A System for Fused Ultrasound-MR Image Guidance for Robot-Assisted Radical Prostatectomy

by

Keith Tsang

BASc, University of Waterloo, 2016

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)

December 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

A System for Fused Ultrasound-MR Image Guidance for Robot-Assisted Radical Prostatectomy

submitted by **Keith Tsang** in partial fulfillment of the requirements for the degree of **Master of Applied Science** in **Biomedical Engineering**.

Examining Committee:

Septimiu Salcudean, Electrical and Computer Engineering
Supervisor

Robert Rohling, Electrical and Computer Engineering
Supervisory Committee Member

Purang Abolmaesumi, Electrical and Computer Engineering
Supervisory Committee Member

Abstract

Robot-assisted laparoscopic radical prostatectomy is a surgical operation where the entire prostate gland is removed. This complex procedure requires the surgeon to establish a fine balance between completely removing the cancer while sparing critical anatomy responsible for continence and potency. Much of the difficulty of this surgery lies in the inability for the surgeon to actively localize the tumours during the surgery and distinguish between cancerous tissue to be removed and healthy tissue to be left behind.

This thesis details the design and development of an image guidance system for radical prostatectomy with the goal of improving patient outcomes by providing the surgeon with improved cancer localization. An image guidance system was proposed, consisting of pre-operatively segmented Magnetic Resonance (MR) images registered to intraoperative transrectal ultrasound rendered in a 3D virtual scene. This system builds upon prior work done by previous members of the lab who developed a TRUS robot which is able to perform a da Vinci-TRUS registration as well as a separate system for MRI-TRUS registration. The registered MRI and TRUS images are presented to the surgeon through the da Vinci surgical console's TilePro display system. Improvements to this prior system in this thesis

include specifically: rendering the MRI and TRUS in a visual representation of the scene, an imaging pipeline framework for capturing video from multiple sources, improvements to the calibration between da Vinci and TRUS coordinate systems, an augmented reality overlay for the ultrasound image displaying the da Vinci instrument location, prostate and tumour boundaries and a real-time registration algorithm for tracking prostate motion over the course of a surgery was integrated into the main workflow.

A series of experiments were conducted in order to validate the system. First, a phantom study was conducted to evaluate the accuracy of the entire registration process. Second, experiments were conducted to measure the runtime performance and latency of the application. Third, the accuracy of the real-time registration algorithm was tested. Lastly, the results of 15 of the ongoing surgical studies using the image guidance system are presented.

Lay Summary

Robot-assisted radical prostatectomy, one of the main treatment options for prostate cancer, is a difficult surgery. Surgeons conducting this surgery must balance between two conflicting goals: the complete removal of cancer, or preserving important nerves responsible for potency and continence. This thesis details work done on an augmented reality image guidance system with the goal of improving the surgeons ability to locate the cancer. MR scans obtained prior to the surgery are capable of indicating accurately the location of the cancer. During the surgery itself, the MR images are deformed to match a real-time ultrasound image. These images are all drawn in a virtual surgical scene which is shown to the surgeon. This system was tested at the operating room for a total of 15 times with positive feedback from the surgeon.

Preface

Material from Chapter 4, Chapter 5, Chapter 6 and Chapter 8 were published in the journal "Medical Image Analysis" titled: "A partial augmented reality system with live ultrasound and registered preoperative MRI for guiding robot-assisted radical prostatectomy". The work was co-authored by Golnoosh Samei (first author), Claudia Kesch, Julio Lobo, Soheil Hor, Omid Mohareri, Silvia Chang, S. Larry Goldenberg, Peter C. Black, and Septimiu Salcudean. Golnoosh Samei was responsible for developing the MRI-TRUS registration system. Julio Lobo and Omid Mohareri were largely responsible for the TRUS robot system. Dr. Peter Black was the main surgeon during the surgical studies. Claudia Kesch was the resident who assisted in the surgical studies. Soheil Hor developed the data collection system used to record surgical data. Silva Chang was responsible for performing the MRI segmentations. Larry Goldenberg is a clinical collaborator and Dr. Septimiu Salcdean was the principal investigator. My contributions to the paper include writing the chapters on system hardware, software and latency. I contributed to the paper by developing the main software systems used as well as assisting in experiments and the surgical studies.

The motor control application developed in Chapter 5 was based on libraries written by Julio Lobo, Andrew Thompson, Troy Adebar and Daniel Da Costa. The modifications to the calibration algorithm was co-developed with Golnoosh Samei. I developed the main application and integrated the DeckLink framework into the system. The phantom study was conducted with Golnoosh Samei as a co-investigator. Golnoosh Samei conducted the data analysis on the study.

The visualization application developed in Chapter 6 was based on an application developed by Orcun Goksel. I was the main developer of the application and the features described in the chapter. The application was co-developed with Golnoosh Samei and Neerav Patel.

The algorithm from Chapter 7 was based on Golnoosh Samei's paper published in IEEE Transactions on Medical Imaging: "Real-Time FEM-Based Registration of 3-D to 2.5-D Transrectal Ultrasound Images". I contributed by developing a C++ implementation of the algorithm and integrating it into the main visualization software.

The surgical study described in Chapter 8 was conducted under an ethics approval from the UBC Clinical Research Ethics Board. The project title is: "Intraoperative TRUS Guidance for RALRP" and it's certificate number is H11-02267.

Table of Contents

List of Tables

List of Figures

Acronyms

API	Application Programming Interface. 34, 47, 80, 83, 91–94
AR	Augmented Reality. 15, 16, 53, 94, 95
CSV	Comma-Separated Values. 91
CT	Computed Tomography. 5, 21
CUDA	Compute Unified Device Architecture. 116, 145
DCE	Dynamic Contrast Enhanced. 5
DSC	Sørensen-Dice Similarity Coefficient. xv, 107–110
DVI	Digital Visual Interface. 47, 56, 60
DVRK	DaVinci Research Kit. 56, 57
DWI	Diffusion Weighted Imaging. 5
EM	ElectroMagnetic. 97

FEM Finite Element Model. 10, 11

FRE Fiducial Registration Error. 69, 70

GPU Graphical Processor Unit. 116, 145

GUI Graphical User Interface. xiv, 38, 61, 63, 64, 83, 84, 92

HD High Definition. 33, 52, 58

HDMI High-Definition Multimedia Interface. 47

IP Internet Protocol. 92

IR InfraRed. 98

LED Light-Emitting Diode. 20

LRP Laparoscopic Radical Prostatectomy. 2, 3, 7, 16

MFC Microsoft Foundation Classes. 61, 78, 80

MIS Minimally Invasive Surgery. 2, 15, 19, 21, 96

mp-MRI MultiParametric Magnetic Resonance Imaging. 5, 6

MRI Magnetic Resonance Imaging. xvi, 5, 6, 8–12, 16, 17, 21, 26, 27, 29, 31, 37, 39, 40, 53, 54, 60, 68, 72, 73, 77, 80, 83, 90, 94, 96, 98, 99, 101, 118, 121, 122, 125

MSM	Master Side Manipulator. 3, 4
NCC	Normalized Cross Correlation. xv, 108, 109, 111, 114
NTSC	National Television System Committee. 57, 58
OR	Operating Room. 57, 67, 121, 140
ORP	Open Radical Prosetatectomy. 2
PSA	Prostate Specific Antigen. 2
PSM	Patient Side Manipulator. 3, 4, 7
RALRP	Robotic Assisted Laproscopic Radical Prosectomy. 2, 4, 6, 7, 9–11, 15–17, 22, 31, 40, 143
RANSAC	RANdom Sample Consensus. xiv, 66–69
RCL	Robotics and Controls Lab. xxi, 9
RGB	Red Green Blue. 24, 52
RP	Radical Prostatectomy. 2, 4, 5
SD	Standard Definition. 58
SDI	Serial Digital Interface. 46, 47, 56, 60, 122
SDK	Software Development Kit. 45, 47, 49, 51
SfM	Shape from Motion. 19, 20
SFP	Small Form-factor Pluggable. 46

SfS	Shape from Shading. 20
SLAM	Simultaneous Localization And Mapping. 19
SSD	Sum of Squared Differences. 100
SXGA	Super Extended Graphics Array. 52
TCP	Transmission Control Protocol. 92
ToF	Time-of-Flight. 20, 21, 98
TRE	Target Registration Error. 69, 74
TRUS	Transrectal Ultrasound. xvi, 6–11, 16, 17, 26, 28, 29, 31, 32, 37, 39, 60, 61, 64, 65, 68, 70–73, 77, 83, 98–100, 117, 118, 120–122, 125
UI	User Interface. xv, 63, 64, 84, 85
UML	Unified Modelling Language. xiv, xv, 47, 48, 62, 63, 81, 85–87, 101, 103
US	Ultrasound. 6–8, 31, 54, 60, 120, 121
VGA	Video Graphics Array. 31

Acknowledgments

I would first like to thank my supervisor, Dr. Septimiu Salcudean for all his supervision during this thesis. He has always provided me with advice and guidance whenever I needed it. I am very thankful for his patience and financial support as I've taken much longer to complete than usual.

I would also like to thank all the members of my team, Golnoosh Samei, Julio Lobo, Soheil Hor, Davood Karimi for all their work and contributions to this project. Without their work, the project wouldn't have gotten to where it is today.

I would like to thank Dr. Peter Black for his for his support of the image guidance project. Thank you for the patience to wait for our long development cycles and letting us re-iterate and revise our setup. Thanks to his continuous feedback we've been able to make many improvements to this project.

I would like to thank Claudia Kesch and Silvia Chang for their assistance throughout the study. Claudia for help in coordinating and recruiting patients as well as her help in registration during all the studies. Silvia for all the MRI segmentations she's had to complete for our project.

I would like to thank all the members of the Robotics and Controls Lab (RCL) lab, especially Megha Kalia, Prateek Mathur, Alaa Eldin Abdelaal and Apeksha Avinash for all the support and advice.

I would like to thank my good friend Varun Jacob-John for all his help and support throughout my program. Thank you so much for reading my thesis so many times and all the suggestions and edits you had to make.

Finally, I would like to thank my family back home in Toronto for continuing to support me over these few years. This has been the first time I've been so far away from home and without your support, I couldn't have gotten this far.

Chapter 1

Introduction

Prostate cancer is the most commonly diagnosed cancer among men in North America and the second leading cause of death. In 2018, prostate cancer accounted for 13.5% of all new cases and 6.7% of cancer mortality among men worldwide [3]. Prostate cancer is considered a malignant tumour as over time, it will spread and invade other parts of the body in a process known as metastasis. Most cases of prostate cancer are diagnosed at an early stage where metastases have yet to occur where the cancer may spread to regional pelvic lymph nodes [14].

At an early stage, the three most common disease management options for prostate cancer are active surveillance, radiation therapy and radical prostatectomy. Active surveillance is a prostate cancer treatment strategy that consists of deferring treatment, monitoring the patient's symptoms and repeating biopsies, until the cancer becomes more active. Radiation therapy is an active treatment option which aims to control or kill the malignant tumour through ionizing radiation. For prostate cancer, this is often done through a treatment known as brachytherapy where many small radioactive source are implanted into the prostate under ultrasound guidance.

Radical Prostatectomy (RP) is a surgical operation involving the complete removal of the prostate gland. It is a well established and accepted treatment option with well studied eligibility criteria. Through histopathological examination of the patient's biopsy tissue samples and Prostate Specific Antigen (PSA) levels, the risk of metastasis could be determined. RP could be completed either as an open surgery, where a large incision is made in the patient in order to access the prostate. Alternatively, RP could be completed as a laparoscopic Minimally Invasive Surgery (MIS) where only small incisions are made for the surgical instruments and an endoscope camera to be inserted into the body. The surgeon manipulates the camera and the instruments through the small incisions.

1.1 Robotic Surgery

RP has three common variations: Open Radical Prosetatectomy (ORP), Laparoscopic Radical Prostatectomy (LRP) and Robotic Assisted Laproscopic Radical Prosectomy (RALRP). Open surgery is a traditional form of surgery in which the surgery is performed through a large incision made in the patient. This form of surgery provides the surgeon with the most visual information as well as the most access for surgical tools. However, open surgery often results in long and painful recovery times as the incisions made are highly invasive with higher chances of infection. Laparoscopic surgery is a form of MIS, a procedure where small ports are placed in the body where cameras and surgical instruments can be inserted. These tools are by the surgeon using an endoscopic camera for operating field visualization. Compared to conventional open surgery, laparoscopic surgery provides advantages such as less blood loss, less pain and faster recovery due to smaller incisions. However, laparoscopic surgery is complex and difficult to perform. Instruments pivot around the surgical ports, so the movement of the distal instrument mirrors the movement of the surgeon's hand. Thus instrument manipulation is counter-intuitive. When combined with a lack of depth perception from a two dimensional endoscopic camera view, this results in a very steep learning curve. As a result, while LRP does provide significant benefits over conventional surgery, it is difficult to master as not all surgeons can be trained to perform them [31].

Two of the main factors limiting the performance of LRP are the loss of freedom of motion due to rigid instruments and the poor visualization from a lack of depth perception by using a monocular endoscope camera. The most common surgical robot in operation, Intuitive Surgical's da Vinci Surgical System, was developed to take advantage of the clinical benefits of laparoscopic surgery and overcome its limitations by providing additional technical advantages. A surgeon operates at a console which displays the surgical scene through a video feed from a stereo endoscopic camera which provides depth perception. The surgeon controls the robot by moving two Master Side Manipulators (MSMs) whose movements are mapped to the Patient Side Manipulators (PSMs). These PSMs are placed into the patient and can be outfitted with a wide selection of instruments to perform the needed surgical tasks according to the surgeon's movements. Unlike standard laparoscopic instruments, these instruments are wristed, which provides two more degrees of freedom over conventional laparoscopic tools. The surgeon's hand motion is mapped to the motion of the instrument tips in an intuitive manner. Compared with laparoscopic surgery, the da Vinci robot has the advantage of better ergonomics, wristed instruments, motion scaling (the surgeon's hand motion is scaled down in order to manipulate small targets) and tremor filtering (physiological tremor is well understood and can be filtered our to a significant extent by the PSMs which only copy a low-pass filtered version of the MSM's motion). These features make robot-assisted laparoscopic operations much easier. Therefore, by 2010 in the United States, RALRP has become the standard of care with over 85% of RP operations performed with robot assistance [6].

Radical Prostatectomy is a difficult surgery where a balance between two conflicting goals must be maintained. Oncological success refers to cancer control or to the complete removal of the cancer without any remaining inside the patient. In order to determine oncological success, surgeons use a surgical margin, where additional tissue at a defined distance surrounding the tumour is being to ensure the complete removal of all cancerous tissue. When the cancerous tissue is examined after a surgery, any cancer located on the edge of the tissue specimen removed from the patient is known as a positive margin, as this indicates that the cancer was not completely removed. A clear distance between the edge of the sample and the tumour is known as a clear margin. While a clear margin is often necessary for determining whether or not the entirety of the tumour is removed, a balance must be maintained as excessive loss of tissue often results in complications to the patient. In the case of prostate cancer, tumours are often located near critical anatomical structures such as the neurovascular bundle which is responsible for sexual potency. As well, if too much tissue is removed in the sphincter area, urinary incontinence follows. Functional success refers to the preservation of continence and potency which is directly correlated to the sparing of critical structures near the prostate. By increasing surgical margins and removing critical structures, the risk of leaving cancer behind decreases; however, this also increases the risk of incontinence and impotence. Much of the difficulty of RALRP lies in the inability of the surgeon to visually localize the cancerous lesions with the naked eye or with the conventional endoscope. This results in requiring larger surgical margins in order to ensure oncological success. Through the use of medical imaging, techniques which create visual representations of the interior of the body, could be used in order to localize tumours as well as the surrounding anatomy. Of all medical imaging types, Magnetic Resonance Imaging (MRI) and ultrasound are of particular interest due to the benefits they can provide for RP.

1.2 Magnetic Resonance Imaging

Magnetic Resonance Imaging is a medical imaging technique which uses magnetic fields in order to generate images of anatomy in the body. Unlike Computed Tomography (CT) imaging, MRI does not use X-rays and thus does not produce ionizing radiation. MRI imaging is capable of producing highly detailed images of the prostate along with the surrounding anatomy. MRI is not without its drawbacks: it is expensive and difficult to access. MRI machines operate with very strong magnetic fields which attracts metal objects in their vicinity. This, combined with their large size, prevents MRI machines from being used in surgical settings.

MultiParametric Magnetic Resonance Imaging (mp-MRI) is a supplemental form of MRI where additional sequences such as Dynamic Contrast Enhanced (DCE) or Diffusion Weighted Imaging (DWI) are used along with conventional T1 and T2 weighted imaging. In multiple studies, as reviewed by Fütterer et al. [9] and Hegde et al. [13], mp-MRI has been shown to improve prostate cancer detection. T2 weighted MRI in particular, is capable of producing highly detailed anatomical images of the prostate and its internal structure such as the seminal vesicles and urethral sphincter [54].

To leverage the accuracy of MRI imaging, MRI in-bore techniques by Susil et al. [46] and Cepek et al. [4] have been developed for brachytherapy and biopsy. While these techniques are capable of providing high quality image guidance, surgery within the bore of the MRI machine is still not feasible with current technology.

Instead, the use of mp-MRI for surgical planning has been reported by groups such as the work by Tan et al. [47]. These authors suggest that MRI could be used to determine if the disease has spread to the seminal vesicles and neurovascular bundles. This could determine if the surgical approach could spare critical structures which will allow for a faster return to continence. MRI imaging has been shown to be effective in changing surgical plans [29], but its use prior to RALRP is limited. The MRI images must be mentally superimposed onto the patient's anatomy and this is difficult to do, especially since there are significant differences in the position of the patient between acquisition and surgery. Surgeons can only approximate the location of any tumours through cognitive fusion and use MRI for surgical planning rather than image guidance.

1.3 Ultrasound Imaging

Medical ultrasound is an imaging technique used to create an image of the body's anatomy through the use of ultrasound waves. Ultrasound imaging uses sound waves at very high frequencies inaudible to humans and are emitted into tissue using a transducer. The tissue's properties affect the way the ultrasound waves are reflected which allows for an image of the tissue to be formed. Transrectal Ultrasound (TRUS) in particular, is the most common form of Ultrasound (US) used for imaging the prostate. While TRUS does have many benefits, such as its low cost, lack of ionizing radiation and real-time imaging capabilities, it does have clear limitations. The biggest drawback of US imaging is that at the high frequencies it operates in, the cancerous tumours cannot be differentiated from the surrounding tissue. As a result, TRUS the boundary between the prostate and the surrounding tissue can be easily seen, but the tumours are indistinguishable from the prostate tissue.

Ultrasound imaging, specifically TRUS is the most common form of ultrasound used for imaging the prostate. Real time TRUS imaging during LRP was first pioneered by [49]. It has since been used in RALRP which have shown improvements in oncological outcomes by a reduced overall PSM by Ukimura et al. [50]. While TRUS has great potential in RALRP, it has a few limitations which need to be considered. The most fundamental problem is the need of an assistant to manually adjust the probe during the surgery. This problem is twofold: the surgeon needs to verbally direct the assistant and the assistant needs access to the probe itself. For RALRP specifically, the da Vinci robot is placed directly adjacent to the operating table which restricts access to objects in that region. Hung et al. [15] developed a robotically manipulated TRUS device controlled using a foot pedal but reported difficulties with fine movement and limited directional control. Han et al. [12] developed another robotically manipulated TRUS system controlled using a joystick and has reported positive results. While these robotic systems serve as substitutes for human operators for the TRUS probe, the responsibility for controlling the probe now falls on the surgeon through the use of custom made control devices such as joysticks or foot pedals. The lack of integration of these devices into the da Vinci robot system places a larger mental overhead on the surgeon.

As a solution to these problems, Mohareri Omid et al. [33] proposed a robotic

TRUS system controlled only by the da Vinci Surgical System. A TRUS probe is mounted on a brachytherapy stepper (microtouch) modified with a motor that rotates the probe. Custom software on the Ultrasound console is responsible for acquiring the ultrasound image data as well as controlling the motor motion. This allows the system to automatically rotate and track the position of the da Vinci instrument tool tip with the TRUS probe. The entire system is integrated into the da Vinci robot workflow thus removing the need for additional operators or manual operation by the surgeon.

