4.1 ± 1.8	4.4 ± 1.9	2.9 ± 1.2

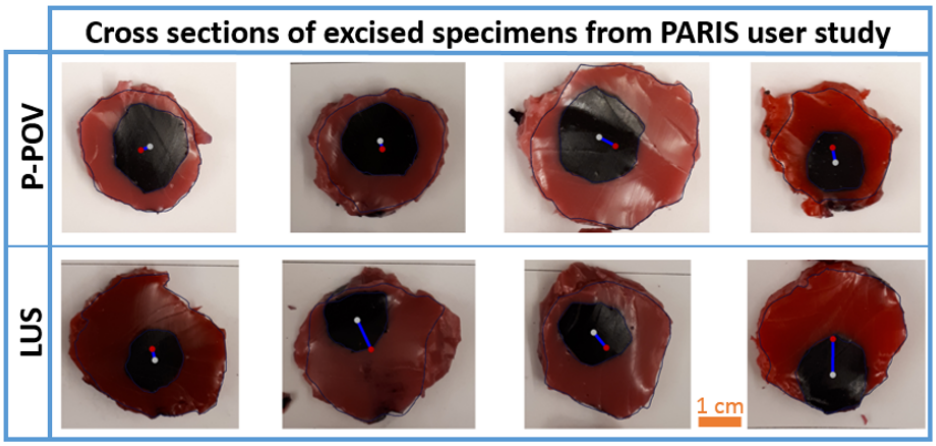

Figure 5.9: Example cross sections of excised specimens using PPOV (top row) and US (bottom row).

ther, example cross sections of the excised specimens are in Figure 5.9. For the novice surgeon, the use of PPOV improved the Hausdorff distance from 19.3 mm to 13.3 mm. For the expert surgeon, the use of PPOV improved the Hausdorff distance from 18.0 mm to 11.0 mm. This indicates a more consistently tight margin with the use of PPOV. In comparing the US and PPOV, the results indicate the surgeons felt more confident (3.3 ± 1.1 to 5.0 ± 0.0) and had a better spatial understanding (3.5 ± 0.8 to 4.6 ± 0.5). These results favour PPOV over US visualization.

5.8 Discussion

This work presents a novel fully-integrated intra-corporeal augmented reality system for intra-operative guidance in soft tissue surgery. Given a 5 mm margin for partial nephrectomy, the error of the subsystems (1.2 mm RMS for the tumour model geometry and 0.8 mm RMS error for dynamic marker re-projection) and the overall tumour localization error (1.7 mm RMS) are small enough to consider PARIS beneficial for guidance.

This work demonstrates that integration of the three components (the DART, the US transducer and the Pico Lantern), vision-based tracking and PPOV projector augmentation are all feasible and practical. Integration of PARIS with the da Vinci® solely requires read-only access to the video feed which eases dissemination. With only one practice trial, PARIS could already be successfully used for guidance, showing that training time is short. The projectors augmentation is clear enough on the surface to be useful in planning the excision. The surgeons indicated that, unlike standalone US imaging, it is helpful to have the projector provide persistent guidance after US scanning. The trade-off is that the projector must be balanced with the surgical light's intensity as to give sufficient contrast to the projection itself.

Key discoveries include: the clear preference of the PPOV orthographic projection over the LPOV; PPOV is an effective visualization, and that orthographic projection parallel to the direction of excision is a good strategy for navigation. With the LPOV, the tumour projection is placed along the line-of-sight to the surgeon, but presents the well-known challenge of depth perception. Moreover, this LPOV is not typically the desired direction for resection. The direction of excision and associated parallel orthographic projection is usually perpendicular to the surface but theoretically could be in any direction that provides a short path and avoids critical anatomy. Such advanced guidance would be implemented by adapting one of the many graphics techniques described in the surgical guidance literature to the projector. In a more general sense, the PPOV is akin to have an eye in the hand of the surgeon, as explored by others, so that the moving projection image gives valuable dynamic visual cues to the surgeon.

The projector, although being adjunct hardware, is a relatively low cost device

that can enhance surface reconstruction density (seen here to be vital in the creation of augmented reality) and that can display realistic augmentations. With the mobile and dynamic projector, the surgeon explores his or her workspace with overlays only possible with the use of a projector. In an *in-vivo* setting, the effect of the anatomy, such as the peritoneum and perinephric fat, on the projections will have to be explored and compensated for.