1.4 MRI-US Fusion

Ultrasound and MRI imaging both have certain advantages and disadvantages. MRI-US fusion is a technique which involves using both modalities at the same time to overcome the limitations of each individual imaging modality. MRI-TRUS fusion techniques have been developed to overcome the limitations of each individual imaging modality. These techniques can be divided into two types: MRI-directed cognitive fusion and MRI-TRUS fusion. Cognitive MRI-directed fusion is a technique in which an operator reviews MRI images prior to TRUS-guided operation and infers the location of tumours relying on their memory to mentally localize the cancerous regions. While this technique has been demonstrated to be superior than using TRUS alone, the onus lies on the operator's spatial cognition for accurately inferring the location of MRI lesions in the TRUS image. Software based MRI-TRUS fusion is an alternative technique in which medical image registration methods are used to perform the image alignment instead of relying on the operator's intuition. MRI-TRUS fusion removes the need for this cognitive task by performing the alignment through registration algorithms where the MRI can be rigidly or non-rigidly registered to the TRUS image.

While there are currently no reports of MRI-TRUS fusion for RALRP, several MRI-TRUS-guided biopsy systems do exist. In a review by Logan et al. [24], MRI-TRUS fusion has generally been shown to provide higher tumour-positive results than alternatives. While these systems prove the feasibility of MRI-TRUS fusion, they are not applicable for RALRP as they require manual re-positioning of the TRUS probe.

1.5 Image Guidance for Radical Prostatectomy

While RALRP has many advantages over conventional laparoscopic procedures, achieving both oncological and functional success is largely dependent on the surgeon's skill. Identification of specific anatomical features such as the neurovascular bundle, prostate apex and bladder neck are especially important as retaining these structures increase the odds of functional success. However, it is often quite difficult to accurately identify these critical structures using only visual cues. These features are mainly identified using preoperative medical imaging techniques such as MRI. By integrating intraoperative medical imaging into the system, localization of critical structures can be improved.

An image guidance system was proposed by the Robotics and Controls Lab (RCL) to improve upon the main limitations of RALRP. The overall objective of this project is the improvement of patient outcomes by providing better cancer localization in the form of augmented reality to the surgeon. To fulfill this objective, an augmented reality guidance system is proposed where a virtual prostate with segmented tumour contours is overlaid on top of the live endoscope camera feed.

Augmented reality overlay is beyond the scope of this thesis, as it is a complex task and currently an active area of research. Instead, this thesis will focus on developing much of the underlying framework towards achieving an augmented reality system. The proposed system consists of a visualization application which positions a virtual prostate model at the same location as the true prostate. To do so, the virtual prostate will be rendered relative to the da Vinci coordinate system to match the true prostate anatomy as closely as possible. The prostate model is constructed from a segmentation of the patient's prostate obtained from preoperative MRI. As mentioned above, the main drawback of MRI is the inability to obtain images in real time. To overcome this limitation, MRI-TRUS fusion can be implemented by registering the MRI to real-time ultrasound. This allows a real-time rendering of the MRI volume along with the tumour segmentation, both of which could be provided to the surgeon.

At the time I joined the project, it already had a few core features completed by prior members of the lab. As mentioned before, Mohareri Omid et al. [33] has developed a system for registered TRUS guidance for RALRP. This system consists of a robotic brachytherapy stabilizer which can rotate a mounted TRUS transducer. This rotation was automated allowing the probe imaging plane to track the da Vinci instrument tip through a calibration procedure completed intraoperatively. The live ultrasound image is provided to the surgeon through their surgeon console.

Work by another previous member of the lab, Goksel et al. [10], while not a direct contribution for this project, can still be adapted and used for the visualization application. Goksel et al. [10]'s work consists of a virtual brachytherapy simulator which contains a framework for simulating ultrasound under tissue deformation caused by needle insertion. While needle insertion simulation is not directly applicable to the current project, the framework for a deformable prostate Finite Element

Model (FEM) can be adapted to image-guided RALRP. Along with this framework is a system for the rapid sampling of an image plane based on the deformation of the FEM. This can be easily modified to sample MRI volumes instead of an ultrasound volume. Golnoosh Samei, a postdoctoral researcher has adapted Goksel et al. [10]'s work. She adapted the prostate visualization from Goksel's simulator as well as the FEM of the prostate and rapid volume sampling. Julio Lobo, a research engineer in the lab has integrated Mohareri's work into the project workflow. This includes the robotic TRUS probe system along with the ultrasound to da Vinci calibration process. Soheil Hor, a masters student developed a data collection system used for recording surgical data.

During the course of the project, Julio Lobo's main contributions were towards the integration of a new ultrasound machine as well as assisting during the patient studies conducted. Golnoosh Samei's contributions were focused primarily towards the volume slicing library and MRI to TRUS registration. She also worked with another post doctoral researcher Davood Karimi, in developing an automatic registration system to replace the manual MRI to TRUS registration. For the contributions detailed in this thesis, Dr. Samei collaborated in the improvement of the calibration process. She was also the author of the real-time registration algorithm later implemented in this thesis.

1.6 Thesis Objectives

This thesis describes the development of additional features towards the overall goals of the project. This consists of developing the software systems required for the rendering and display of a virtual prostate, improving the robustness of the registration through improved calibration and prostate tracking, and initial steps towards augmented reality such as an ultrasound overlay system. These objectives aim to complete the missing features and improve upon the limitations of the current project setup and are required to fulfill the overall objectives of the augmented reality radical prostatectomy project. The main goals of the thesis are as follows:

1. Design and develop a system handling the input of real-time ultrasound images and output of the virtual prostate scene video stream.
2. Integrate a new, and much better quality, ultrasound machine into the motor control system.
3. Improve the reliability and robustness of the calibration process and validate the system through a phantom study.
4. Develop an application for the rendering of a virtual prostate model with registered MRI along with live ultrasound video augmented with an overlay which displays the prostate and tumour contours.
5. Develop a real time prostate tracking system to ensure a consistent registration for accurate tumour localization.

1.7 Thesis Outline

Chapter 2 provides an overview of current augmented reality surgical guidance methods. The chapter then covers the current state of augmented reality research. This also includes work done on registration and tracking techniques and their importance for augmented reality.

Chapter 3 serves as an overview and background on the current state of project prior to the contributions of this thesis. This chapter provides an overview of the hardware components required for the guidance system to function.

An overview of the current software systems is also provided along with additional tools which were used during the development of the project.

Chapter 4 presents the development of a input/output framework used for retrieving images from a new ultrasound machine as well as sending images to the da Vinci surgical console. This chapter also presents work done to provide depth perception to the 3D prostate simulation through the stereopsis effect.

Chapter 5 covers the development of an application used for the registration between ultrasound and da Vinci coordinates systems. A new improved calibration procedure is introduced in order to improve the robustness and reproducibility of the registration. This chapter also includes an phantom study performed to determine the overall accuracy of the registration process.

Chapter 6 details the work on developing the main visualization software of the image guidance system. This includes the integration of the framework developed in Chapter 4. This chapter also presents work done on developing a augmented reality overlay for the ultrasound video feed.

Chapter 7 reports on the work done on developing a real-time registration system. This chapter first covers an overview of registration and tracking algorithms and their importance to the project. This chapter then details how the algorithm was implemented along with an experiment for determining the performance to benchmark the algorithm.

Chapter 8 discusses observations which were made during surgical studies of the image guidance system. This chapter presents qualitative results of the system as well as feedback from the surgeon who used the system.

Chapter 9 is the final chapter which concludes the thesis as well as describing the contributions and possible future work.

Chapter 2

Overview of Augmented Reality Guidance Methods

2.1 Augmented Reality

Minimally invasive surgery has become increasingly popular due to its multiple benefits to the patient. However, minimally invasive surgery also increases the difficulty of the procedure due to the restrictive field of view. Medical imaging presents one possible way to alleviate these difficulties, but raises a problem in how to best create and present useful visualizations. Augmented reality aims to solve these challenges by presenting additional information to the operator through virtual overlays on top of the patient anatomy.

Augmented Reality (AR) for robotic assisted surgery has the potential for many benefits over conventional RALRP. Not only is there a standardized display and interface through the da Vinci surgical console, but also stereoscopic vision for depth perception. Augmented reality for laparoscopic MIS can allow the surgeon to quickly identify subsurface targets and critical structures. AR can also reduce the mental overhead of the surgeon by matching medical images to their respective sources in the surgical scene. Lastly, AR can potentially reduce the rate of positive surgical margins by displaying the margins on a virtual model. Reviews by Bernhardt et al. [2] and Lin et al. [21] outlines the potential as well as challenges in implementing augmented reality for laparoscopic surgery.

Simpfendörfer et al. [43] introduced an AR navigation system for TRUS-guided conventional LRP. Fiducials in the form of needles with coloured heads are inserted into the prostate. These fiducials are used for assistance for both registration and tracking. While this system is quite promising with a successful in-vivo study, the need for artificial invasive fiducials is a significant limitation. Not only will this increase surgical time as the needles must be planted, the use of fiducials may also increase the chances of bleeding or other complications.

Kolagunda et al. [20] presents work done on a mixed reality guidance system for RALRP. A 3D reconstruction of the prostate is obtained using the stereoscopic images of the endoscope camera. This surface model is then registered to preoperative MRI. Unfortunately, this system does not function in real-time and requires a camera calibration to be performed prior to the surgery which will prevent the surgeon from being able to change the focus of the endoscope camera. These limitation greatly limits the practicality of this solution.

Porpiglia et al. [40] has one of currently most sophisticated AR system for RALRP. Their system produces a 3D prostate model based on mp-MRI which is then overlaid on top of the laparoscopic video. The alignment process is manually performed using their custom software. The resulting scene is presented to the surgeon using the da Vinci TilePro system. Throughout the surgery itself, tracking is done by having an expert manually aligning the model. While this system has been successfully tested on 30 patients with positive results, the need for manual alignment complicates the usage of the system.

Ukimura et al. [51] introduced a similar system which requires a manual alignment for alignment. Their system uses MRI-TRUS fused 3D model for RALRP. The placement of the model requires significant expertize.

One of the most significant challenges for implementing a practical augmented reality effect is the need for accuracy. Reaching sufficient accuracy sufficient for clinical use is a significant challenge in itself, but maintaining the accuracy in a real time in a highly dynamic surgical scene becomes a major challenge. During an RALRP intervention, not only will the surgeon be moving the camera and surgical instruments, but there will also be blood and smoke from resection and cautery tool usage along with tissue deformation and resection. Another factor is the passive patient motion from breathing and heart beats. This motion is close to periodic and can be estimated and compensated as reported by Mourgues et al. [35] and Mountney and Yang [34].

In order to keep a registration accurate in real time, most methods are computationally expensive. To fulfill the requirements for real-time performance, two main criteria must be satisfied. The first criteria is the need for an acceptable latency or time delay. Reported in Marescaux et al. [28], 330 ms is the limit of acceptable latency. The second criteria is the need for a reasonable video frame rate ideally greater than 25 Hz.

2.1.1 Registration

The first major challenge for laparoscopic augmented reality is obtaining the initial registration between imaging modalities and the surgical scene. This registration challenge is a well established problem with a decade of research. Most proposed methods for laparoscopic AR registration can be categorized within four main approaches: interactive, point-based, surface-based and volume-based.

Interactive techniques largely rely on manual user inputs from an expert to register pre-operative medical images to an endoscopic scene. The effectiveness of these techniques largely depending on the amount of interaction involved, quality of interfaces as well as the overall expertise and familiarity of the expert when using the system.

Point-based techniques rely on an automatic identification of natural or artificial landmarks between pre-operative and endoscopic imaging. Work by Conrad et al. [5] uses anatomical landmarks in the liver for the automatic alignment between 3D reconstruction and laparoscopic imaging. The limitation of anatomical landmarks lies in the difficulty in identifying such landmarks, especially when the surgery involves mainly soft tissue. To alleviate this problem, artificial landmarks through fiducials or markers could be introduced prior to the acquisition of pre-operative data such as work by Megali et al. [30]. These markers often require a tracking system and also face the problem of anatomical deformation between preoperative acquisition and intervention.

Surface-based approaches aim to estimate the intraoperative scene through surface reconstruction either actively by the use of additional hardware to the endoscope or passively by analyzing visual cues obtained from endoscopic images.

Passive Techniques

- Shape from Motion (SfM) is a passive technique that infers depth of a surface using motion across frames. While SfM can be used for any endoscope, the main drawback of the techniques lies in the need for motion. SfM is also works best with non-deformable scenes but there has been recent work on non-rigid SfM such as work by Parashar et al. [39].

- Stereovision is a well established technique in which 3 dimensional geometric data can be obtained from images obtained simultaneously from two cameras at different viewing angles. In the stereo reconstruction step, the relative depth in features between “left” and “right” images is estimated and used to generate a stereo disparity map. While stereovision is fast and accurate, it depends heavily on scenes with rich textures in order to perform the feature matching. Real-time stereo reconstruction for robotically assisted MIS has demonstrated in work such as reported by Stoyanov et al. [45].

- Simultaneous Localization And Mapping (SLAM) is a technique developed by the robotics community and is mainly used for autonomous robotic navigation and localization. SLAM algorithms are very computationally intensive and generally rely on elements in a scene to remain static. A subset of these algorithms known as MIS-SLAM improves on these limitations by allowing for some scene deformation and has been demonstrated to be feasible in a few studies such as the work by Mountney and Yang [34]. MIS-SLAM is very new and suffers largely from its complexity and lack of established frameworks and tools.

- Shape from Shading (SfS) is a passive surface reconstruction technique which uses the shading and lighting conditions of the scene to predict depth. Some research, e.g., by Malti and Bartoli [27] and Visentini-Scarzanella et al. [55], has shown promising results for surface reconstruction but only in a controlled environment. SfS suffers from poor robustness as it must make significant assumptions about the lighting conditions of the environment. SfS is limited in use when used on its own but has potential when combined with other passive techniques such as hybrid SfS/SfM or SfS/stereovision.

Passive techniques have the main advantage over active surface reconstruction techniques where they function with any endoscope. However, most current passive techniques are not designed for laparoscopic conditions. While passive techniques work well in controlled environments with good lighting and rigid motions, they lack robustness and fail in dynamic environments with inconsistent lighting, textureless features and deformable scenes.

Active Techniques

- Structured light is an active technique which involves the usage of infrared light in a known pattern as a way of determining the depth of objects and has been used in the works such as Lin et al. [22].

- Time-of-Flight (ToF) is a technique in which the discrepancy of phase shift between emitted and reflected optical signals is measured. These optical signals vary from infrared light, to pulsed Light-Emitting Diodes (LEDs) or lasers. Reports of their use include work by Gudmundsson et al. [11], Kim et al. [18] and Fuchs and May [8].

The main drawback of active techniques lies in the introduction of additional hardware components. These projection devices often require a trade-off between resolution and sensor size as these sensors need to fit in conventional MIS environments. The introduction of new hardware face challenges in clinical testing as new medical devices will often require a lengthy regulatory approval process. Active techniques are able to deal with scenes with low texture complexity unlike passive techniques. However, in a study by Maier-Hein et al. [25], stereovision was demonstrated to have greater accuracy than both structured light and ToF techniques. Surface based techniques largely still suffer from a lack of robustness where dynamic changes in the surgical scene such as blood or smoke remains a major issue.

Volume-based techniques involves the use of intraoperative 3D imaging systems aside from endoscopes. This can vary from ultrasound transducers or open MRI scanners. The main goal of this approach is to use a secondary form of intraoperative imaging in which pre-operative data can be registered to. The main weakness of this technique lies in the need for additional imaging systems. These secondary imaging systems are also highly dependent on the type of OR environment available. Another drawback is the large size of volumetric data which often requires computationally expensive algorithms to process.

2.1.2 Tracking Techniques

The second major challenge of laparoscopic augmented reality is the need to maintain the registration over the course of an intervention where endoscope movement and scene deformation may occur. While continuously performing the registration in real-time is possible, such as the work of Vagvolgyi et al. [53] where CT data is registered to video, this is largely limited to rigid and simple scenes. As the current project is concerned primarily with RALRP, the camera movement tracking problem can be simply solved by obtaining the endoscope geometric information from the robot directly. The scene deformation problem however, remains a major challenge. The cause of laparoscopic scene deformation can largely be attributed to patient motion from heartbeat and breathing or surgeon interaction with tissue or tissue resection. While the former is periodic and can be compensated, the latter tends to be unpredictable and may cause permanent changes to the scene. The most common approach for maintaining scene tracking is from either natural or artificial landmarks.

Natural landmarks can found within a laparoscopic scene as features from edges, corners or textures. By keeping track of these features over time, the dynamics of the scene could be captured. These features could be obtained either automatically or manually specified by a user. Yip et al. [56] introduced a framework for acquiring dense 3D features from stereo endoscopic data. The main limitation of natural landmarks are similar to the problems faced by surface-based registration where certain laparoscopic scenes may lack distinct features as well as irregular illumination, specularity, as well as scene changes introduced during intervention such as blood and smoke.

Artificial landmarks can be introduced into a laparoscopic scene in the form of fiducials or markers to help facilitate their detection. These algorithms tend to be quite robust and tracking can be maintained for a long duration. In the work of Teber et al. [48], needles with coloured heads were introduced into tissues in the laparoscopic scene. Nakamoto et al. [36] also proposed the use of wireless tracking in the form of tiny EM transponders placed within the tissue. Singla et al. [44] introduced the user of 3D printed markers which can be placed on top of tissues of interest. The main drawback of artificial landmarks lie in their invasiveness. The insertion and removal of fiducials often take additional surgical time and could potentially increase the risk of complications. This invasive nature also reduces the number of markers that could be potentially introduced thus limiting the overall effectiveness of the technique.

Specularities

One of the main challenges in passive surface reconstruction and tracking techniques is the presence of specular reflections on the tissue surface. Specularity is a characteristic which represents the amount of reflectivity of a surface. In the surgical scene, specular highlights appear on tissue surface due to their non-lambertian nature. These highlights often occlude surface details such as texture and may appear as features. The problem of specularities can be largely be separated into two tasks: the detection of specular reflection and the recovery of the surface information. While recent research has shown positive results, the problem still remains unsolved. In work by Liu-Yin et al. [23], an algorithm is introduced for non-rigid 3D reconstruction with specular reflections. This work uses a prior obtained reflectance model created from a template to predict the position of specular reflections and has shown positive results. Unfortunately, it still lacks robustness as the reflectance model cannot account a dynamic surgical scene such as when blood or smoke is introduced. Mallick et al. [26] introduced a surface reconstruction technique with specular reflection removal using a series of images captured at multiple illumination levels.

2.1.3 Real-Time Registration Techniques

Real time surface reconstruction techniques, while possible, have largely been limited to static and non-rigid scenes. KinectFusion by Newcombe et al. [37] has introduced a real-time dense surface mapping and tracking system for static scenes which uses a Microsoft Kinect for depth mapping. Follow up work by Newcombe et al. [38] in DynamicFusion extended the algorithm to real-time non-rigid environments. Further work by Innmann et al. [16] further improves robustness using sparse Red Green Blue (RGB) colorspace feature matching. Unfortunately, these system's need for a depth camera makes it unsuitable for our current application.

2.2 Summary

To date, an accurate and dense real-time reconstruction of the laparoscopic scene has yet to be developed. Different approaches for registration and tracking have their specific advantages and drawbacks. For our specific application, the need for a robust solution in a highly dynamic environment greatly reduces the options available.

We propose the use of a real-time 3D to 2.5D ultrasound registration process to alleviate the main drawbacks of the volume-based approach. The work of a member of the project, Samei et al. [41] details an algorithm for using ultrasound to perform a volume-based registration. Not only is ultrasound readily available and inexpensive, it is also a component of our current system setup. 2D ultrasound is used, as 3D volumetric data is not available at real-time rates for prostate imaging.

Chapter 3

Augmented Reality System for Radical Prostatectomy System Setup

3.1 Introduction

The main objective of this project is the design and development of the software and hardware systems for an augmented reality image guidance system. Creating a realistic overlay which does not "float" over the target video is a difficult task and currently still an active area of research. The augmented reality overlay is beyond the scope of this thesis and will instead focus on establishing the frameworks and systems required for such a system.

3.1.1 System Requirements

A summary of the core features required for the main visualization software system as follows:

1. Display prostate model created from pre-operative MRI segmentation
2. Render prostate in position and orientation matching that of true prostate
   (a) Register MRI volume to ultrasound
   (b) Register ultrasound to da Vinci coordinate system
3. Render scene from position and orientation of da Vinci endoscope cameras
4. Control medical imaging using da Vinci instruments
   (a) Display MRI plane projected on prostate
   (b) Display ultrasound plane based on current transducer rotation
5. Present visualization to surgeon by sending 3D stereo scene to da Vinci surgeon console

The initial step of the project consists of an application for displaying a virtual prostate at the same location as the true prostate seen in the endoscope video. The prostate model is to be constructed from a preoperative MRI and updated to match the true prostate through MRI-TRUS fusion. The fusion will allow the MRI to deform to match the current state of the prostate as closely as possible. To render the virtual prostate in the same position as the true prostate, a transformation from the MR coordinate frame to the da Vinci coordinate frame must be determined. This could be accomplished by obtaining an ultrasound to da Vinci transformation through a calibration process. The scene can then be drawn in the same orientation and positive as the da Vinci endoscope camera video. The resulting visualization will display the prostate as close to the true prostate as possible. With the prostate at the correct location, a system is required to allow the surgeon to view the current segmented tumour locations in the MRI volume. This is completed by using the da Vinci instrument tool as a virtual cursor which could be used to scroll through the different planes of the MR volume. The ultrasound plane will also have to rotate accordingly to match the position of the da Vinci instruments. With the visualization and control systems established, a separate framework for sending the rendered scene to the da Vinci surgeon console is required. Another system will also be required for drawing the scene in two separate frames for stereopsis.

To achieve the features specified for the main visualization system, a few subsystems are required. Some of these subsystems have already been developed or can be adapted from existing work from previous students. These subsystems comprise both hardware and software interfaces to the hardware. A summary of the required tasks for these systems are as follows:

1. Obtain a data stream from the da Vinci surgical robot
2. Obtain ultrasound images from the new bk3500 ultrasound machine
3. Motorize the rotation of the ultrasound transducer probe
4. Send stereo video signals to the da Vinci surgical console
5. Record and save surgical data

The first subsystem consists an interface between the da Vinci surgical robot and the desktop computer. This system interfaces with the robot by establishing a connection and reporting the position and orientation of the surgical robot instruments and endoscope camera. This system has already been developed prior in the work of Mohareri Omid et al. [33] and Adebar et al. [1] and only needs small modifications to be used for the current project.

The second tasks requires the need for a method of retrieving ultrasound images from the ultrasound machine. The bk3500 ultrasound machine required a new interface for the ultrasound images to be transmitted into the software system. The only constraint for this requirement is that the images must be obtained in real-time with as little latency as possible.

The third system consists of a method of controlling the rotation of an ultrasound probe using the desktop computer system. This will require a combination of hardware systems along with the software for controlling the rotation. Once again, this system was already developed in the work of Adebar et al. [1].

The fourth requirement is a system for sending a generated video signal to the da Vinci surgical console. This consists of two videos streams for a stereo rendering of the virtual prostate scene to be viewed by the surgeon.

The last requirement is a system for recording data to be analyzed later and used for further development. This consists of recording the endoscope cameras of the da Vinci robot along with the ultrasound signal all synchronized to the data stream of the da Vinci instrument positions. This system was developed during this thesis by another member of the lab Soheil Hor.

3.1.2 Prior Work

This project was started by other members of the lab prior to the work completed in this thesis. As such, a few of the key subsystems have already been developed. First, the robotic TRUS probe system has already been developed by Adebar et al. [1] and tested in studies by Mohareri Omid et al. [33]. A basic visualization application has also been developed based on the brachytherapy simulator from Goksel et al. [10]. The two applications have been integrated together through the efforts of Julio Lobo, Golnoosh Samei and Soheil Hor. Their work includes an interface to the da Vinci robot which establishes a connection and retrieves data from the robot instrument locations as well as the data recording system used for collecting surgical data for later review. Dr. Samei further developed an application for a manual registration between MRI and TRUS over the course of the work completed in this thesis.

Aside from further work in the main visualization software, the subsystems not yet completed at the start of this thesis work consisted of a system for interfacing with the ultrasound machine and a system for sending stereo video signals to the da Vinci surgeon console. As mentioned above, a new ultrasound machine has been obtained by the lab which offers much better image quality than the previous one used by Mohareri Omid et al. [33]. This however, required a new interface in both the visualization software in development, as well as the robotic TRUS software developed by Mohareri Omid et al. [33].

3.2 Hardware Overview

The current guidance system consists of five main components. These are the da Vinci surgical robot, an ultrasound machine, a robotic TRUS probe system, a data recorder system, and an external desktop computer. Shown in Figure 3.1 is a diagram of how these components were connected together.

3.2.1 da Vinci Robotic Surgical System

The da Vinci robotic surgical system serves as both a data source as well as the final destination for the guidance system visualization. The da Vinci robot is developed

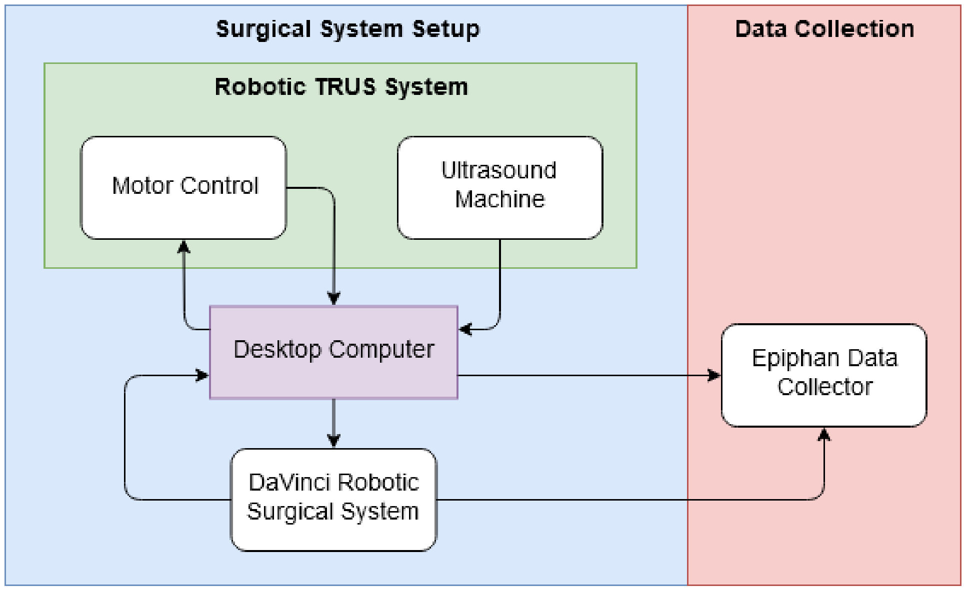


Figure 3.1: Overview of system hardware setup

by Intuitive Surgical and currently has models across four generations. While each generation has their own specifications, the core principle of the machine remains the same. The da Vinci system consists of three main components: the patient cart, the vision cart, and the surgeon console. The patient cart consists of the surgical robot itself which contains four independent arms with three instruments and one endoscope camera. The vision cart contains the camera controller systems which are used by the endoscope camera. The surgeon console serves as the main surgeon interface which contains the stereo screen and manipulators used by the surgeon to control the robot.

For the current project specifically, a 2^{nd} generation da Vinci S model was used for development while the studies were performed in the collaborating hospital where a 3^{rd} generation da Vinci Si model was used. While the two systems do have some differences, these differences are minor and do not affect the overall workflow of the system. Both machines use the same tools and provide the same data through their research interfaces. The main differences between the two machine mainly lies in the surgical console.

The TilePro system is a feature of the da Vinci surgeon console which provides a multiple input display system. The surgeon console screens normally only show the stereo camera video from the endoscope. When the TilePro system is enabled, it allows for the simultaneous display of the surgical scene along with video from up to two additional sources [52]. These additional video inputs are displayed below the standard camera view which can be seen in Figure 3.2. TilePro is often used for providing medical imaging such as MRI or ultrasound. The TilePro video inputs can either be displayed side by side or can be fused together to form a single stereo video. This TilePro system serves as the main platform for the final output of the proposed system. The virtual surgical scene produced by the guidance system is displayed in the console TilePro window, which serves as the main visual interface for the surgeon to view the MRI-US fusion.

In the current system setup, a Video Graphics Array (VGA) to composite adapter is used to connect the current computer desktop's screen with the TilePro composite ports. This allows the desktop to clone its display on the TilePro screen. While this method if simple to setup and use, unfortunately it only provides a mono-display to the surgeon.

3.2.2 TRUS Robot

As part of Adebar et al. [1]'s work, a robotic TRUS probe system was created to help facilitate RALRP. The system was designed with the purpose of eliminating the need of an external assistant for controlling the ultrasound. Instead, a robo-

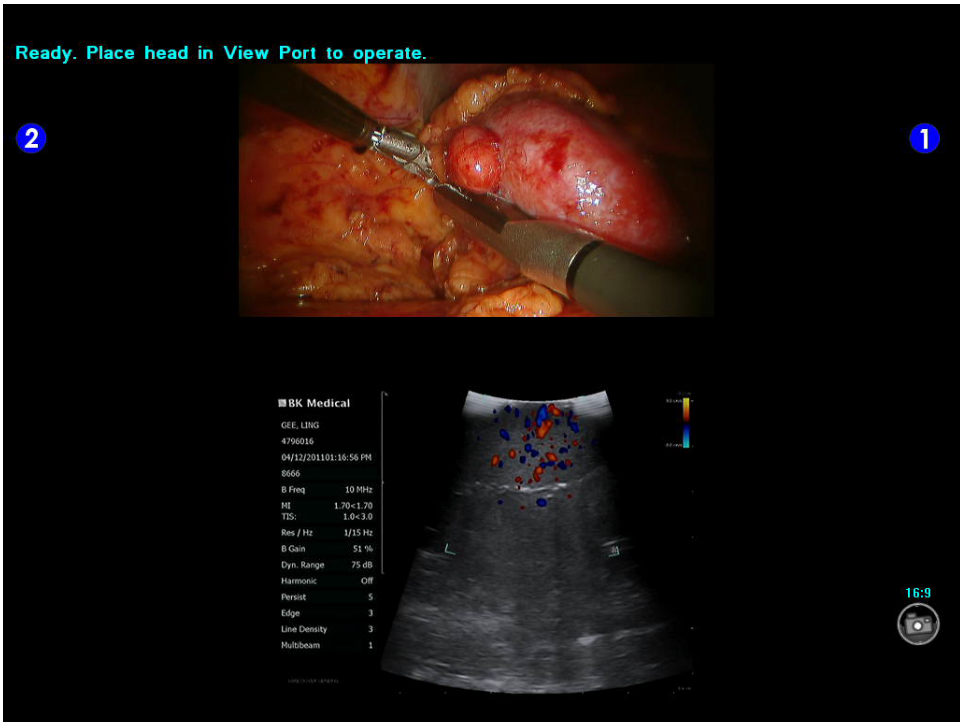


Figure 3.2: Ultrasound for partial nephrectomy using TilePro system

tized probe automatically rotates to follow the surgeon's instruments thus always keeping imaging plane focused on the area the surgeon is working on.

The system was tested in 20 patient study in the work of Mohareri Omid et al. [33]. A standard brachytherapy stabilizer system (Micro-Touch 610-911) was modified to remotely rotate a mounted ultrasound probe. A Sonix ultrasound machine with a sagittal/transverse biplane TRUS transducer was used to image the prostate. This stabilizer setup was mounted to the operating table, underneath the patient side manipulators of the da Vinci robot.

3.2.3 Multimedia Recorder

The Epiphan Pearl is a multimedia recorder which is used as the main medium for data collection. This data collection is independent from the other hardware components as it only receives data from the da Vinci endoscope cameras as video from the desktop computer. The Epiphan Pearl is capable of simultaneously capturing up to four High Definition (HD) video sources at once. This allows for the full recording of the surgery through saving the left and right channels of the da Vinci endoscope video, the ultrasound video and the virtual surgical scene produced by the guidance system. The da Vinci robot instrument positions and orientations is also recorded by encoding the data into an audio stream. The benefit of using this recording device is its ability to record all the channels simultaneously which allows the endoscope video to be synchronized properly with the robot instrument data. This allows for the full reconstruction and simulation of a surgical case which can be used to validate and test the system.

3.3 Software Setup

Prior to the start of this thesis, a collection of software applications and interfaces have already been developed which could be re-purposed for this project. A few of these applications have already been integrated into a basic system which performs the intended purpose of the guidance system. This system however, is still rather basic and lacks many features listed as part of the requirements. An overview of the original software layout can be seen in Figure 3.3.

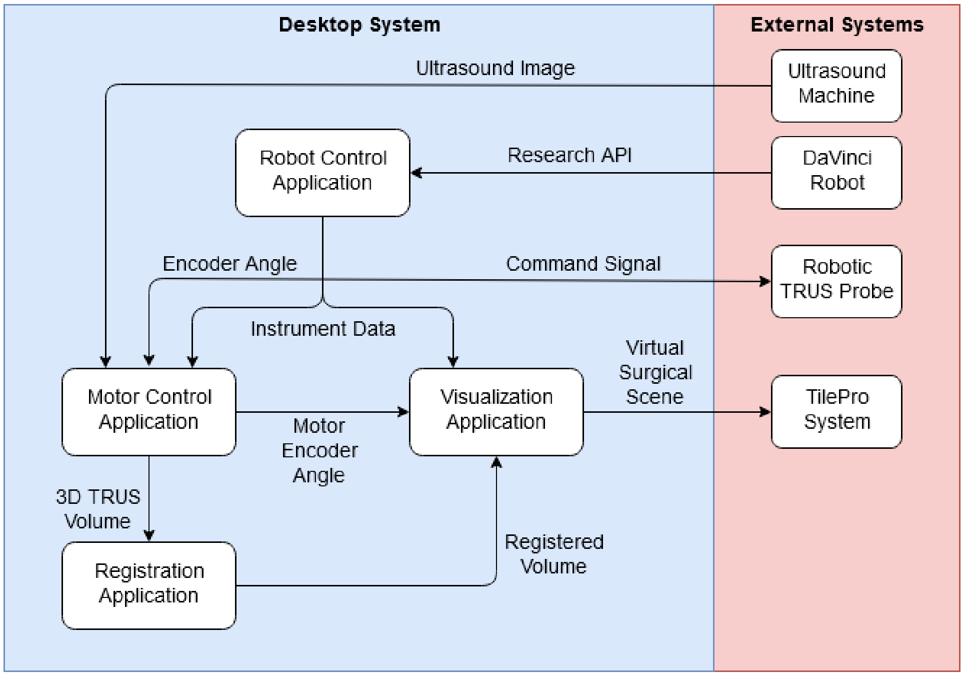


Figure 3.3: Overview of system software setup

3.3.1 Robot Control Application

The Robot Control Application serves as the central hub of the software system. It's main purpose is to establish communication with the da Vinci robot using the research Application Programming Interface (API) library provided by Intuitive Surgical. This library provides access to the robot through a read only interface and provides a variety of data once the connection is established. For this project specifically, the data of interest is the joint data from the da Vinci patient side manipulators. From this data, the position and orientation of the current da Vinci instrument tips can be derived through forward kinematics. The position and orientation of the robot endoscope camera can also be obtained through this way.

One of the main limitations of the research interface is that there can only be one connection established to the robot at a time. In Mohareri Omid et al. [33]'s work, this limitation was not a problem as only the motor controller needed to connect to the robot. However, for the current project objectives, the number of applications needing the instrument data has increased. As a workaround around this limitation, a socket-based server/client inter-process communication library was developed. The robot control application remains as the only application that connects to the da Vinci robot. Once the data stream is established, it serves as a server and broadcasts the data to any subscribing client processes. A diagram of these connections can be seen in Figure 3.4.

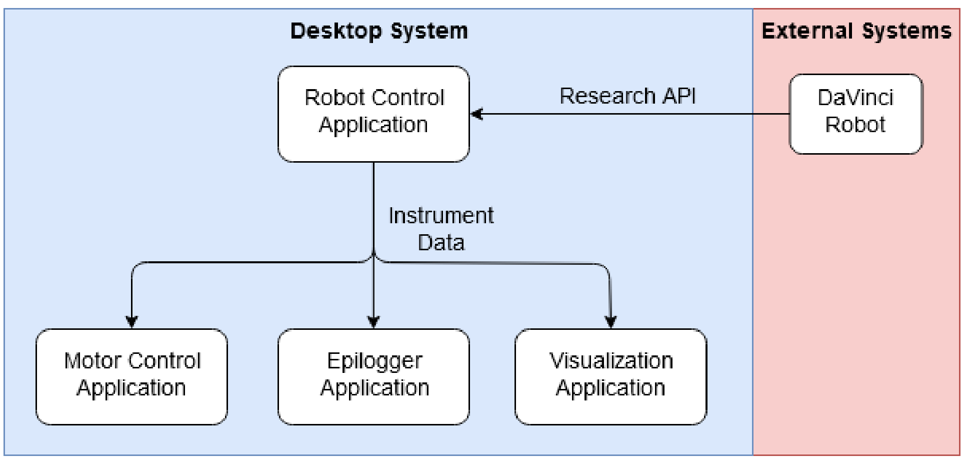


Figure 3.4: Overview of control signal clients

3.3.2 Motor Control Application

The current motor control system has been adapted from Mohareri Omid et al. [33]'s work to perform two specific tasks: track the da Vinci instrument tool tip and obtain a 3D volumetric sweep of the prostate. The system itself consists of two applications, the eScan project and the motor control application.

In order to track the da Vinci instrument tool tip, a calibration process is required for determining the transformation between the da Vinci instrument and the ultrasound machine. The full algorithm is explained in greater detail in Mohareri Omid et al. [33], but a summary is as follows:

1. The surgeon touches the prostate with the da Vinci instrument, moving the instrument up and down to create an artifact that can be localized in the ultrasound image
2. User manually rotates the ultrasound probe until instrument tool tip artifact is visible
3. The artifact coordinates are selected in the Application
4. The current da Vinci instrument location along with the ultrasound coordinates and motor roll angle are recorded
5. { 1. through 4. } are repeated four times for different points on the prostate
6. The rigid transformation between the two sets of coordinates is found from the two sets of robot and ultrasound artifact coordinates.

The eScan project is responsible for interfacing with the Sonix ultrasound machine and also features an interface for selecting points of interest. The motor control application is a simple console application with different options for enabling the different functions of the motor. One particular mode is the sweep mode which is used after the calibration process. This sweep mode has the motor move from its minimum rotation angle to its maximum angle. During this movement, each plane of the ultrasound image is saved into a file resulting in a full 3D scan of the prostate.

3.3.3 Registration Application

The registration application was developed for the purpose of performing a manual registration between the MRI volume and the 3D TRUS volume obtained during the surgery. An initial version of this application was already developed prior to the start of this thesis with many new features and improvements introduced over the course of this thesis.

To perform the manual registration, a segmented MR volume is first specified. Based on this segmentation, a tetrahedral mesh is generated along with organ-specific elasticity values that will be used by the model. These values are taken either from generic parameters reported from the literature or computed from elastography. This mesh is used throughout the project as the main 3D representation of the prostate. After this mesh is created, the 3D TRUS volume is loaded. In order to properly use this volume, it must be first processed through a scan conversion algorithm. This scan conversion is required to transform the Cartesian coordinate volume back into polar coordinates. With this completed, the boundary of the 3D mesh is projected on the current ultrasound plane. This boundary in blue, and can be seen in Figure 3.5 in both the transverse and sagittal planes of the current ultrasound volume. The segmentation is performed by a trained resident during the surgery by translating, rotating or deforming this prostate boundary. After the resident is satisfied that the registration is accurate, the registration is saved to be later used by the visualization application.