The first study has a small scope and size, limiting the ability to make statistically significant conclusions. However, the main outcome is preliminary evidence that the PPOV orthographic projection aided the surgeon in centering the tumour within the ideal excision boundaries in all excisions performed with that method. Note in this case, the surgeon is instructed to produce a zero-margin resection in an effort to judge the ultimate accuracy of the different guidance methods. Thus, positive margins are expected and are not indicative of actual surgery using 5 mm margins.

The second study is able to compare two users with different knowledge levels and how PARIS can insist them. Using PARIS, a surgeon is able to achieve a statistically significant reduction in healthy tissue excised regardless of experience. The surgeons also felt more confident and had a better self-reported spatial understanding of the underlying anatomy.

This work concludes that PARIS is a simple, easily integrated system with potential to provide valuable guidance with sufficient ease and accuracy in laparoscopic surgery. The dynamic marker provides a tumour-centric reference for relative measurements of the USTransducer and projector to minimise errors. The dual use of the projector for additional features and guidance information is feasible. This guidance is an adjunct, not replacement, to standard practice. Further study is needed to demonstrate utility *in-vivo*, where the challenges of bleeding, smoke, and specular reflections arise, and whether or not parallax needs to be accounted in using the stereo laparoscope.

Chapter 6

Conclusion and Future Work

This chapter reviews the work proposed in this thesis. The contributions, limitations, and future work are described as well.

6.1 Author's Contributions

This thesis has presented three augmented reality systems, NGUAN, NGUAN+, and PARIS. The systems stem from the idea that, with US-based augmented reality, the inherent challenges of laparoscopic surgery can be mitigated intra-operatively and the surgeon can improve upon the current standard of care.

These systems leverage computer-vision based tracking of a fiducial marker, and operate in the same manner. After inserting the DART, a tracked ultrasound scan is performed, yielding a reconstructed volume. The tumour of interest is segmented from this volume, and the resulting model provides guidance to the surgeon in both the intra-operative planning and excision stage. The systems were evaluated using simulated RAPNs, a type of laparoscopic surgery.

The first system, NGUAN, showed the technical feasibility of using intra-operative ultrasound to generate augmented reality. It showed that the DART concept could fit within the work space of the surgeon, and that it is possible to provide guidance in spite of tissue deformation. NGUAN+ used a mix of medium and low fidelity augmentations to produce significant improvements in surgery. NGUAN+ was able to reduce the depth of the cut beyond the tumour. Then a novel intra-corporeal

projector system in PARIS illustrated, for the first time, the use of a projector to augment the kidney itself and inform the surgeon. PARIS was able to position the surgeon in a better spot and improve his or her ability to intra-operatively plan. With these systems, the surgeon gains an unprecedented information in which to explore and perceive his or her workspace.

This work required engineering to do the following:

- interface with the ultrasound machines,
- calibrate the ultrasound transducer to its tracked marker,
- design and development of the DART,
- track and register one or multiple fiducials,
- register the kinematics-based tracking of the da Vinci® with the computer-vision based tracking of the DART,
- track and control a projector,
- render augmentations including virtual viewpoints, surgical instruments,
- perspective and orthographic projections, point clouds and meshes, and
- evaluation of the systems; and the design, implementation and analysis of user studies

There were limitations. For the systems and their evaluations, these limitations can be categorised into the system and its components, the principle of operation, and evaluation of the system.

6.1.1 System and Components

While the components of the systems have been thoroughly evaluated, the obvious limitation is that the system evaluation has only been on plastic PVC phantoms. The porcine model, shown to be suitably representative for humans, would be the next step [70].

Considering the components of the systems it should be noted that alternative ultrasound calibration methods may improve results while projector and camera calibrations were suffice,. Najafi et al. [50] illustrated a ultrasound calibration technique which produced a submillimeter pinhead reconstruction accuracy, an order of magnitude more accurate than the geometric calibration performed here [50]. As well, in considering the fact a stereo camera pair was available, using triangulation may improve the system accuracy error [24, 35].

For the framework to be broadly used, it must be extended to include commonly used ultrasound transducers and machines. Aloka and BK Ultrasound machines are used at our local center at VGH. Provided that these machines can send a video signal out, or either the ultrasound image itself or the entire screen, then they can be supported. Transducers can similarly be supported, but doing so will require the generalisation of the pose estimation. With the planar fiducials, only transducers with flat faces can be used. 3D printing transducer covers or exploring model-based tracking methods may overcome this.