3.3.4 Visualization Application

The visualization application is the main focus of the project and produces the virtual prostate scene. In order to render the prostate at the proper position, a series

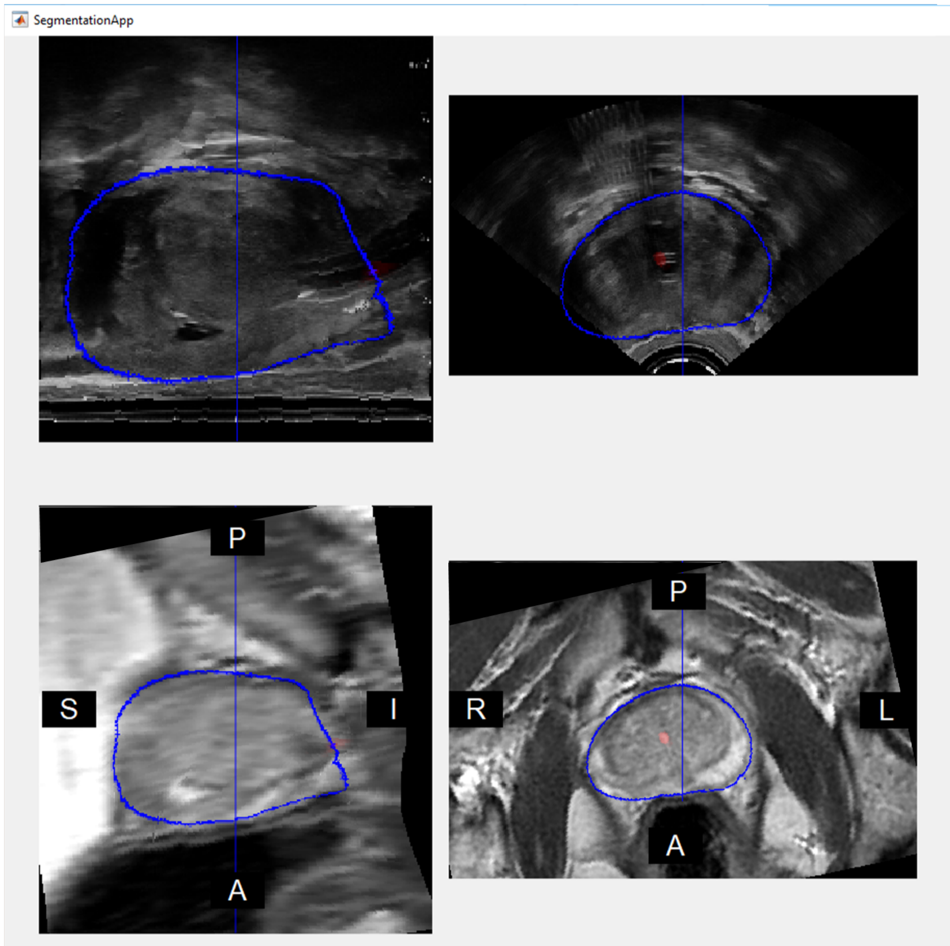


Figure 3.5: GUI of manual registration application

of transformations shown in Figure 3.6 are applied. The prostate object mesh in the scene is the same mesh created by the registration application. This mesh is then deformed and transformed using the registration previously created by the expert resident. This brings the mesh from the MR coordinate frame into the ultrasound coordinate frame. The mesh can now be in the da Vinci coordinate system by applying the transformation found during the motor calibration step.

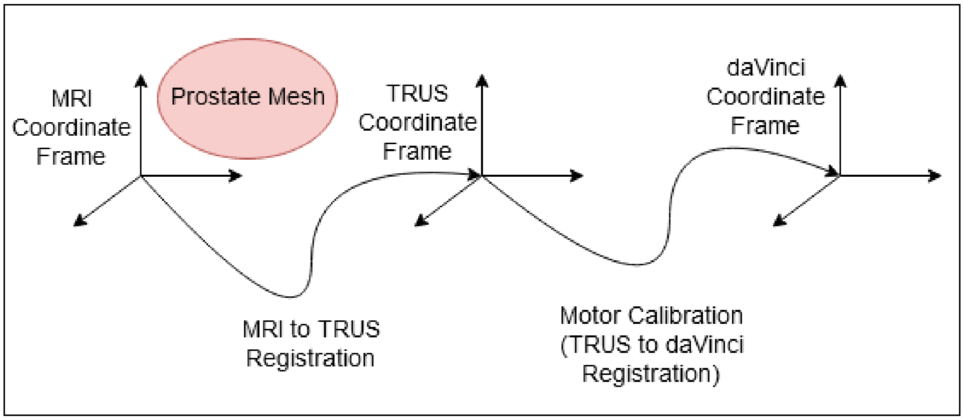


Figure 3.6: Transformations for visualization application

With this transformation complete, the prostate mesh is rendered in the da Vinci coordinate system. This allows the endoscope camera position and orientation to be used directly as the renderer camera parameters and positions the virtual scene in the same orientation and position as the true prostate. The da Vinci tool tip position is represented in the form of a virtual instrument whose position is updated in real-time.

3.3.5 MRI-TRUS Fusion Framework

MRI is often acquired weeks in advance of surgery. The location of the prostate during surgery changes, because the patient is in a different position and orientation in surgery *vs.* the MRI scanning session; as well, there may be changes in the prostate that are due to the different times at which the MRI and TRUS are acquired. In order to compensate for these differences, a deformable registration between the MRI and TRUS must be performed. Prior to the surgery, the MRI volume is segmented by an expert along with any cancerous lesions. A finite-element model constructed from this segmentation is used to represent this deformation.

Slicer Framework

One of the main challenges of an image guidance system for RALRP is the need for real-time performance. Once the calibration and registration process are completed, the mesh can be updated. However, to sample the MRI volume based on this registration requires calculating the entirety of the deformed MRI volume and is too computationally expensive for real-time performance.

One possible approach for solving this sampling problem is to first deform the entire volume before determining the image intensities in the region of interest. While this approach is commonly used, it is impractical for our application. Not only will the deformed voxels no longer have a regular grid structure and require expensive interpolation techniques to solve for intensity values, but the deformation must be applied to the entire volume which is computationally expensive for real time performance.

As an alternative, Goksel et al. [10] proposed a method for determining the subset of the volume which have an effect on the intensity values of the area of interest. As the user will only view a single MRI plane at a time, only a single slice sample of the MRI volume needs to be deformed based on the registration. The main algorithm consists of three main steps:

1. For every pixel p in then image plane:
2. Find the finite element e enclosing the pixel p
3. Find the location of the undeformed pixel p_0 based on the nodal displacements of the element e
4. Interpolate the volume intensities at pixel location p from the values at the deformed nodes enclosing it.

Both steps 2 and 3 are constant time operations and have many known solutions available. Step 1 on the other hand, is a very expensive operation and the main challenge of the algorithm. This step is known as a “point location problem”, in which the region where a point lies needs to be determined. First, a set of all the elements which are intersected by the desired image slice are determined. This set is topologically sorted with their positions discretized into image pixels. This image is then traversed using parallel scan lines to determine the relevant element contributing to the current pixel.

Implementation

To show the MRI volume, the mesh is sliced at the axial position corresponding to the position of the da Vinci instrument tool tip. At this plane, the respective projection of the registered MR image slice is displayed. In this version of the application, the display is split into two, with the left showing the virtual prostate scene with the right panel displaying the transverse plane of the MRI volume.

Chapter 4

Augmented Reality Imaging Pipeline

4.1 Introduction

The first objective of this thesis is the integration of a new ultrasound machine acquired by the lab. This new ultrasound machine, the bk3500, while featuring superior performance over the previous Sonix ultrasound machine, has a few limitations in which a workaround must be developed.

4.2 Ultrasound Integration

A bk3500 ultrasound machine from BK Ultrasound with an E14CL4b 14-4 MHz biplane transducer is to be integrated into the guidance system. As a biplane transducer, the probe can be used to image in two separate planes at the same time if desired. These planes consists of the transverse and sagittal planes as shown in Figure 4.1.

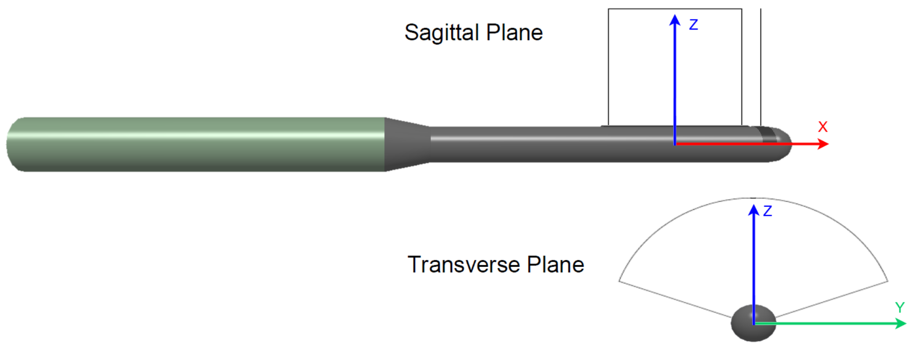


Figure 4.1: Transducer probe coordinate system and anatomical planes

There are two options when using this transducer probe to view the entirety of the prostate and not just a single plane. While using the sagittal imaging plane, the probe can be rotated about the X-axis to form a full sweep of the prostate. Alternatively, while viewing the transverse plane, the probe in can translated along the X-axis back and forth. The sagittal sweep does not change the transducer configuration inside the patient, and for this reason it is fundamentally safer. As well, the sagittal linear array has a constant lateral resolution. The rotational movement was determined to be sufficient for imaging the prostate. The ultrasound machine is configured to acquire ultrasound images with a size of 6.5 cm $\times$ 6.5 cm at a resolution of 3 pixels/mm at 34 Hz. This machine is primarily used to image the prostate and is configured to provide an image which delineates the prostate boundary from its surroundings.

While the bk3500 provides much superior image quality than the Sonix ultrasound machine used previously, it has certain limitations that causes it to be incompatible with the previous workflow. In the prior system, the ultrasound image was obtained by streaming the B-Mode image directly from the machine. While this was attempted with the BK 3500, the images would arrive only after a 10 second delay which prevents any real-time usage of the image.

Alternatively, it was attempted to reconstruct the B-mode images from the RF data of the machine, but the reconstructed images were of much inferior quality than the images seen on the machine's display. This can be seen in Figure 4.2 when the captured B-mode image are of much superior quality to the reconstructed B-mode image. As the algorithms for decoding the images are proprietary to the ultrasound manufacturer, and typically involves many years of tuning the signal processing pipeline, it was not possible for the image to be replicated with the same quality as the native bk3500. As a workaround, it was determined that the simplest solution would be to directly obtain the images seen on the ultrasound machine screen through a hardware-based frame grabber card.

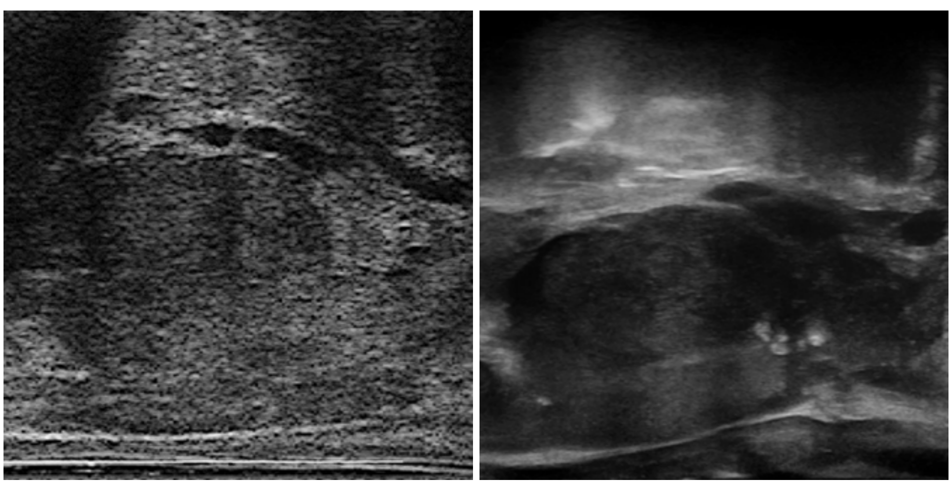


Figure 4.2: Comparison between RF-reconstructed (left) and native B-mode (right) ultrasound images

4.3 DeckLink Framegrabber Card

Conventional desktop computers are only capable of producing video outputs and cannot acquire external video signals. For a computer to do so, an additional expansion card known as a frame grabber is required. Framegrabber cards are have conventionally been mainly used by the film and broadcast industry. For many of their applications, computers with frame grabbers serve as intermediaries between systems. As such, many frame grabbers are designed with both input and output in mind.

For the purposes of this project, with the long term goals in mind, at least five input/output ports are desired for the frame grabber card. The inputs consist of the ultrasound video stream as well as two ports for left/right endoscope camera video feed. The outputs consist of the left/right channels of the virtual prostate video stream.

While there is a collection of different frame grabber cards available commercially, it was determined that the DeckLink Quad 2 was the most suitable frame grabber card available. The DeckLink Quad 2 features 8 bi-directional channels which can be configured independently with different video format standards. Along with the frame grabber, a Software Development Kit (SDK) is also included. This SDK is available for both Windows and Unix based operating systems.

Most other frame grabber systems were rejected as they did not have sufficient ports for the current project. While it is possible to install multiple frame grabber cards into a single desktop computer, it will be much more expensive than a single card. Multiple cards may also cause synchronization problems between video from each card. Another consideration was the type of port used by the frame grabber.

Many frame grabbers use the Small Form-factor Pluggable (SFP) or CoaXPress standard which are incompatible with the video sources used by the da Vinci robot. While some of the ports can be converted into a Serial Digital Interface (SDI) connector through some additional hardware, this adds more complexity and cost to the system than necessary.

This frame grabber system was determined to be the ideal solution for the ultrasound interface as it solves multiple problems through a single hardware system. Not only can the video stream of the ultrasound machine be obtained, but a video signal can also be sent to the da Vinci surgical console with the same hardware system. Figure 4.3 shows how the frame grabber card now fits in the overall system.

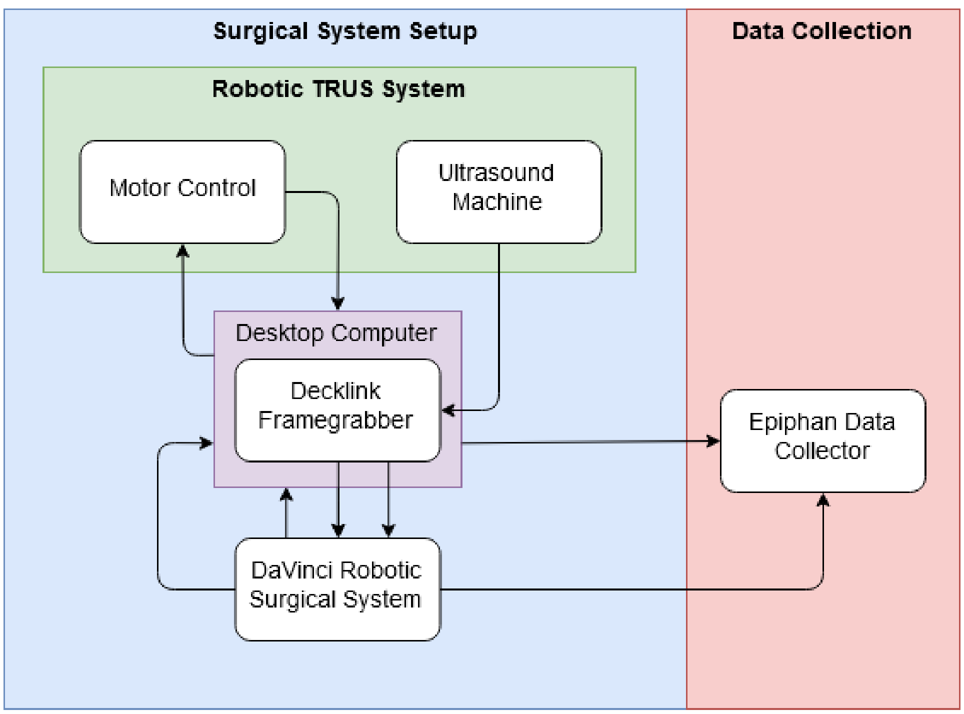


Figure 4.3: Block diagram of system setup with DeckLink framegrabber card

4.3.1 Limitations

The main hardware limitation of the DeckLink frame grabber card lies in its interface type; the DeckLink card features only eight mini-SDI connectors. SDI connectors are not a frequently used connection type, and there are video sources which need to be captured that only have conventional video output ports. More specifically, the bk3500 machine required by the project only has a Digital Visual Interface (DVI) connection. In order to acquire a video signal from such a port, some form of conversion is required. For the specified DVI port, the signal was first converted into an High-Definition Multimedia Interface (HDMI) signal then converted into an SDI signal.

4.4 Imaging Pipeline Framework

The imaging pipeline consists of a software library developed to interface with the DeckLink Quad 2 frame grabber card through. While Blackmagic Design Design has provided with a comprehensive SDK for interfacing with the frame grabber card, the SDK API is rather complex and does not offer much documentation. The imaging pipeline framework is largely designed as a wrapper around this third party SDK. The primary goal of the imaging pipeline library are primarily to make the interface to the framework card as simple as possible, which ensures there is little cost to the performance of the system. The full Unified Modelling Language (UML) diagram of the framework can be seen in Figure 4.4.

4.4.1 General Workflow

The operation of the frame grabber framework begins with the creation of a manager object. This object is responsible for opening and closing ports as well main-

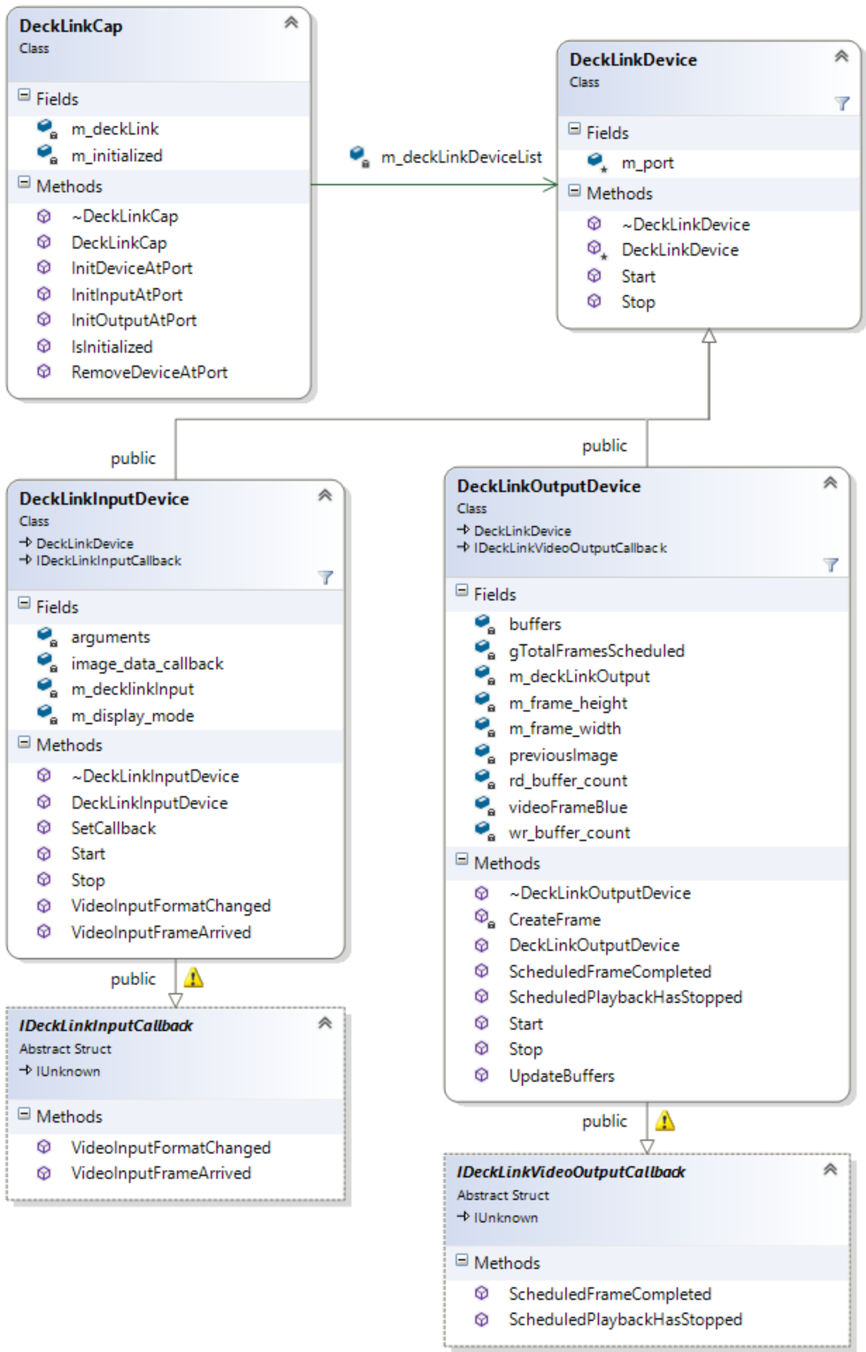


Figure 4.4: UML diagram of DeckLink framework

taining the list of initialized ports. A port is initialized with a port number and as either an input or an output. Once initialized, ports can be started or stopped to begin or end the capture or output process. Input ports need to be provided with a callback function for handling input frames. In order to simplify usage as much as possible, all Blackmagic Design SDK specific logic is abstracted away from the frame grabber library.