In the case of PARIS, the surface reconstruction algorithm used is not the best available in the field. It is a readily available algorithm that, when compared to a small set of other algorithms, gave the best trade off between speed and accuracy. Improving the choice of stereo reconstruction algorithm used is warranted. Further, in the original Pico Lantern paper, the use of the projector was in simulating a wide baseline stereo reconstruction with only a monocular laparoscope and the projector. Making this approach real-time and automatic would be suited to make PARIS more accessible.

6.1.2 Principle of Operation

The ultimate goal is that guidance in NGUAN, NGUAN+ and PARIS is used clinically *in-vivo*. To achieve that goal, the issues of ultrasound scanning and segmentation, tissue deformation, renal artery clamping must be addressed in the overall framework.

For simplicity, manual segmentation was used in experiments. This is time consuming and incorporates additional human error given the difficulty in interpreting ultrasound images. In practice, segmentation time can be reduced by using

(semi-)automatic methods. Alternatively, simply approximating the tumour segmentation with a bounding sphere may suffice, and be highly efficient. That said, *in-vivo* automatic segmentation of tumours is more difficult than segmentation of phantoms. In regard to reliably scanning the full tumour, knowing the tumour's size from pre-operative imaging may be beneficial. Using this, the surgeon's ultrasound could be augmented with a "bounding circle" to guide them in maximizing how much tumour he or she capture.

The guidance used in this work treats the DART and tumour as rigid bodies that, from FEM simulations, are relatively rigid to one another. However these simulations did not simulate the kidney during excision itself, where severe shearing and tissue tearing occurs. The forces during excision are likely greater than during planning. To that end, incorporating a deformable model which can register continuously to the laparoscopic image would be beneficial. As well, the DART is also used as a means to circumvent surface tracking of the featureless kidney. However, through the use of the Pico Lantern to project additional features and the real-time tissue tracking presented by Yip et al. [85], these challenges may be mitigated.

As with any navigation system, one must consider the characterization of the system itself. While NGUAN and NGUAN+ are both real-time, limited only by the frames per second of the laparoscopic video, PARIS is not real-time. PARIS in fact has a latency of 300 ms. This is largely attributed to the computation of ray-surface intersection.

The tracking performed is dependent on the markers. These markers are unobstrusive, safe for use in the patient, and provide full 6-DOF tracking with sufficient accuracy and robustness. However, in all systems, the number of tracked instruments depends on the number of markers in the scene. This work does not explore more than two markers at any given time. Due to the topographical filtering performed during pose estimation, it remains to be seen whether mutliple markers can be tracked at once. Ways to mitigate this issue would be variance in marker size and colour. The ability to track flexible instrumentations is also a remaining issue. The working volume of the systems have been shown to be adequate and simulate a patient's abdomen.

6.1.3 Evaluation

The phantoms presented for here contained only a single endophytic tumour. A human kidney will contain branches of the renal artery known as segmental arteries. These traverse the kidney and surgeons simply cut them, which requires reconstruction after tumour excision. Modeling arteries into these phantoms, and observing how the guidance could help the surgeon understand his or her location (or avoid them all together) would be beneficial.

The qualitative questionnaires used in this thesis were primarily adapted from the System Usability Scale. However, none of the qualitative feedback received considered or measured the mental, temporal and physical demands on the surgeon or more situational specific constraints. To that end, in future studies, the Surgery Task Load Index should be used [82]. This validation measure was developed using fundamentals of laparoscopic surgery peg transfer task, and is well suited for the evaluation of introducing interventions into the surgical environment [82].

6.2 Future Work and Recommendations

What was not discussed was the potential of fluorescence imaging in identifying tumour boundaries. Firefly™ is an integrated real-time fluorescence-imaging mode that is available for the da Vinci® system. It uses near-infrared fluorescence imaging after the injection of the indocynanine green contrast agent. Such imaging may be an excellent compliment to the guidance described in this work. By design, the images are co-registered with Firefly's camera, simplifying registration of yet another component. With Firefly, one can assess the presence of blood vessel structures, and whether the kidney has been adequately clamped off from blood. In the case of dense perinephric fat, Firefly may have difficulty imaging, but US may be of benefit as it can image through said fat [52]. Regardless, the imaging of blood vessels with either modality may be useful in providing real-time localization and guidance.

While developed and tested using robot-assisted partial nephrectomy as a model procedure, this system and its concepts is extensible to conventional laparoscopic surgery and can be applied other soft-tissue surgeries such as the hepatectomy or prostatectomy. With the exception of the stereo surface reconstruction in Chap-

ter 5, the rest of the instrumentation for these systems can translate well into the conventional laparoscopic approach. Addition of fiducial markers onto the laparoscopic instruments can replace the need for kinematics-based tracking. The general workflow of the system does not require significant engineering input to perform, and so this work is relatively low barrier to go from laboratory settings to clinical usage. Robust evaluation to observe the significance of augmented reality on long-term outcome, operation time, and usability are needed.