4.4.2 DeckLink Capture Class

`DeckLinkCap` is the main object used for initializing the DeckLink frame grabber system. During the initialization process, all the possible ports of the currently installed frame grabber card are initialized and their handlers saved into an array. These handlers will be later used to produce `DeckLinkDevice` objects for direct interaction with a specific port. During this process, a failure to detect any ports signifies improperly installed drivers or no frame grabber card detected. By calling a port initialization function, an object of the corresponding type will be created from the port handling object. During this process, a list of all initialized ports are kept in order to ensure that no two ports are used at the same time. This class also provides a port de-initialization function. This is needed for two primary reasons: to allow ports to be initialized so another application can take control of the port, and to change the type of the port.

4.4.3 Base Abstract Class

`DeckLinkDevice` is an abstract header which serves as the parent for the other two device objects. This object contains the functions Start and Stop which are to be implemented by all sub-classes. The Start function is used by both sub-classes to begin video capture or output. The Stop function is used to properly de-initialize port objects and end the capture or output process.

4.4.4 Video Input

`DeckLinkInputDevice` is the object produced after a successful port initialization. The main purpose of this object is to implement the Blackmagic library input callback interface. This provides the object with two main callback functions. One callback is triggered whenever a new frame is obtained and provides a video frame object which contains the characteristics of the frame such as resolution and size. Another callback function provided is called whenever the video format changes. This callback can be used to update the video format if it ever changes.

4.4.5 Video Output

`DeckLinkOutputDevice` is the other object produced from an output port initialization. The purpose of this object is to maintain a buffer of video frames and output them at the rate defined by the specified frame rate. The buffer can be updated with new image frames anytime by calling the corresponding function. An internal callback function will be triggered whenever a new frame is required for the output. As mentioned before, a consistent frame rate must be maintained in order to conform to the video standard requested by the frame grabber system. Unfortunately, the rate at which frames can be drawn may not necessarily match the rate at which frames need to be output. In order to handle such a problem, a multi-frame buffer is introduced. This buffer saves the last few frames provided to the frame output framework. If output frames are slower than the required frame rate, then duplicates of the last frame will be output to maintain the frame rate. If output frames are faster than the required frame rate, then the oldest frames will be skipped. This bypasses the strict frame rate requirements of the DeckLink SDK.

4.4.6 Limitations

There are a few limitations of the frame grabber card that need to be considered when developing the software for the system. Primarily, as a professional quality card designed for broadcast applications, the DeckLink frame grabbers must conform to broadcast video standards. This results in the system requiring very specific resolutions, frame rates and aspect ratios. While the resolution and aspect ratio does not pose too large a problem, the frame rate requirements can be a potential problem. The DeckLink framework was designed to ensure that the system output is maintained at the required frame rate in order to produce a steady video signal.

Video Standard Restriction

The most inconvenient limitation of the DeckLink frame grabber its limited set of broadcast video format standards. When initialized, a specific resolution and frame rate must be specified for the video signal of interest. The list of possible resolution and frame rates is quite restrictive and vary depending on the model of the frame grabber card. If the currently selected resolution is smaller than the provided video signal, the frame grabber will simply ignore the signal as it considers it a mismatched signal and therefore invalid. This mismatch problem also applies if the chosen frame or field type of the video signal is different from the input signal where the frame grabber ignores the input. While there is no way to bypass the field and frame rate restriction, a larger resolution can be chosen if no other similar resolution is available. This input signal mismatch problem can be somewhat alleviated by introducing an automatic format detection system. While this function can prevent the frame rate or field type mismatch problem, it cannot prevent the resolution restriction problem completely.

To bypass the resolution restriction, a larger resolution can be chosen but will require some post processing to resize and crop the images to the desired original resolution. This resolution restriction is especially inconvenient when the desired video source is in a different aspect ratio. In order to read a 4:3 aspect ratio video such as Super Extended Graphics Array (SXGA), the closest available resolution is HD 1080 which results in a highly distorted and stretched image which must be later resized for the proper resolution. These conversions not only degrade the video quality, they add additional overhead which cannot be easily eliminated.

Image format

When video frames are obtained from the DeckLink frame grabber card, they arrive in UYVY YUV 4:2:2 pixel format. In this format, the image bytes are stored in the form of macro pixels. Each macro pixel contains the colour and intensity information for two pixels. In order to obtain an RGB image, a conversion process needs to be applied to the image pixels.

4.5 Stereo Output

Depth perception is the ability to perceive the dimensions and relative distance of an object. This illusion of depth is provided through a variety of depth cues. While accurate depth perception generally requires binocular depth cues, there are also monocular depth cues which aid in the perception of depth. Stereopsis or also known as "binocular disparity" is the dominant cue which is most responsible for providing the depth effect. While stereopsis is the most important depth cue used by humans to judge the depth of objects, other secondary cues also play a role and may detract from the overall effect if not implemented properly. Binocular disparity occurs when observing images of a scene from two slightly different angles. A stereo pair, or the image seen from each individual left and right eye, is slightly different and allows the distance to an object to be obtained through triangulation.

One of the major limitations of the da Vinci surgical system is the lack of haptic feedback. Without haptic feedback surgeons must rely solely on their depth perception to ensure their instruments are where they should be. As such, all da Vinci systems have stereo endoscope cameras. When operating, surgeons view the surgical scene through the da Vinci console which provides a separate display screen for each eye. This provides the stereopsis depth cue during the operation.

While the endoscope system has this 3D stereo mode, the custom inputs provided to the da Vinci TilePro system do not have a stereo effect by default. While it is possible to configure the da Vinci console to display two separate input signals for each individual display screen, the stereopsis effect must be manually introduced. While depth perception is largely unnecessary when presenting 2D information such as Ultrasound of MR images, depth cues are helpful when presenting virtual scenes or for AR. In the virtual scene rendered in the developed application, a prostate is simulated along the instruments of the da Vinci system. Through this virtual scene, the goal is to provide additional context and information to the surgeon such as the location of tumour or simply the MRI of the prostate. Thus, for our current application, depth cues will greatly increase the effectiveness of this information, which can potentially lead to improved positive margins and thus achieve the major goal of the project.

4.5.1 Algorithm

In order to set up a stereo effect, two cameras or two images are required. The problem lies on how to position these cameras in order to obtain the ideal stereo effect. There are two commonly used methods known as the converged projection or "toed-in" method and the parallel projection method. When dealing with physical cameras, each method has their specific advantages and disadvantages. However, in a virtual environment, there is an ideal solution which generates an effect with the best of both solutions.

Asymmetric Frustum Parallel Projection Planes is method of producing a stereo effect in a virtual environment without any of the drawbacks of other techniques. This technique involves two parallel cameras at an eye separate distance with both cameras pointed straight at a target which can be seen in Figure 4.5. The two cameras' image planes are offset slightly which is analogous to the horizontal shift performed with physical camera images. This image plane offset can be done with an asymmetrical frustum by modifying the camera projection matrix. In most 3D rendering frameworks such as OpenGL, this off-axis method can be easily performed.

4.6 TilePro Simulator

The da Vinci TilePro system is often used for facilitating research for the da Vinci platform. Research applications range from Magnetic Resonance Imaging and Ultrasound guidance [33] for robotic surgery to Augmented Reality systems [44] for

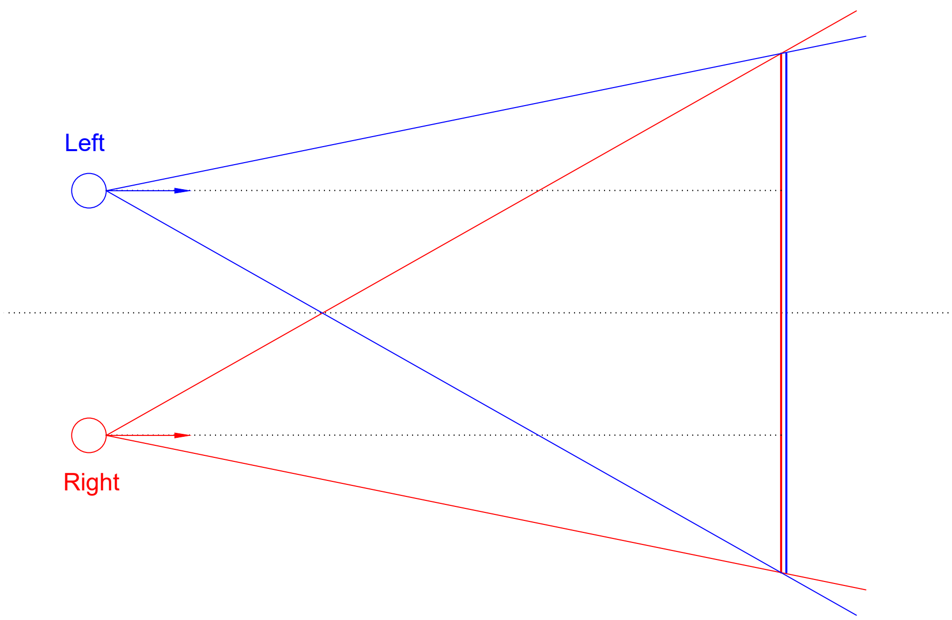


Figure 4.5: Asymmetric Frustum Parallel Projection Planes for stereo effect

the da Vinci robot. In a majority of research for the da Vinci platform, additional data needs to be presented to the surgeon console during its operation. However, the standard endoscope video in the console cannot be modified in any way that may endanger the patient if any problem is encountered. As a result, the TilePro system is an invaluable tool as it allows for additional video inputs without any risk to the patient.

While TilePro is available for the latest da Vinci models, its usage is more limited in older da Vinci models. On systems without TilePro functionality, important display based research cannot be implemented and tested, severely limiting the usefulness of these otherwise functional older models for research. This presents a problem as research tends to be conducted with older systems no longer used in the operating room. The DaVinci Research Kit (DVRK) is a collection of components from the original da Vinci robot used as a research platform for surgical robotics research. As it is based on the original da Vinci model, it too does not have TilePro functionality. As running the robot in DVRK mode is the only way of controlling the robotic arms without the conventional da Vinci operating system, it would be beneficial for TilePro to be introduced for the DVRK system as well.

The TilePro system has been available for the da Vinci surgical system ever since the second generation model. The TilePro system allows for up to two video inputs; they can either be displayed side by side or they can be fused together to form a single stereo video. The 3rd generation da Vinci robot model's surgical console has a touchscreen panel which allows for the adjustment of TilePro settings. The 2nd generation da Vinci robot has no such interface and relies on the console foot pedals to enable or disable the TilePro function. The TilePro system specifications can be seen in Table 4.1.

Table 4.1: da Vinci console TilePro specifications by generation

Robot Model	da Vinci S	da Vinci Si
Resolution	1024×768	1280×1024
Frame Rate	59.94	59.94
Connections	DVI	DVI, SDI

4.6.1 System Setup

The TilePro simulation consists of two main components, a procedure for setting up the da Vinci robot and supporting software. This setup process depends on the robot model as the console specifications are different. In either case, the main setup consists of a desktop computer to serve as an intermediary system between the da Vinci endoscope camera and console display. While this intermediary system cannot be used for the Operating Room (OR), it serves a simulation on how the TilePro system will work for a newer da Vinci Si model robot.

A conventional desktop computer with a frame grabber card installed was used as the main platform for the proposed system. By using the previously created DeckLink framework, frames from the endoscope camera can be acquired, while a custom fused scene can be output to the da Vinci console. In order to properly replicate the surgical scene, two channels are required for the stereo camera output frames to he grabbed into the desktop computer. An additional two channels are required as output from the desktop computer to send the fused stereo frame back to the da Vinci console, by using the da Vinci console native camera ports. Thus, a minimum of four channels are required, with an additional channel necessary for each extra input source wanted. The reason why such a setup cannot be used in surgery is that there is now a computer interfaced between the endoscope camera (the surgeon's eyes into the patient) and the console. If anything goes wrong with the computer, the surgeon can no longer see the surgical scene.

The configuration for the simulation system changes slightly depending on the model of da Vinci robot used. In order to properly replicate the camera signal without any compatibility problems for the 2^{nd} generation da Vinci robot, an DeckLink output video at 1920×1080 interlaced with a frame rate of 59.94 is required. These settings can be configured in the DeckLink framework produced in Chapter 4. The first generation da Vinci model running the DVRK system requires an output signal with the 525i National Television System Committee (NTSC) analog standard with a frame rate at 59.94. This can be configured in two ways, by either producing a NTSC signal from the frame grabber framework, or by producing a standard HD signal and enabling the convert to Standard Definition (SD) setting in the frame grabber firmware settings.

4.7 Conclusion

A framework was video input and output has been developed to be used for multiple purposes. First, the framework is used to integrate the new bk3500 ultrasound machine into the project. Second, the framework can now be used to push images into the TilePro system of the da Vinci surgeon console. A TilePro simulator system has also been setup to help facilitate development for older da Vinci robots. Depth perception has been introduced to the 3D prostate simulation through stereopsis or binocular disparity. This effect was created by producing two images of the same scene from different perspectives to produced a stereo effect. Surgeon feedback has been positive with the surgeon capable of perceiving the depth of certain objects relative to others.

Chapter 5

Motor Control and Ultrasound to da Vinci Registration Application

5.1 Introduction

Ultrasound imaging is the central part of the system workflow with its ability to provide real time visualization of the prostate anatomy. While ultrasound images are highly informative and provide a good view of the prostate, they are all relative to their own reference frame. If there is to be any interaction between the da Vinci system and the ultrasound images such as the virtual surgical scene or augmented reality display, a transformation between the two coordinate frames must be obtained.

The ultrasound machine to da Vinci robot coordinate system transformation is calculated through a process known as registration. Point set registration is a process where the spatial transformation between two sets of points which provides the best alignment is determined. This transformation can be either rigid or non-rigid. The problem at hand involves a single rigid body transformation from one coordinate system to another. The points required for this registration are obtained through a calibration procedure performed during the surgery.

Once this calibration process is completed, the motor control application has two more purposes to fulfill. First, a 3D volume of the prostate must be obtained for the MRI-TRUS registration algorithm. This is completed by having the TRUS robot move from its minimum angle to its maximum angle which each ultrasound image is saved into a file. Once this is completed, the tracking mode is enabled in which the TRUS robot automatically moves the robot to follow the da Vinci instrument tool tip using the transformation obtained by the registration process.

As introduced in Chapter 3, all of these functions have already been developed prior by [33]. However, with the introduction of the new ultrasound machine, the DeckLink framework developed in Chapter 4 must be integrated into the application. This new ultrasound interface also renders the previously developed *eScan* application redundant. In order to simplify the workflow, a new application which combines the motor control system and ultrasound image selection interface will be developed.

5.1.1 Ultrasound Integration

The bk3500 features an option to clone the current screen display as an output to another screen through a DVI connection. By converting this DVI connection into a SDI one, the US machine display was successfully connected to the frame grabber system. Once connected, the frames were obtained with a resolution of 1920×1080 and a frame rate of 59.98 Hz. As mentioned before, one of the limitations of the frame grabber system is its limited range of usable image sizes. While the US machine clones its display at a resolution of 1280 × 1024, there is no choice but to use a higher resolution or the frame grabber system will simple ignore the signal. Once obtained, the image is re-scaled to its proper resolution and cropped for the US image. During normal operation, the pixel location of the US image does not change. As such, the image can be repeatedly cropped without needing to adjust this pixel offset coordinate.

At this time, the pixel to mm scale can also be determined through the unit bars seen the image. This information is used during the US interaction step later to determine the location of the user selected pixel coordinates in mm.

5.2 Motor Control Application

The motor control application is designed as a Graphical User Interface (GUI) interface for manipulating the TRUS robot used for manipulating the ultrasound transducer. The application allows for three modes of motor operation: manual rotation control, position to rotation tracking, and full rotational motor sweep. The manual motor control is used to perform a calibration procedure for determining the registration between ultrasound and da Vinci robot coordinate systems. The position to rotation tracking is responsible for tracking the da Vinci instruments by using the same framework previously developed by Adebar et al. [1]. The mode of operation is used to obtain a 3D volume of the prostate by sweeping through the full rotation of the probe motor.

The application itself is an Microsoft Foundation Classes (MFC) dialog-based application which serves as an interface to the motor control library previously developed by members of the lab. The GUI is separated into two regions: one for changing the motor states, the other for displaying a preview of the live ultrasound

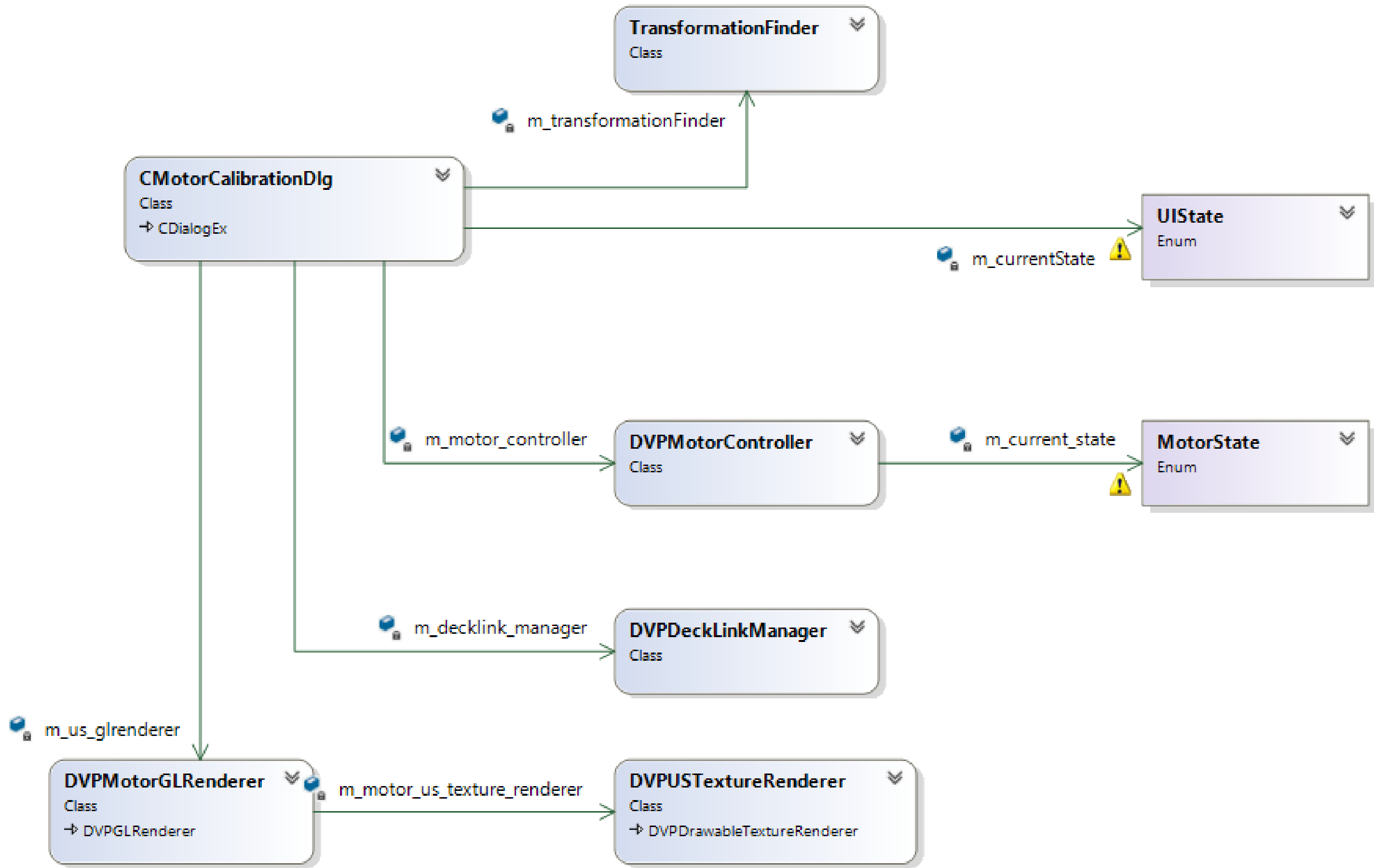


Figure 5.1: Simplified UML diagram of motor control application

image. A UML diagram of the system design can be seen in Figure 5.1.