In conclusion, the systems here are innovative approaches to surgical navigation for minimally invasive surgery that can be broadly disseminated. Doing so, based off the studies here, has the potential to make challenging cases more feasible and reduce the learning curve for performing the surgery. The work has the potential to increase the availability of procedures. This in turn can increase the number of patients that undergo laparoscopic surgery, improving patient care at a large scale.

Bibliography

[1] O. Akca, H. Zargar, R. Autorino, L. F. Brandao, H. Laydner, J. Krishnan, D. Samarasekera, J. Li, G.-P. Haber, R. Stein, et al. Robotic partial nephrectomy for cystic renal masses: a comparative analysis of a matched-paired cohort. *Urology*, 84(1):93–98, 2014. → pages 20

[2] H. O. Altamar, R. E. Ong, C. L. Glisson, D. P. Viprakasit, M. I. Miga, S. D. Herrell, and R. L. Galloway. Kidney deformation and intraprocedural registration: a study of elements of image-guided kidney surgery. *Journal of endourology*, 25(3):511–517, 2011. → pages 33

[3] 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. → pages 29

[4] R. Autorino, A. Khalifeh, H. Laydner, D. Samarasekera, E. Rizkala, R. Eyraud, R. J. Stein, G.-P. Haber, and J. H. Kaouk. Robot-assisted partial nephrectomy (rapn) for completely endophytic renal masses: a single institution experience. *BJU international*, 113(5):762–768, 2014. → pages 20

[5] M. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *ACM SIGGRAPH Computer Graphics*, volume 26, pages 203–210. ACM, 1992. → pages 30

[6] 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. → pages 7, 8, 10, 27, 29, 30, 32, 33, 74

[7] M. Borghesi, E. Brunocilla, A. Volpe, H. Dababneh, C. V. Pultrone, V. Vagnoni, G. La Manna, A. Porreca, G. Martorana, and R. Schiavina.

Active surveillance for clinically localized renal tumors: an updated review of current indications and clinical outcomes. *International Journal of Urology*, 22(5):432–438, 2015. → pages 17

[8] J.-Y. Bouguet. Camera calibration toolbox for matlab. 2004. → pages 42

[9] M. Camara, E. Mayer, A. Darzi, and P. Pratt. Soft tissue deformation for surgical simulation: a position-based dynamics approach. *International journal of computer assisted radiology and surgery*, 11(6):919–928, 2016. → pages 58

[10] C. Cheung, C. Wedlake, J. Moore, S. Pautler, and T. Peters. Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 408–415, 2010. → pages 30, 31

[11] J. E. Choi, J. H. You, D. K. Kim, K. H. Rha, and S. H. Lee. Comparison of perioperative outcomes between robotic and laparoscopic partial nephrectomy: a systematic review and meta-analysis. *European urology*, 67 (5):891–901, 2015. → pages 6, 20

[12] T. Collins, A. Bartoli, N. Bourdel, and M. Canis. Robust, real-time, dense and deformable 3d organ tracking in laparoscopic videos. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 404–412. Springer, 2016. → pages 40

[13] T. Collins, P. Chauvet, C. Debize, D. Pizarro, A. Bartoli, M. Canis, and N. Bourdel. A system for augmented reality guided laparoscopic tumour resection with quantitative ex-vivo user evaluation. In *Computer-Assisted and Robotic Endoscopy: Third International Workshop, CARE 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 17, 2016, Revised Selected Papers*, volume 10170, page 114. Springer, 2017. → pages 88

[14] B. J. Dixon, M. J. Daly, H. Chan, A. D. Vescan, I. J. Witterick, and J. C. Irish. Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. *Surgical endoscopy*, 27(2):454–461, 2013. → pages 34

[15] P. Edgcumbe, C. Nguan, and R. Rohling. Calibration and stereo tracking of a laparoscopic ultrasound transducer for augmented reality in surgery. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 258–267. Springer, 2013. → pages 58

[16] P. Edgcumbe, P. Pratt, G.-Z. Yang, C. Nguan, and R. Rohling. Pico lantern: Surface reconstruction and augmented reality in laparoscopic surgery using a pick-up laser projector. *Medical image analysis*, 25(1):95–102, 2015. → pages 78, 79