During the calibration procedure, points on the ultrasound image are selected by clicking within the preview screen. When a point is selected, a crosshair is drawn on the screen to indicate the current chosen location which can be seen in Figure 5.2. This selected point is retained even if the dialog window size is changed.

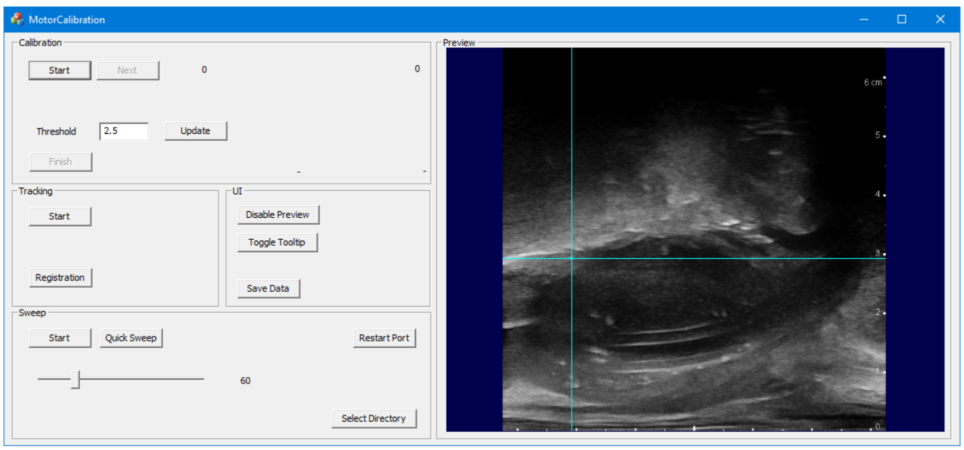


Figure 5.2: Motor control GUI application during calibration mode

Taking full advantage of the DeckLink framework, the real time ultrasound image is also sent to the da Vinci surgical console during the calibration procedure. Not only will the surgeon be able to see the current ultrasound image, they will also be able to see the current point selected by the user. This allows the surgeon to provide their feedback on whether or not the select point is on/off plane.

5.2.1 Motor Control Dialog

`CMotorCalibrationDlg` is the main class responsible for the GUI elements of the project. The class is responsible for handling User Interface (UI) interaction. The class is also responsible for creating the objects needed for the DeckLink framework and ultrasound texture rendering. In order to better manage the different states of the application, a state machine was implemented to control the behaviour of the different GUI elements. Depending on the current system state, certain UI elements will be enabled or disabled. This allows for the system to be state to be strictly controlled and prevents non-deterministic actions from happening.

5.2.2 Motor Controller

The `DVPMotorController` is a class responsible specifically for interacting with the TRUS robot motor. The class also uses a similar state machine setup like that of the ultrasound interface with three main states: manual motor control, sweep mode, and tracking mode. In the manual motor rotation mode, the keyboard arrow keys are used for controlling the robotic probe movement. Unlike the prior system where the motor keeps moving until a stop command is given, the current system is designed to stop the motor movement the moment the keyboard button is released. This allows the ultrasound probe to be more easily positioned and, and makes it easier for the operator to learn how to use it. Along with this, an additional change was introduced in the form of a modifier key, which, when held, doubles the speed of the motor. This dual speed control provides the user with two modes of operation, a fast mode for rotating the probe quickly and a slow mode for the accurate positioning of the probe.

5.3 Calibration Process

Calibration methods can be broadly classified as either automatic or manual methods. Determining the transformation between medical imaging and da Vinci coordinate frames is a well established problem with many different solutions. Some automatic methods use anatomical landmarks seen in both the da Vinci endoscope cameras and medical images to perform the calibration. Unfortunately for our specific problem, there aren't enough landmarks that are visible in both the US image and da Vinci endoscope cameras. We have chosen to use a manual calibration procedure in order to obtain a set of points required for the registration process.

The manual calibration procedure functions by locating features on both the US image and the da Vinci system in which the coordinates are known. For our setup specifically, this calibration procedure functions by locating the tip of a da Vinci instrument placed on the anterior surface of the prostate. As the location of the da Vinci instruments can be determined through the da Vinci research interface, only the location of the instrument artifact on the US image needs to be identified.

The calibration technique begins with the surgeon palpating the surface of the prostate with the da Vinci instrument tip (monopolar curved scissors). The palpation greatly aids in the identification of the tool tip of the instrument in the US image by providing motion which can be more easily identified. The TRUS robot probe is manually rotated until this motion can be seen. When the tool tip artifact can be identified, the angle where the tool tip artifact has the highest intensity is then determined in order to determine the centre of the tool tip. Once found, the pixel coordinates of the tool tip is determined through the user interface. This pixel location is converted into mm through the scaling provided by the US image. This procedure is repeated at least three times. With a minimum of three sets of coordinates, a registration can be performed to determine a rigid transformation between the ultrasound and the da Vinci coordinate frames.

While the procedure done by Adebar et al. [1], Mohareri Omid et al. [33] used four points in order to perform the calibration, we have extended the technique to function with as many points as desired. While more points may not necessarily provide a superior registration, they allow for greater redundancy and a chance for recovering from a poor registration. The major limitation of the previous work was the system's inability to handle outliers or incorrectly chosen coordinates. As a result, no matter how many points are chosen, the transformation solver only creates the transformation which best fits all the specified points. The system has no feedback mechanism informing the user if the chosen coordinates provide an accurate transformation or not. The calibration process was improved in three primary ways. First, a new algorithm for outlier rejection was introduced which will be further discussed in Section 5.3.1. Second, two feedback mechanisms were introduced to provide the user with a qualitative mechanism for determining if the specified coordinates are accurate or not; these are discussed in Section 5.3.2. Third, the calibration process is no longer restricted to four points and more points can be chosen if it is determined that the previously chosen points are inaccurate.

5.3.1 Outlier Rejection

Two algorithms were introduced as a method of improving the reliability of the calibration. Both algorithms function near identically with the main difference in how they select calibration points for generating their models. The first algorithm is an exhaustive algorithm where all the possible combinations of points are tested. The second algorithm is RANdom Sample Consensus (RANSAC) which the points selected are chosen randomly instead. The core algorithm consists of the following steps:

Exhaustive Algorithm

1. Generate all non-repeating 3 point combinations of indices of the given calibration point list.
2. Fit a model to the selected points from given indices.
3. Calculate fitting error based on consensus set points and keep model with lowest error and most amount of consensus set points.

RANSAC Algorithm

1. Shuffle list of calibration points
2. Select a subset of list (3 points)
3. Fit a model to the selected points
4. Determine the number of remaining points that fit the model
5. Repeat until the maximum number of iterations (100 times)

The model which provides the largest number of fitted points with the lowest error is kept and the remaining points excluded as outliers. By using this algorithm, points that are chosen incorrectly can be ignored completed instead of skewing the transformation. An example of this can be seen in Figure 5.3 where a single outlier results in the entire transformation shifted right. With the algorithm applied, the outlier point is rejected which allows for a better fitting transformation.

While the exhaustive algorithm will most likely provide the best transformation, its performance does not scale well as each additional calibration point would increase computational time exponentially. As such, if the list of calibration points exceed 10, the RANSAC algorithm will be used instead. When testing in the OR in a surgical study, it is highly unlikely that the number of calibration points will

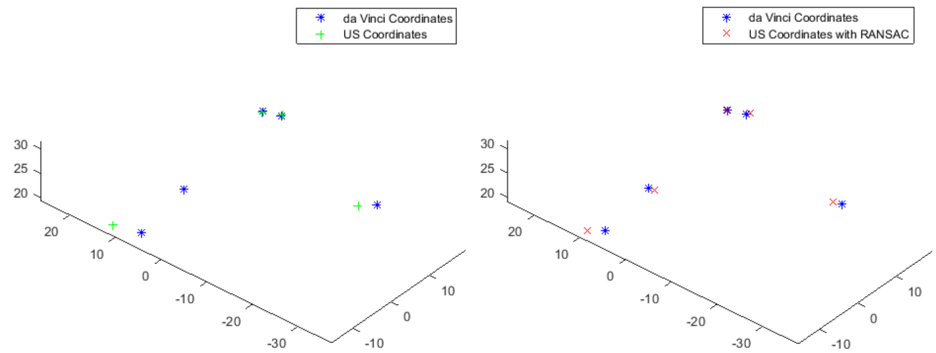


Figure 5.3: Comparison of transformation with and without RANSAC

exceed 6 points. The RANSAC algorithm is largely in place as a backup in case it is required.

5.3.2 Feedback Mechanisms

The current most frequent point of failure of the visualization application lies in the calibration process. While the calibration process is quite simple, the difficulty lies in determining if the calibration obtained is accurate. Currently, the only way to determine this is to run the visualization application and compare the camera orientation with the current endoscope camera position. Unfortunately, this currently cannot be done until the MRI-TRUS registration is completed and by that time, it is too late to modify the calibration. In order to improve the robustness of the calibration procedure, a few more tools are provided along with the RANSAC outlier rejection algorithm.

Visual Feedback

A visual feedback system was also introduced in the form of da Vinci instrument position indicator. On top of the ultrasound preview screen, a crosshair which indicates the position of the da Vinci instrument tool tip is drawn. This tool tip indicator can only be shown when at least three points have been chosen to provide the initial estimation of the transformation. This tool tip indicator provides a real time comparison for the user on how accurate the current registration as the distance between the indicator and the tool tip artefact is the Target Registration Error (TRE) of the system. As the calibration improves, the TRE should decrease with the crosshair beginning to match the artefact position. This provides the user with a quick way to determine if their chosen points are reducing the TREs of the calibration. Another benefit of the crosshair is in providing a rough guide on the plane of the tool tip artifact. This is of particular use when locating points which are of a significant distance away.

Exhaustive and RANSAC Parameters Feedback

The fitting error from the sample consensus algorithms can be shown which provides an indicator on how well the currently chosen points fit the determined transformation. This fitting error, also known as Fiducial Registration Error, is a measurement of how well the points fit the current transformation. Along with the error value, the number of consensus points is also reported. This lets the user to determine if additional points are necessary and if these new points help contribute to the accuracy of the calibration. With the current calibration protocol of four to five points, a single outlier may not of significant concern but any more than one outlier may indicate that the current calibration is not very accurate.

The display of both this values serve primarily as a sanity check for the system. It should be noted that there is no correlation between TREs and Fiducial Registration Errors (FREs) as reported by Fitzpatrick [7]. Neither the FRE nor the number of fitted points directly indicate whether or not the registration is accurate. Instead, a high FRE or low number of consensus points indicate a high chance that something is going wrong.

5.3.3 Point Selection

The main calibration protocol has been revised to use five calibration points. Instead of the previous system of selecting two points in the mid before the left and right points, three points are chosen in the mid region to provide the initial estimate of the transformation. This initial transformation is unlikely to be accurate as the three points co-linear, but it allows the crosshair overlay to be generated. This provides a rough guide on the locations of the left and right artifacts as well as more points for the sample consensus algorithms. When picking the next two points, additional points can also be added if the fitting error is too high or if there are too many outliers for a reliable transformation.

5.4 Results

5.4.1 Experimental Setup

While it is possible to quantify the improvements of the calibration process alone, it was determined that evaluating the combined performance of the entire registration pipeline is more valuable. In order to evaluate the performance of the combined registration, a phantom experiment was performed. The purpose of this study was to evaluate the combined error of both MR-TRUS registration and TRUS-da Vinci calibration in a user study. A custom prostate phantom, fabricated from a material commonly used in ultrasound phantoms, was created with markers placed within. Users were asked to position the da Vinci instruments to locations marked in the

TilePro visualization. The instrument location is then compared with the markers locations which provides a quantitative error score which evaluates the overall accuracy of the registration.

Phantom Creation

The phantom creation processs was developed by another member of the lab Leo Metcalf. The goal of the phantom is to replicate surgical conditions as much as possible. As such, the prostate phantom was produced from a 3D-printed mould based on a patient's anatomy. The phantom itself was made from an agar mixture [cite?] at different densities for two separate regions: one for the prostate itself and one for the bulk tissue surrounding the prostate. The agar 3D-printed phantom mould consists of four main components: (i) an outer container for the entire phantom, (ii) a negative mould for creating the prostate depression, (iii) a shaft for producing a cylindrical canal for the TRUS probe and (iv) a 3D-printed prostate mould. On the negative mould there are impressions which allow eight plastic beads to be placed into the phantom. These beads serve as fiducials which can be easily localized when seen in the ultrasound image.

The phantom fabrication process itself consists of the following steps as detailed in Figure 5.4

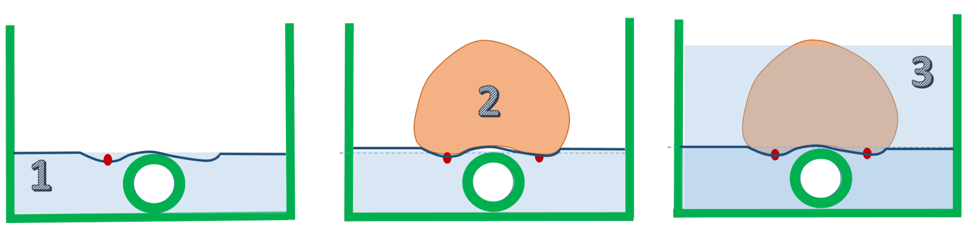


Figure 5.4: Phantom fabrication process

1. The probe canal is inserted into the outer phantom mould then half filled with the agar mixture with the negative mould inserted and allowed to settle
2. The prostate phantom is produced separately with a different agar mixture density using the prostate mould
3. Plastic beads are inserted into the depression imprinted by the negative mould and the prostate phantom is inserted into the tissue depression
4. The first agar mixture is added to the mould again until it completely covers the prostate phantom. Once the phantom is set, the probe canal can be removed to create the canal for the TRUS probe.

Phantom Study

With the phantom created, an experiment can now be conducted to determine the accuracy of the overall registration process. The experiment proceeds as follows:

1. Perform calibration using prostate phantom
2. Obtain TRUS sweep of phantom
3. Perform TRUS-MRI registration
4. Replace phantom with container filled with water
5. Have user place virtual tool tip on MRI bead locations
6. Obtain TRUS sweep of tool tip
7. Repeat for each bead (8 times)

The first step of the experiment is to perform the calibration procedure using the prostate phantom. Once completed, a full sweep of the phantom is completed in order to obtain a 3D TRUS volume. From this volume, the fiducial locations can all be located and will later be used to evaluate the accuracy of the registration.

With the 3D TRUS volume, the TRUS-MRI registration is performed with the same process detailed in Chapter 3. Once both registrations are completed, the phantom is removed without moving the ultrasound probe and is replaced with a container filled with water. Water is good at transmitting ultrasound and allows objects not in contact with the probe to be imaged. For this experiment, the purpose of the water tank is to make it easy to locate the da Vinci instrument tool tip in the ultrasound image. The visualization application is enabled with the da Vinci console display replaced by the virtual prostate scene. A user, by controlling the da Vinci robot instruments, attempts to move the virtual tool tip to the positions of the fiducials marked on the MRI image. The position of the instrument is then obtained by performing another TRUS sweep and repeated eight times for each fiducial. This tool tip position in the ultrasound image is compared to the actual fiducial position to obtain the total registration error of the system. The study was performed by four separate subjects and results of the experiment can be seen in Table 5.1.

Table 5.1: Results of phantom study experiment

Subject	Error
1	3.1 ± 1.2
2	3.0 ± 1.5
3	2.8 ± 1.2
4	3.9 ± 1.2
mean	3.19 ± 1.3

5.4.2 Discussion

The accuracy of the current registration process has improved significantly from the previous system setup which would result in errors over 8mm. While the current

TRE is not insignficiant, it is sufficient for the purposes of this thesis. As the goal of the thesis is to only assist the surgeon in determining whether or not a nerve-sparing operation is possible, rather than aid in the actual surgical procedure, 3mm or error is sufficient for that purpose.

Currently the calibration process is a manual process which is subject to human error. Determining the location of the da Vinci tool tip in the ultrasound image is often a difficult task. While localizing the artifact in the current ultrasound plane is a simple task, the difficulty lies in determining the proper angle in which the plane is located. Searching for the tool tip along the mid-line is not as difficult as not much rotation is required. However, the two points left and right of the mid-gland are more difficult to locate due to the amount of rotation involved. Searching for the current angle for the probe and changing the probe angle is often a time consuming task especially if the current ultrasound image is of poor quality.

Improvements could be introduced in the form of automatic tool tip detection such as the work of Mohareri et al. [32]. Not only do such algorithms have the potential of increasing the accuracy of the artifact detection, they can also decrease the amount of time required for the calibration procedure.

5.5 Conclusion

The calibration procedure of the project used for determining the registration between ultrasound and da Vinci coordinate system has been improved by introducing a new ultrasound machine as well as an improved workflow. The new ultrasound machine was integrated into the calibration application through the DeckLink framework introduced prior in Chapter 4. The new calibration algorithm is capable of handling outliers and provides feedback to the user if the current calibration is adequate. The new procedure also provides tools for revising the registration if necessary. The new process was evaluated through a phantom study to quantify the total registration accuracy of the system. With the new algorithm, registration error has been reduced to an average of 3.19 $\pm$ 1.3 mm.

Chapter 6

Image Guidance Visualization Application

6.1 Introduction

The visualization application is the central focus of this project as it's responsible for creating the virtual prostate rendering scene which will be seen by the surgeon. This application has already been in development by other members of the project. However, the application was largely just a proof-of-concept prototype. In order for this application to be used in the operating room for a study, a well designed and robust application is required. A great deal of refactoring is also required in order to complete the software requirements laid out in Chapter 1.

6.1.1 System Requirements

The core requirements of the visualization application is as follows:

1. Render virtual prostate, ultrasound probe and da Vinci instruments in da

Vinci coordinate frame

2. Render scene from position and orientation of da Vinci endoscope cameras
3. Display MRI plane projected on prostate
4. Draw live ultrasound image beside virtual prostate scene
5. Augment ultrasound image with overlay of current prostate and tumour boundaries
6. Present visualization to surgeon by sending 3D stereo scene to da Vinci surgeon console

Of these requirements, numbers 1-3 have already been completed prior to the start of this thesis. The main design considerations of this chapter consists of developing the remaining requirements as well as improving the overall performance of the application.