[17] G. Falcao, N. Hurtos, J. Massich, and D. Fofi. Projector-camera calibration toolbox. *Erasumus Mundus Masters in Vision and Robotics*, 2009. → pages 79

[18] S. S. Fenton, D. E. Schaubel, M. Desmeules, H. I. Morrison, Y. Mao, P. Copleston, J. R. Jeffery, and C. M. Kjellstrand. Hemodialysis versus peritoneal dialysis: a comparison of adjusted mortality rates. *American Journal of Kidney Diseases*, 30(3):334–342, 1997. → pages 21

[19] I. Figueroa-Garcia, J.-M. Peyrat, G. Hamarneh, and R. Abugharbieh. Biomechanical kidney model for predicting tumor displacement in the presence of external pressure load. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, pages 810–813. IEEE, 2014. → pages 33

[20] J. M. Fitzpatrick. Fiducial registration error and target registration error are uncorrelated. In *SPIE Medical Imaging*, pages 726102–726102. International Society for Optics and Photonics, 2009. → pages 64

[21] H. Fuchs, M. A. Livingston, R. Raskar, K. Keller, J. R. Crawford, P. Rademacher, S. H. Drake, A. A. Meyer, et al. Augmented reality visualization for laparoscopic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 934–943. Springer, 1998. → pages 27, 28

[22] N. Grenier, J.-L. Gennisson, F. Cornelis, Y. Le Bras, and L. Couzi. Renal ultrasound elastography. *Diagnostic and interventional imaging*, 94(5): 545–550, 2013. → pages 39, 41, 68

[23] C. Hansen, J. Wieferich, F. Ritter, C. Rieder, and H.-O. Peitgen. Illustrative visualization of 3d planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5(2): 133–141, 2010. → pages 29, 33

[24] R. I. Hartley and P. Sturm. Triangulation. *Computer vision and image understanding*, 68(2):146–157, 1997. → pages 99

[25] M. Hew, B. Baseskioglu, K. Barwari, P. Axwijk, C. Can, S. Horenblas, A. Bex, J. de la Rosette, and M. L. Pes. Critical appraisal of the padua classification and assessment of the renal nephrometry score in patients undergoing partial nephrectomy. *The Journal of urology*, 186(1):42–46, 2011. → pages 16

[26] H. Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2):328–341, 2008. → pages 82, 83

[27] A. Hughes-Hallett, E. K. Mayer, H. J. Marcus, T. P. Cundy, P. J. Pratt, A. W. Darzi, and J. A. Vale. Augmented reality partial nephrectomy: examining the current status and future perspectives. *Urology*, 83(2):266–273, 2014. → pages 27

[28] A. Hughes-Hallett, E. K. Mayer, H. J. Marcus, P. Pratt, S. Mason, A. W. Darzi, and J. A. Vale. Inattention blindness in surgery. *Surgical endoscopy*, 29(11):3184–3189, 2015. → pages 34

[29] A. Hughes-Hallett, E. K. Mayer, P. Pratt, A. Mottrie, A. Darzi, and J. Vale. The current and future use of imaging in urological robotic surgery: a survey of the european association of robotic urological surgeons. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 11(1):8–14, 2015. → pages 26, 27

[30] A. Hughes-Hallett, P. Pratt, J. Dilley, J. Vale, A. Darzi, and E. Mayer. Augmented reality: 3d image-guided surgery. *Cancer Imaging*, 15(1):O8, 2015. → pages 31

[31] A. Hughes-Hallett, P. Pratt, E. Mayer, M. Clark, J. Vale, and A. Darzi. Using preoperative imaging for intraoperative guidance: a case of mistaken identity. 2015. → pages 27

[32] S. Isotani, H. Shimoyama, I. Yokota, T. China, S.-i. Hisasue, H. Ide, S. Muto, R. Yamaguchi, O. Ukimura, and S. Horie. Feasibility and accuracy of computational robot-assisted partial nephrectomy planning by virtual partial nephrectomy analysis. *International Journal of Urology*, 22(5): 439–446, 2015. → pages 29

[33] U. L. Jayarathne, E. C. Chen, J. Moore, and T. M. Peters. Freehand 3d-us reconstruction with robust visual tracking with application to ultrasound-augmented laparoscopy. In *SPIE Medical Imaging*, pages

978617–978617. International Society for Optics and Photonics, 2016. → pages 31

[34] F. Jolesz. Advanced multi-modality image guided operating (amigo) suite. In *Proc. Fourth Image-guided Therapy (NCIGT) Workshop*, pages 15–6, 2011. → pages 7