One of the requirements stated is the introduction of augmented reality for the ultrasound image. The goal here is to provide the surgeon with an alternative method for visually determining the location of cancerous tumours. One of the main drawbacks of ultrasound is its inability for the viewer is differentiate a tumorous lesion from the surrounding tissue. If these ultrasound lesion artifacts could be highlighted, the surgeon will be able to recognize and track the location of the cancer. By using the transformation found by the calibration and MRI-TRUS registration applications, the prostate tetrahedral mesh could be used as a reference for the current prostate position. As an added benefit, the surgeon will be able to gain a perspective on the position of the prostate relative to the MRI as they can use the ultrasound image as a reference point.

6.1.2 System Overview

The visualization application is an MFC win32 application developed in C++. The details of the current system setup along with input and outputs each application is responsible for is elaborated in more detail in Figure 6.1. In order to communicate with other processes, the visualization application implements the master-slave library previously mentioned in Chapter 3. This allows the application to receive the instrument and camera position and orientation data from the Robot Control application along with the current motor angle from the Motor Control application. Once the virtual scene is generated, the resulting image frame is sent to the da Vinci TilePro system using the DeckLink framework developed in Chapter 4.

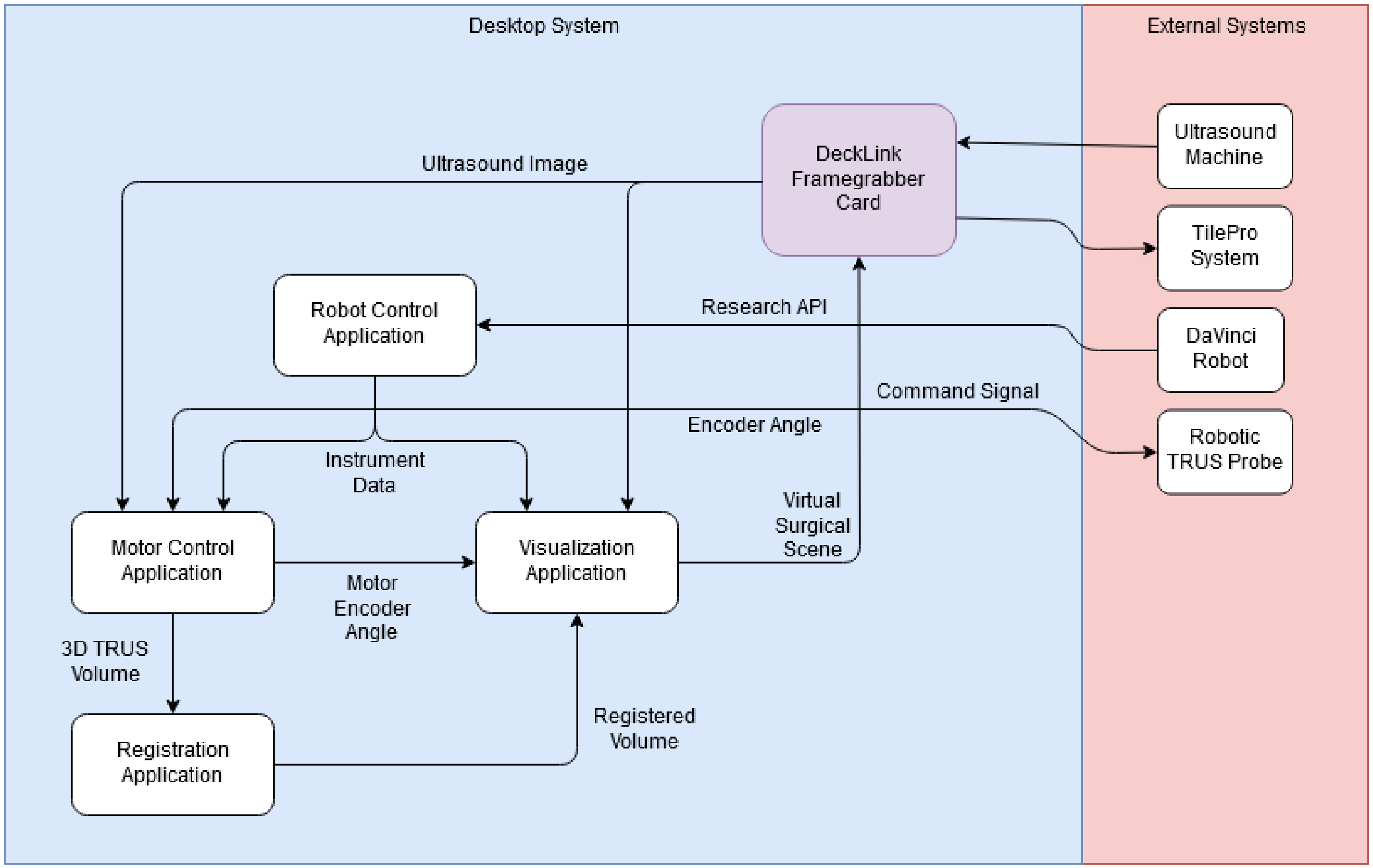


Figure 6.1: Block diagram of software layout with new components

6.2 Software Design

As the initial visualization application was written in MFC, the design of the application aims to follow the Document/View architecture paradigms as closely as possible. The document/view architecture separates data management into two classes, View objects which are responsible for displaying data and managing user interaction and Document objects which handle the data storage and management. For this project specifically, the system architecture is very simple as there is only a single document along with a single view object which can be seen in Figure 6.2.

One of the design considerations of this system is the need to handle data output and input at different frequencies. The data arriving from the robot API, motor control application, and ultrasound video stream all arrive at different rates and need to be handled separately. As a solution to this problem, the application is designed to function with multiple threads all running simultaneously. A diagram of these threads and their interactions can be seen in Figure 6.3.

The application contains six separate asynchronous operations used for performing the processing tasks necessary for the image guidance system. Of these six, four are used only for handling data input/output. These consists of two callback functions used for handling the motor and robot API data input and two threads used for handling the DeckLink framework input/output. The remaining two threads are used for the draw and update loops of the application. The main draw loop is responsible for rendering the virtual prostate scene. The update thread is used to call the slicer framework used for sampling the MRI image. Both of these threads run at 60 Hz which is more than sufficient for real-time performance.

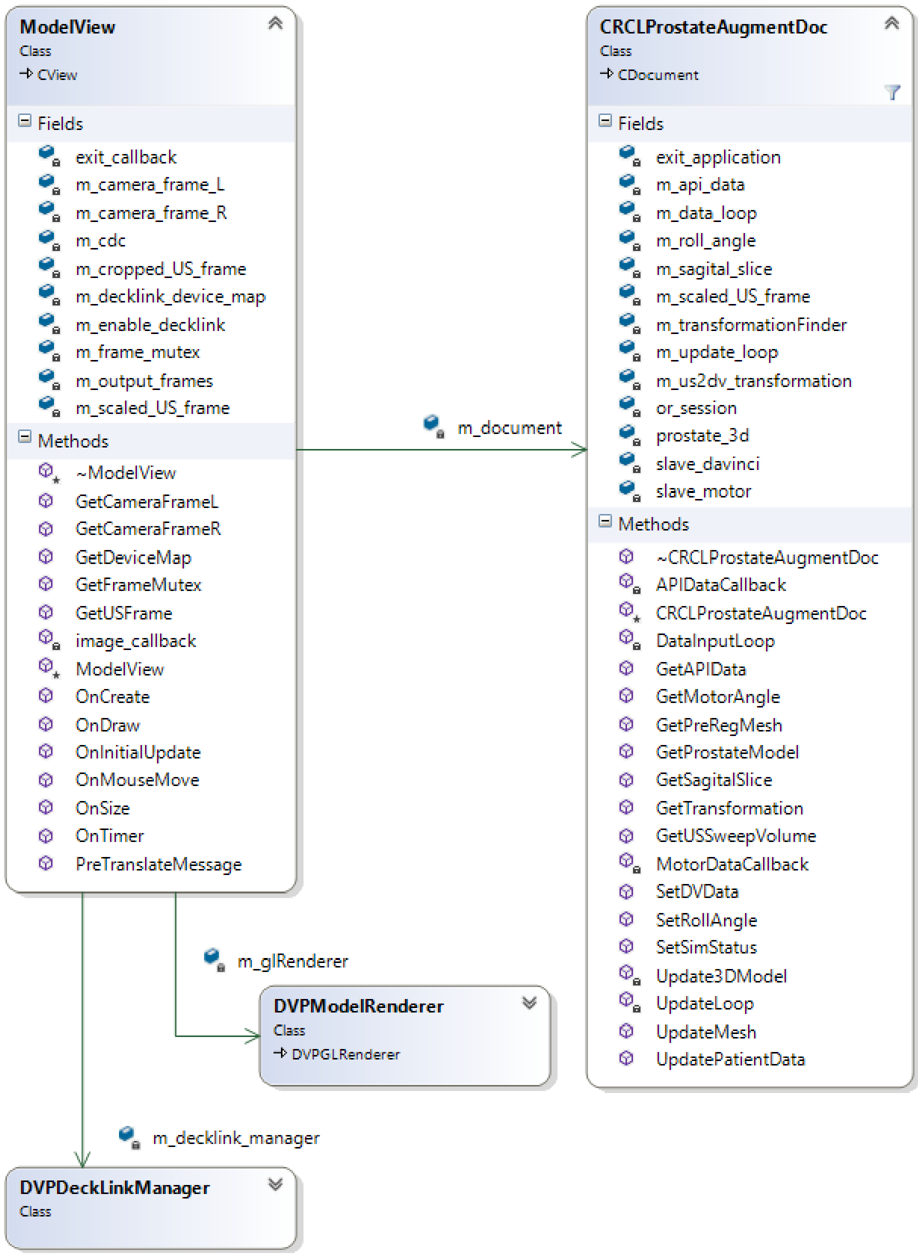


Figure 6.2: Simplified UML of visualization software

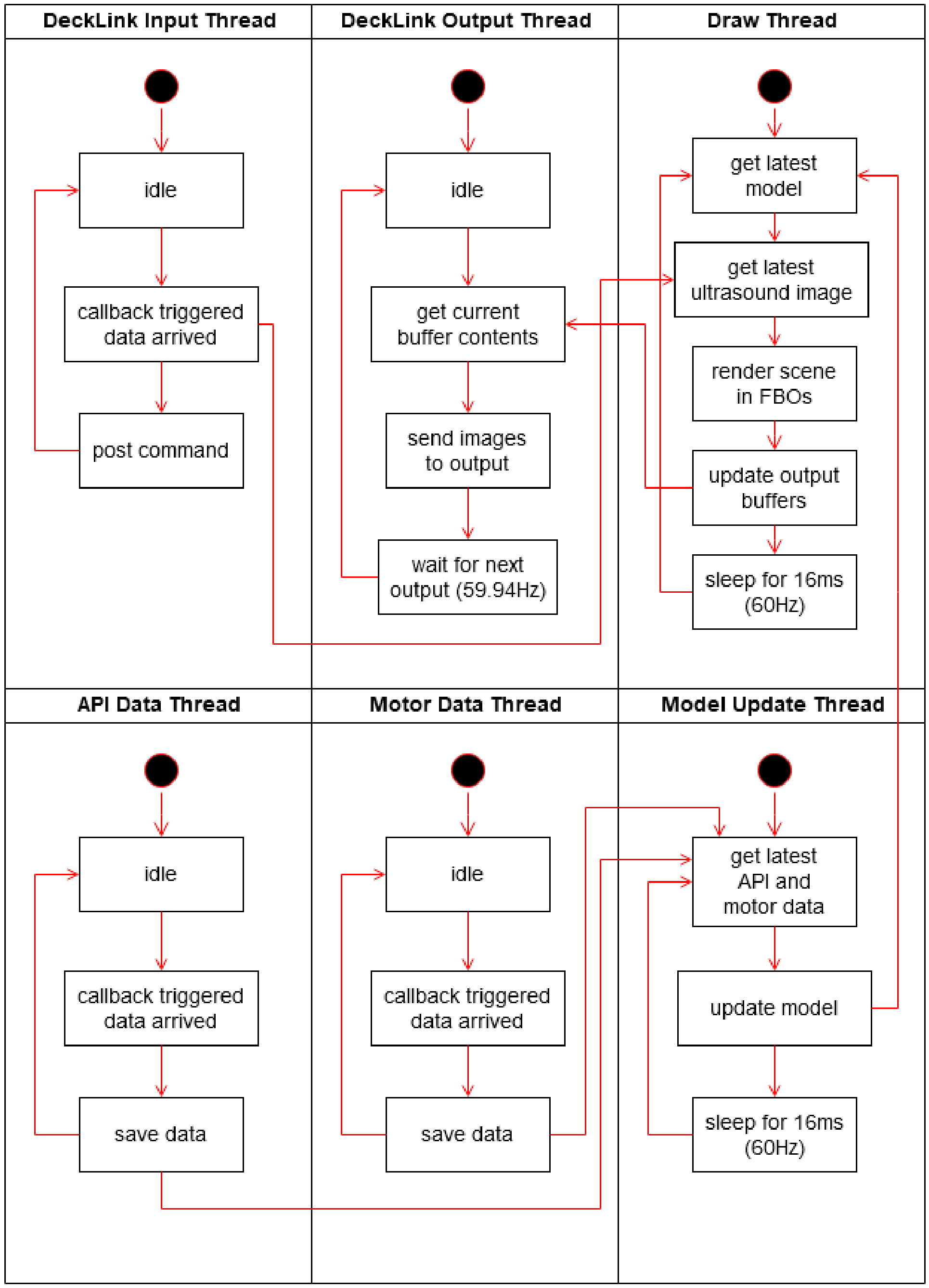


Figure 6.3: Asynchronous system functions and their interactions

6.2.1 Document Class

The `CRCLProstateAugmentDoc` is the only document in the current application and is responsible for most of the back-end tasks of the application. These tasks consists of the following:

1. Connect to the Robot Control Application
   (a) Obtain current camera position/orientation
   (b) Obtain current instrument position/
2. Connect to the Motor Control Application
   (a) Obtain TRUS robot rotational angle
3. Load tetrahedral prostate mesh
4. Update mesh based on registrations and transformations
5. Obtain slice of MRI volume based on robot instrument position

In order to perform these tasks simultaneously, a multi-threaded design was implemented. Specifically, tasks 1 and 2 are callback functions which will be called whenever new data arrives. Task 4 and 5 are performed by using the slicer framework described in Chapter 3 and is executed by a thread which uses the latest obtained API and motor data.

6.2.2 Model Class

`ModelView` is the main View object which handles the image rendering and user interactions. The layout of the GUI application is to be designed as shown in Figure 6.4, where the virtual prostate scene is drawn on the left and the live ultrasound video stream is shown on the right.

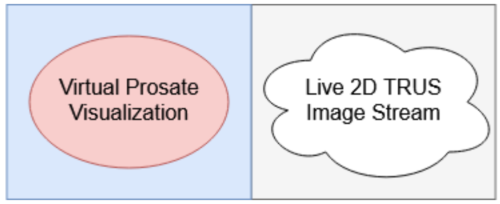


Figure 6.4: Layout of application UI elements

Stereo Output Rendering

While drawing both the virtual scene along with the ultrasound video stream is not a difficult task as most GUI libraries allow for multiple UI elements side by side in the same window, the main challenge lies in making this compatible with the TilePro system.

In the system setup prior to the start of this thesis, a video signal for the TilePro system was produced by connecting the surgeon console to the desktop video out port and setting the Windows system to cloning the main display. However, as a 3D stereo virtual scene is a requirement for the new visualization application, this previous method cannot be used. In order to produce a convincing 3D stereo effect, the Asymmetric Frustum Parallel Projection Planes algorithm explained in Chapter 4 needs to be implemented. This involves rendering the same virtual scene but from two different camera positions and with different projection matrices. Once rendered, the two video frames will be sent to the TilePro system where the 3D TilePro mode can be enabled through the surgeon console. The need for two channels also prevents the ultrasound image from being sent through the second TilePro port. Instead, the ultrasound image must be combined with the virtual prostate scene in a single image frame as shown in Figure 6.5.

Due to the need for a single rendering context, all the virtual scene rendering

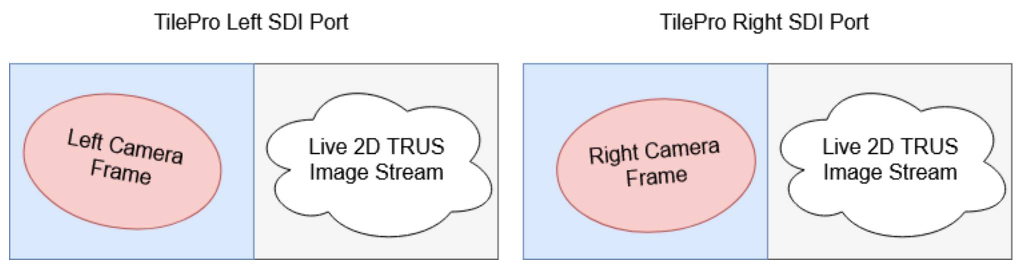


Figure 6.5: Stereo UI layout

needs to be completed in a single View object. This view object is instead responsible for creating all the objects required for rendering to the view along with rendering to the TilePro system. As seen in the UML from the previous Figure 6.2, the view object contains references to both the DeckLinkManager class and the `DVPModelRenderer`.

6.2.3 Framegrabber Integration

The `DeckLinkManager` class is an object which is responsible for performing all the DeckLink framework related tasks for this application. This class handles the initialization and destruction of `DeckLinkDevice` objects as well as handling the starting and stopping of the ultrasound video and output to TilePro. In this class, the port numbers used for input and output are specified in an Enum variable which can be easily modified if necessary.

6.2.4 Rendering Framework

As mentioned before, all the rendering for this application needs to be done in a single rendering context in order for the entire scene to be captured for the TilePro output. With this design consideration in mind, the rendering class of the current application was refactored and moved into a separate library in a series of classes.

This is done not only for implementing the features required for the application, but also to simplify the application and abstract away some of the more complex setup tasks. This also allows the same rendering functions to be reused by other applications such as the motor calibration application. The simplified UML of the rendering framework can be seen in Figure 6.6.

Base Abstract Renderer Object

The `DPVGLRender` is an abstract class which serves to abstract away most of the OpenGL initialization and setup processes. This class provides a simple interface where a series of public functions wrap protected functions which will be implemented by subclasses. This allows the openGL setup to all be performed by this class which greatly simplifies any subclass implementation.