[35] K. Kanatani, Y. Sugaya, and H. Niitsuma. Triangulation from two views revisited: Hartley-sturm vs. optimal correction. *In practice*, 4:5, 2008. → pages 99

[36] A. Kutikov and R. G. Uzzo. The renal nephrometry score: a comprehensive standardized system for quantitating renal tumor size, location and depth. *The Journal of urology*, 182(3):844–853, 2009. → pages 16, 39

[37] D. M. Kwartowitz, S. D. Herrell, and R. L. Galloway. Toward image-guided robotic surgery: determining intrinsic accuracy of the da vinci robot. *International Journal of Computer Assisted Radiology and Surgery*, 1(3): 157–165, 2006. → pages 6, 53, 58

[38] J. S. Lam, J. Bergman, A. Breda, and P. G. Schulam. Importance of surgical margins in the management of renal cell carcinoma. *Nature clinical practice Urology*, 5(6):308–317, 2008. → pages 18, 22, 26

[39] T. Langø, S. Vijayan, A. Rethy, C. Våpenstad, O. V. Solberg, R. Mårvik, G. Johnsen, and T. N. Hernes. Navigated laparoscopic ultrasound in abdominal soft tissue surgery: technological overview and perspectives. *International journal of computer assisted radiology and surgery*, 7(4): 585–599, 2012. → pages 25

[40] W. K. Lau, M. L. Blute, A. L. Weaver, V. E. Torres, and H. Zincke. Matched comparison of radical nephrectomy vs nephron-sparing surgery in patients with unilateral renal cell carcinoma and a normal contralateral kidney. In *Mayo Clinic Proceedings*, volume 75, pages 1236–1242. Elsevier, 2000. → pages 18, 21

[41] D. C. Leslie, A. Waterhouse, J. B. Berthet, T. M. Valentin, A. L. Watters, A. Jain, P. Kim, B. D. Hatton, A. Nedder, K. Donovan, et al. A bioinspired omniphobic surface coating on medical devices prevents thrombosis and biofouling. *Nature biotechnology*, 32(11):1134–1140, 2014. → pages 60

[42] J. Leven, D. Burschka, R. Kumar, G. Zhang, S. Blumenkranz, X. Dai, M. Awad, G. Hager, M. Marohn, M. Choti, et al. Davinci canvas: a

telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 811–818, 2005. → pages 30

[43] Q.-l. Li, H.-w. Guan, F.-p. Wang, T. Jiang, H.-c. Wu, and X.-s. Song. Significance of margin in nephron sparing surgery for renal cell carcinoma of 4 cm or less. *Chinese medical journal*, 121(17):1662–1665, 2008. → pages 22

[44] G. M. London. The clinical epidemiology of cardiovascular diseases in chronic kidney disease: Cardiovascular disease in chronic renal failure: Pathophysiologic aspects. In *Seminars in dialysis*, volume 16, pages 85–94. Wiley Online Library, 2003. → pages 21

[45] N. Mahmoud, I. Cirauqui, A. Hostettler, C. Doignon, L. Soler, J. Marescaux, and J. Montiel. Orbslam-based endoscope tracking and 3d reconstruction. *arXiv preprint arXiv:1608.08149*, 2016. → pages 40

[46] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Photonics for industrial applications*, pages 282–292. International Society for Optics and Photonics, 1995. → pages 26, 27

[47] O. Mohareri, G. Nir, J. Lobo, R. Savdie, P. Black, and S. Salcudean. A system for mr-ultrasound guidance during robot-assisted laparoscopic radical prostatectomy. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 497–504. Springer, 2015. → pages 28, 29

[48] T. Möller and B. Trumbore. Fast, minimum storage ray/triangle intersection. In *ACM SIGGRAPH 2005 Courses*, page 7. ACM, 2005. → pages 88

[49] L. J. Moore, M. R. Wilson, J. S. McGrath, E. Waine, R. S. Masters, and S. J. Vine. Surgeons display reduced mental effort and workload while performing robotically assisted surgical tasks, when compared to conventional laparoscopy. *Surgical endoscopy*, 29(9):2553–2560, 2015. → pages 6

[50] M. Najafi, N. Afsham, P. Abolmaesumi, and R. Rohling. A closed-form differential formulation for ultrasound spatial calibration: multi-wedge phantom. *Ultrasound in medicine & biology*, 40(9):2231–2243, 2014. → pages 99

[51] J.-J. Patard, O. Shvarts, J. S. Lam, A. J. Pantuck, H. L. Kim, V. Ficarra, L. Cindolo, K.-R. Han, A. De La Taille, J. Tostain, et al. Safety and efficacy of partial nephrectomy for all t1 tumors based on an international multicenter experience. *The Journal of urology*, 171(6):2181–2185, 2004. → pages 21

[52] N. Pavan, T. Silvestri, C. Cicero, A. Celia, and E. Belgrano. Intraoperative ultrasound in renal surgery. In *Atlas of Ultrasonography in Urology, Andrology, and Nephrology*, pages 137–146. Springer, 2017. → pages 25, 26, 101

[53] T. Peters and K. Cleary. *Image-guided interventions: technology and applications*. Springer Science & Business Media, 2008. → pages 7, 8, 9, 10

[54] T. M. Peters and C. A. Linte. Image-guided interventions and computer-integrated therapy: Quo vadis?, 2016. → pages 1, 7, 8, 32, 33, 34

[55] P. M. Pierorazio, H. D. Patel, T. Feng, J. Yohannan, E. S. Hyams, and M. E. Allaf. Robotic-assisted versus traditional laparoscopic partial nephrectomy: comparison of outcomes and evaluation of learning curve. *Urology*, 78(4): 813–819, 2011. → pages 6, 20

[56] P. Pratt, A. Di Marco, C. Payne, A. Darzi, and G.-Z. Yang. Intraoperative ultrasound guidance for transanal endoscopic microsurgery. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 463–470, 2012. → pages 31, 32

[57] 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*, 6(1):23–31, 2012. → pages 31

[58] P. Pratt, A. Jaeger, A. Hughes-Hallett, E. Mayer, J. Vale, A. Darzi, T. Peters, and G.-Z. Yang. Robust ultrasound probe tracking: initial clinical experiences during robot-assisted partial nephrectomy. *International journal of computer assisted radiology and surgery*, 10(12):1905–1913, 2015. → pages 32, 38, 45

[59] D. C. Rizzo. *Fundamentals of anatomy and physiology*. Cengage Learning, 2015. → pages 13, 14, 15, 16

[60] J. Sauro. Measuring usability with the system usability scale (sus), 2011. → pages 70

[61] C. Schneider, C. Nguan, M. Longpre, R. Rohling, and S. Salcudean. Motion of the kidney between preoperative and intraoperative positioning. *IEEE Transactions on Biomedical Engineering*, 60(6):1619–1627, 2013. → pages 25, 26, 32

[62] C. Schneider, C. Nguan, R. Rohling, and S. Salcudean. Tracked pick-up ultrasound for robot-assisted minimally invasive surgery. *IEEE Transactions on Biomedical Engineering*, 63(2):260–268, 2016. → pages xiv, 37, 38

[63] C. M. Schneider, G. W. Dachs II, C. J. Hasser, M. A. Choti, S. P. DiMaio, and R. H. Taylor. Robot-assisted laparoscopic ultrasound. In *International Conference on Information Processing in Computer-Assisted Interventions*, pages 67–80. Springer, 2010. → pages 30

[64] R. Shiroki, N. Fukami, K. Fukaya, M. Kusaka, T. Natsume, T. Ichihara, and H. Toyama. Robot-assisted partial nephrectomy: Superiority over laparoscopic partial nephrectomy. *International Journal of Urology*, 23(2): 122–131, 2016. → pages 6, 20

[65] R. L. Siegel, K. D. Miller, and A. Jemal. Cancer statistics, 2016. *CA: a cancer journal for clinicians*, 66(1):7–30, 2016. → pages 16

[66] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays: A literature review of augmented reality. *Journal of Display Technology*, 4(4): 451–467, 2008. → pages 27

[67] T. Simpfendörfer, C. Gasch, G. Hatiboglu, M. Mueller, L. Maier-Hein, M. Hohenfellner, and D. Teber. Intraoperative computed tomography imaging for navigated laparoscopic renal surgery: First clinical experience. *Journal of Endourology*, 30(10):1105–1111, 2016. → pages 28

[68] 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. → pages 28

[69] S. E. Sutherland, M. I. Resnick, G. T. Maclennan, and H. B. Goldman. Does the size of the surgical margin in partial nephrectomy for renal cell cancer really matter? *The Journal of urology*, 167(1):61–64, 2002. → pages 22

[70] M. Swindle, A. Makin, A. Herron, F. Clubb Jr, and K. Frazier. Swine as models in biomedical research and toxicology testing. *Veterinary pathology*, 49(2):344–356, 2012. → pages 98

[71] H.-J. Tan, E. C. Norton, Z. Ye, K. S. Hafez, J. L. Gore, and D. C. Miller. Long-term survival following partial vs radical nephrectomy among older patients with early-stage kidney cancer. *Jama*, 307(15):1629–1635, 2012. → pages 21

[72] 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. → pages 28, 47

[73] R. H. Thompson, S. A. Boorjian, C. M. Lohse, B. C. Leibovich, E. D. Kwon, J. C. Cheville, and M. L. Blute. Radical nephrectomy for pt1a renal masses may be associated with decreased overall survival compared with partial nephrectomy. *The Journal of urology*, 179(2):468–473, 2008. → pages 18, 21

[74] R. H. Thompson, B. R. Lane, C. M. Lohse, B. C. Leibovich, A. Fergany, I. Frank, I. S. Gill, M. L. Blute, and S. C. Campbell. Renal function after partial nephrectomy: effect of warm ischemia relative to quantity and quality of preserved kidney. *Urology*, 79(2):356–360, 2012. → pages 18, 22, 35

[75] O. Ukimura and I. S. Gill. Imaging-assisted endoscopic surgery: Cleveland clinic experience. *Journal of Endourology*, 22(4):803–810, 2008. → pages 28

[76] S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 13(4):376–380, 1991. → pages 63

[77] C. Våpenstad, A. Rethy, T. Langø, T. Selbekk, B. Ystgaard, T. A. N. Hernes, and R. Mårvik. Laparoscopic ultrasound: a survey of its current and future use, requirements, and integration with navigation technology. *Surgical endoscopy*, 24(12):2944–2953, 2010. → pages 26

[78] R. Venkatesh, K. Weld, C. D. Ames, S. R. Figenshau, C. P. Sundaram, G. L. Andriole, R. V. Clayman, and J. Landman. Laparoscopic partial nephrectomy for renal masses: effect of tumor location. *Urology*, 67(6): 1169–1174, 2006. → pages 17, 21, 90

[79] F. Volonté, N. C. Buchs, F. Pugin, J. Spaltenstein, B. Schiltz, M. Jung, M. Hagen, O. Ratib, and P. Morel. Augmented reality to the rescue of the minimally invasive surgeon. the usefulness of the interposition of

stereoscopic images in the da vinci robotic console. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 9(3):e34–e38, 2013. → pages 29

[80] R. Wang, Z. Geng, Z. Zhang, and R. Pei. Visualization techniques for augmented reality in endoscopic surgery. In *International Conference on Medical Imaging and Virtual Reality*, pages 129–138. Springer, 2016. → pages 29

[81] E. Wild, D. Teber, D. Schmid, T. Simpfendörfer, M. Müller, A.-C. Baranski, H. Kenngott, K. Kopka, and L. Maier-Hein. Robust augmented reality guidance with fluorescent markers in laparoscopic surgery. *International journal of computer assisted radiology and surgery*, 11(6):899–907, 2016. → pages 28

[82] M. R. Wilson, J. M. Poolton, N. Malhotra, K. Ngo, E. Bright, and R. S. Masters. Development and validation of a surgical workload measure: the surgery task load index (surg-tlx). *World journal of surgery*, 35(9):1961, 2011. → pages 101

[83] X. Wu and X. Shu. Epidemiology of renal cell carcinoma. In *Renal Cell Carcinoma*, pages 1–18. Springer, 2017. → pages 16

[84] Z. Yaniv and K. Cleary. Image-guided procedures: A review. *Computer Aided Interventions and Medical Robotics*, 3:1–63, 2006. → pages 7

[85] M. C. Yip, D. G. Lowe, S. E. Salcudean, R. N. Rohling, and C. Y. Nguan. Real-time methods for long-term tissue feature tracking in endoscopic scenes. In *International Conference on Information Processing in Computer-Assisted Interventions*, pages 33–43. Springer, 2012. → pages 100

[86] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. 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. → pages 47

[87] L. Zhang, M. Ye, P.-L. Chan, and G.-Z. Yang. Real-time surgical tool tracking and pose estimation using a hybrid cylindrical marker. *International Journal of Computer Assisted Radiology and Surgery*, pages 1–10, 2017. → pages 31

[88] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11): 1330–1334, 2000. → pages 42, 43

[89] P. T. Zhao, L. Richstone, and L. R. Kavoussi. Laparoscopic partial nephrectomy. *International Journal of Surgery*, 36:548–553, 2016. → pages 20, 21