Concrete Renderer Class Implementation

`DVPGLModelRenderer` is the concrete implementation of the `DPVGLRender` abstract class specific to the visualization application. This class contains two objects, `DVPProstateGLRender` and `DVPUSTextureRender` which are responsible for rendering the virtual prostate and ultrasound texture. This class does not perform the low-level OpenGL calls, but creates the objects responsible for that task. This draw process in performed on the main thread of the application and creates the fused scene which will be sent to the DeckLink library. By using the OpenGl function `glViewport`, specific regions of the screen can be specified for a draw process. This divides the screen for each object to perform their respective draw calls. This draw process is repeated three times, once for the current application, once for the left frame and once for the right frame. For the left and

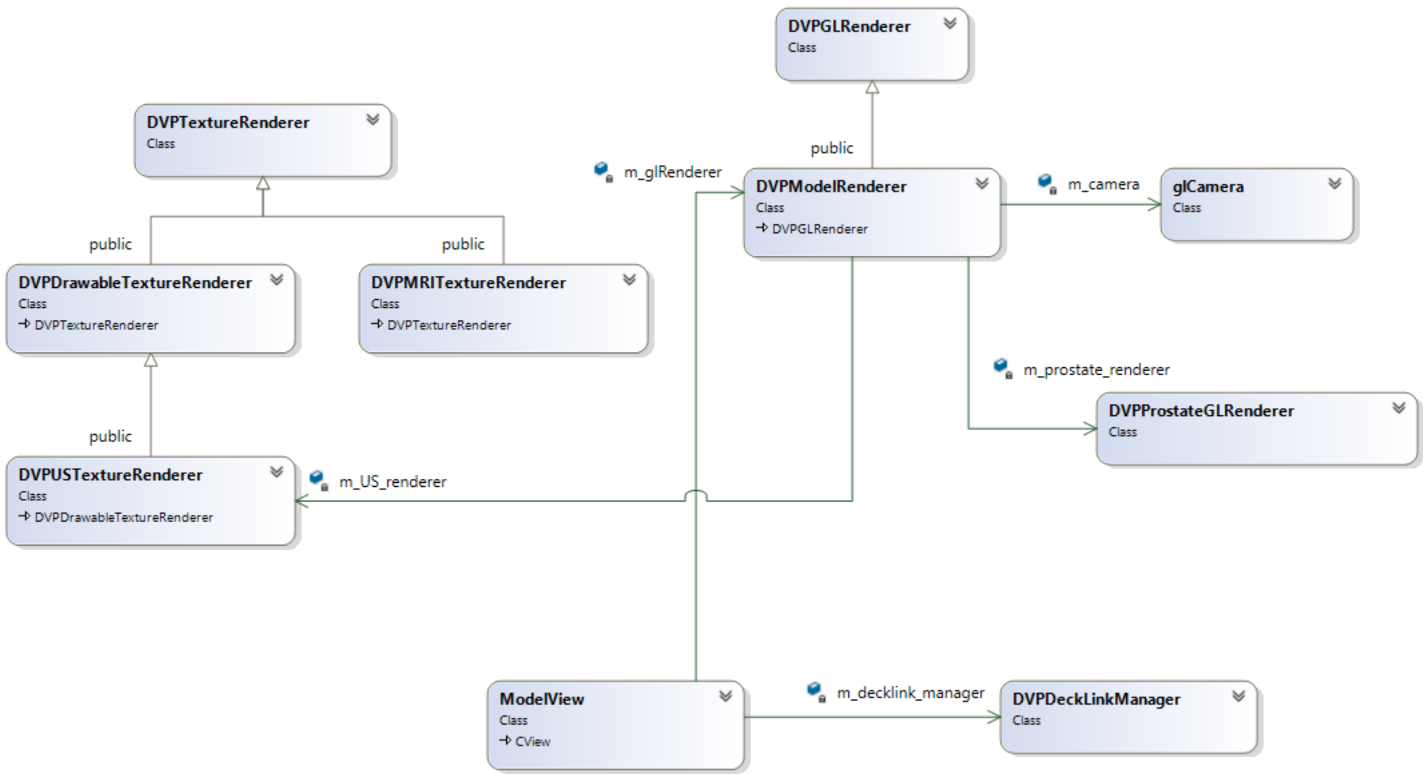


Figure 6.6: Simplified UML of rendering framework

right frames, the scene is drawn to a `FrameBufferObject` which is used for offscreen rendering. During each draw process, the `glCamera` class is used to specify the location of the camera for the `DVPProstateGLRender` object. This handles the camera placement necessary for the stereo effect.

Base Texture Renderer Object

The `DVPTextureRender` class is an object which handles the rendering of a texture when provided an OpenGL rendering context. This base class is designed to render any given texture. The object handles resizing the given texture to the proper size while preserving its aspect ratio.

Drawable Texture Renderer Object

The `DVPDrawableTextureRenderer` class is a subclass of `DVPTextureRender` which extends it the ability to draw overlays over the given texture. This feature is designed to be overloaded by subclasses for more specific implementations.

Ultrasound Texture Renderer Object

`DVPUSTextureRenderer` is a subclass of `DVPDrawableTextureRenderer` specifically designed for drawing the ultrasound images obtained from the DeckLink framework. The texture is as grayscale only to record the intensity values of the ultrasound image. As a subclass of `DVPDrawableTextureRenderer`, additional elements such as crosshairs which are used to indicate the current tooltip position are drawn on top of the ultrasound image.

Prostate Model Renderer Object

`DVPProstateGLRenderer` is the class which contains all the OpenGL function calls required for rendering the prostate mesh. Most of the prior work used for rendering the prostate model has been moved into this class for improved readability. The virtual prostate model can be seen in Figure 6.7.

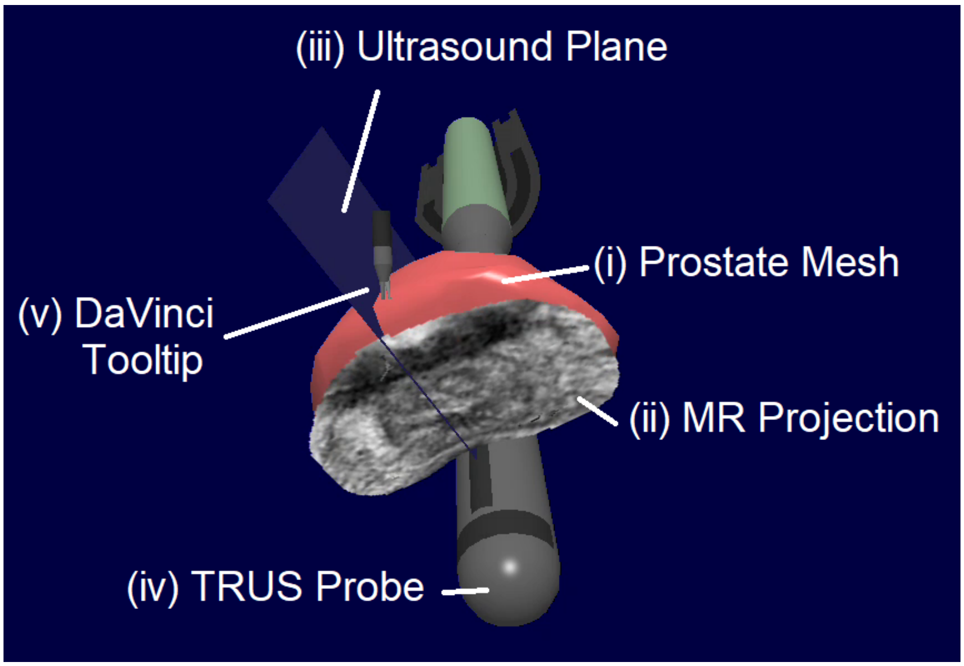


Figure 6.7: Labelled rendering of virtual prostate scene

6.3 Ultrasound Augmented Reality Overlay

Using the slicer frame adapted from Goksel et al. [10]'s work, given a tetrahedral mesh, a volume can be sampled in an arbitrary plane to obtain a deformed 2d slice sample of the volume. This framework is generic and could be used on any given volumetric data.

The process where the tetrahedral prostate mesh is generated from the MRI segmentation was modified to produce an additional volume. This volume is based on the same mesh but instead of containing the MRI voxel intensities, the volume stores only the boundaries of the prostate and tumour segmentations.

By sampling this mask along the sagittal direction using the mentioned slicer algorithm, a registered 2D image of the prostate boundary and tumour can be obtained. By applying the transformations determined in the registration steps, this sampled image can be transformed into the ultrasound coordinate frame.

6.3.1 Overlay Texture Renderer

The overlay drawing process was implemented in the `DVPMRITextureRenderer` object as a subclass of `DVPTextureRender`. This class renders prostate boundary as an alpha-blended texture on top of the ultrasound image. This allows both the ultrasound and boundary to be seen at the same time. The tumour boundary is drawn with a different colour to differentiate it from the prostate boundary. From Figure 6.8, the prostate region can be seen in blue while the tumours locations can be seen in red.

6.4 Robot Data Simulator

The robot data simulator serves as an offline alternative to the robot control application. The main purpose of the simulator is to generate the robot instrument position and orientation data stream based on previously recorded surgical studies. The da Vinci data simulator consists of two main components, a script for decoding recorded data and an application for sending the processed data to subscribing applications.

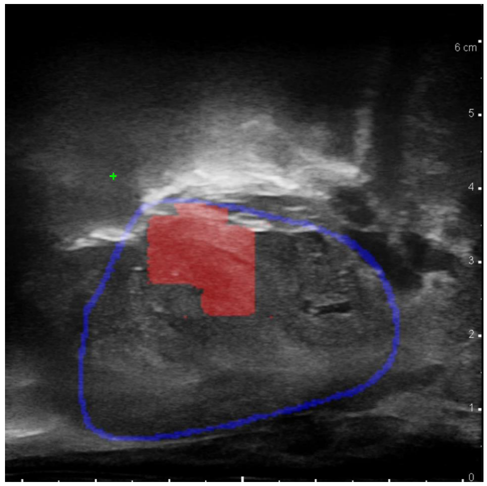


Figure 6.8: Overlay indicating tumour region

A script is used to decode the data recorded from the Epiphan Pearl system which provides multiple samples of the robot instrument data. This data is further processed by checking the timestamps to ensure the data is in the proper sequential order. If any part of the data is missing, it is replaced with null values. Once this processing is completed, the data is saved in a Comma-Separated Values (CSV) file.

In the current project workflow, there are multiple applications which need to pass data from one to another. To facilitate such an operation, a master / slave data transfer library was created. By using the windows sockets library or Winsock, ports between applications can be opened and data sent bidirectionally. The current library design only features unidirectional data passing and is used by multiple applications. The robot control master application serves as the main broadcaster of the robot data stream. One of the main limitations of the intuitive robot API is the restriction which prevents multiple applications from accessing the robot at once. In order to bypass this limitation, the robot control application is the only application which access the da Vinci robot and broadcasts the data to all the other applications which requests for this data. By replacing this control application and sending data with the same port number, it is possible to simulate the data stream of the robot to every application used by this project.

The main simulation application functions through a master slave paradigm and serves as an alternative data broadcaster which replaces the control application. By loading one of the processed decoded audio files, the entire history of a robot's instrument data can be replicated. This GUI application provides an interface for manipulating a previously extracted audio data file. Not only can the data be played back, the data can be modified to test the software behaviour. This allows for the free control of the endoscope camera position and orientation as well as the instrument tool tip position and orientation.

6.5 Latency Test

The RCLAugmentApp is a multi-threaded application designed for producing the reconstructed 3D surgical scene and sending it to the da Vinci console TilePro system. The application has two primary input sources:

1. The da Vinci Research API stream is established with the da Vinci robot through a standard ethernet cable using a Transmission Control Protocol (TCP)/Internet Protocol (IP) connection.

2. The BlackMagic DeckLink Quad 2 is a frame grabber card used for both input and output. As input, the DeckLink Quad 2 captures frames from the bk3500 ultrasound machine at a resolution of 1920×1080. The signal is

obtained by converting the ultrasound source using a BlackMagic HDMI to SDI mini converter.

Aside from the inputs, the application produces two output signals for the da Vinci TilePro console. These outputs consists of two channels for a stereo 3D simulated surgical scene and are sent to the TilePro system using the DeckLink Quad 2 card through SDI cables. This output system has handled separate in a separate thread by the DeckLink Quad API. In order to produce the 3D surgical scene for the TilePro console, the application uses two threads:

1. The algorithm update thread is responsible for updating the mesh clipping plane based on the current da Vinci tooltip coordinates obtained from the research API.

2. The draw thread is responsible for rendering the surgical scene in 3D stereo. After rendering the surgical scene, it fuses it with the ultrasound image to form a single image.

To evaluate the latency of the DeckLink Quad 2 system, a stopwatch was started on a test computer. This display signal was then captured by the DeckLink Quad 2 card and sent to the da Vinci Console. A camera was then used to capture all the displays in order to obtain the time-stamp on each screen. The results can be seen in Table 6.1. To compute the total latency of this multi-threaded system, the worse case scenario is considered where the individual latency of each separate thread is added up. Since the ultrasound input is not involved in the main processing loop, its latency is not added to the scene rendering computation time. The US signal has a total latency of 219.7 ms. The total latency time for the rendered scene in the da Vinci console was 144.4 ms and consisted of the latency of the da Vinci API, the scene rendering time, deformed MRI rendering and the TilePro output.

Table 6.1: Transmission latency and computation time for signals and connections

Signal / Connection	Computation Time / Latency
da Vinci API Data Stream (60Hz)	2.0 ± 1.67 ms
3D Scene Rendering (60Hz)	15.1 ± 2.0 ms
Deformed MRI Rendering (60Hz)	2.6 ± 0.29 ms
Ultrasound Input (59.94Hz)	75.0 ± 14.8 ms
Output to da Vinci TilePro (59.94Hz)	125.0 ± 17.6 ms
Total Latency of Rendered Scene	144.7 ms
Total Latency of Displayed US Image	219.7 ms

6.5.1 Discussion

From the results of the latency test, it can be seen that the latency quite high. From Table 6.1, it can be seen that the most significant impact to performance is caused by the framegrabber retrieval and output of video frames rather than the processing of the application itself. Comparing the results of our experiment to the work of Kibsgaard and Kraus [17] who performed a similar experiment using the same framegrabber card, it can be seen that they've also obtained similar results with an average latency of 237.7 ms.

While this latency is not insignificant, it is lower than the limit of acceptable latency of 300 ms from Marescaux et al. [28]. This latency is also lower than the average limit for human reaction time of 250 ms [28]. Another consideration would be the relative latency between endoscope video and AR overlay would be much lower than absolute latency. As the endoscope video will also be read by the DeckLink framework and sent to the TilePro system, the effective latency between camera and AR could be as low as 19.7 ms.

While much of this latency may simply be a technological limitation of current generation framegrabber systems, there are still potential improvements which can be made. The DeckLink Quad 2 is capable of a feature known as keying in which input data is processed while simultaneously overlaying it with computer graphics. This feature could potentially help reduce latency by improving the delay between input and output.

6.6 Conclusion

A visualization application based on the requirements set out in Chapter 3 was developed building upon the work by other members of the project. This includes integrating the imaging pipeline framework introduced in Chapter 4 for displaying the real-time ultrasound video as part of the virtual surgical scene. As a key contribution of this project, a 3D stereo effect of the virtual scene was produced by rendering the scene from two different perspectives and sending the resulting frames to the TilePro system through the imaging pipeline framework. An augmented reality system for the ultrasound image was introduced in the form of an prostate and tumour boundary overlay. In order to test and debug the system, a robot simulator was developed to playback the data from previous surgical cases. With all these features complete, a latency test was performed with the functional latency of the system averaging around 144.7 ms and total latency with ultrasound at 219.7 ms. This delay while not insignificant, remains within the acceptable limits of latency for augmented reality systems.

Chapter 7

Real-time Intraoperative Prostate Registration

7.1 Introduction

Augmented reality for a laparoscopic MIS environment is a difficult problem with two significant challenges. The first major challenge is the need for an accurate registration between the surgical environment and the imaging modality. The second challenge is the need for maintaining the registration among the dynamics of the surgical environment.

In the current state of the project, the first major challenge of determining the registration between the da Vinci robot and MRI imaging modality has already been solved through combining the transformations obtained during the registration and calibration process. While this current method has been shown to produce an accurate initial registration, it lacks a system for accounting for surgical dynamics. As such, whenever the prostate is moved or deformed from any surgical movement, the registration will become inaccurate as there is no current way of maintaining the registration during such movement. In order for the image guidance system to be practical, a method for accounting for surgical dynamics or tracking the prostate motion and deformation is required.

7.2 Summary of Tracking and Registration Techniques

For the current problem, the goal is a method for maintaining a registration over the course of a surgery. The first category of techniques considered are known as tracking techniques in which the goal is to track the motion and deformation of the tissue of interest. Tracking can either be done either through the use of natural or artificial landmarks in the surgical scene. Alternatively, registration techniques which could be performed periodically or in real-time may also be considered. While registration techniques differ from tracking techniques, they still remain many similarities between the two problems. Registration can be largely divided into interactive, point-based, surface-based and volume-based techniques.

7.2.1 Tracking Techniques

Natural landmarks are features which could be normally found in a laparoscopic scene. These features often have distinct edges or corners or rich textures distinct from their surroundings.

Artificial landmarks can be introduced in the laparoscopic scene as fiducials or other markers which can be easily detected. These fiducials vary in form from colour needles to ElectroMagnetic (EM) transponders placed within the tissue.

7.2.2 Registration Techniques

Interactive techniques rely on manual user input such as from an expert to perform the registration during a surgery. The current MRI to TRUS registration algorithm serves as an example of such a technique where a resident will perform the registration manually during the surgery. While this technique is highly accurate, the need for a user input makes extending the algorithm to real-time application is highly impractical if not impossible.

Point-based techniques are similar to tracking techniques in which specific points in the surgical scene are chosen as landmarks and used for alignment. These landmarks can either be natural anatomical landmarks located on both medical imaging and in the surgical scene or artificially introduced in the form of fiducials.

Surface-based techniques aim to estimate the surgical scene through surface reconstruction and can be largely divided into active and passive techniques. Active techniques use additional hardware devices aside from endoscopic data such as structured light or ToF to obtain depth from the scene.

Volume-based techniques use intraoperative 3D imaging aside from visual cues from endoscopic images to perform the registration. These imaging sources may range from open MRI scanners to real-time ultrasound transducers.

7.2.3 Real-Time Registration

From the list of possible options, it can be seen that many are not without their drawbacks which renders them unfeasible for the current application. Some solutions, while promising, require additional hardware components such as InfraRed (IR) emitters or ToF cameras where regulatory and ethics approval will be difficult. Vision based approaches while promising, lack robustness and will require significant work in order to account for surgical dynamics. A volume based approach using a real-time implementation of a FEM-based 3D to 2D TRUS registration algorithm was determined to be the most feasible solution. Instead of tracking the prostate, a real-time registration algorithm will perform a similar function and account for prostate motion and deformation. This algorithm was developed by Dr. Samei, a post-doctorate researcher who has been working on this project. Her algorithm is explained in more detail in Samei et al. [42].

7.3 Algorithm

Prior to running the algorithm, a registration must already have been completed, as described in Chapter 5 and Chapter 8. A MRI volume is required for generating the corresponding FEM mesh along with a TRUS volume for the initial registration. The real time registration method is an iterative algorithm which solves for a new transformation given its initial registration TRUS volume along with a new current TRUS image and angle. A quick summary of the algorithm can be broken down to the steps:

1. Calculate node coordinates of mesh with transformation applied
2. Solve for new transformation parameters by iterating through all ultrasound planes in buffer
    (a) Determine deformation model of TRUS volume at current angle
    (b) Solve for deformed TRUS volume slice
    (c) Calculate sum of squared differences between current ultrasound and deformed slices
    (d) Calculate derivative / hessian

3. Repeat n iterations until convergence

The registration involves finding the transformation which aligns a previously obtained reference TRUS volume (I_r) to a set of currently obtained 2D TRUS slices (I_s). While it is possible to deform the reference frame then sample it to obtain the pixel intensity values for comparison, it will be computationally very expensive. Instead, the pixel positions corresponding to the pixels of Is are determined instead. Once obtained, the two images can be compared using a Sum of Squared Differences (SSD) as a cost function for alignment matching. This optimization problem is an unconstrained nonlinear multi-variable minimization. In order to perform the optimization, a trust region optimizer is used to minimize the SSD cost function. The first and second order derivatives are pre-calculated with higher order terms ignored.

This real-time registration algorithm is designed to have two components, a non-rigid and rigid registration. For the purposes of this thesis, only the rigid registration will be implemented. The focus of this chapter is on the implementation of the framework and integration of the algorithm into the visualization application. Once that is completed, the addition of the real-time rigid registration algorithm could be added to the system.

7.4 Software Design

7.4.1 Real-Time Registration Library

In order to integrate the algorithm into the main visualization system, a separate library was created which handles all registration specific functions. To perform the actual optimization, the third party optimization toolkit dlib [19] was used. By providing the toolkit with a specific class object, a trust region minimization can be performed. The library was designed to be operated asynchronously from the main visualization application which allows for both the MRI slicing algorithm and real-time registration algorithms to be run concurrently. This also has the added benefit of allowing for multiple instances of the registration algorithm to be used. A UML diagram of the library can be seen in Figure 7.1

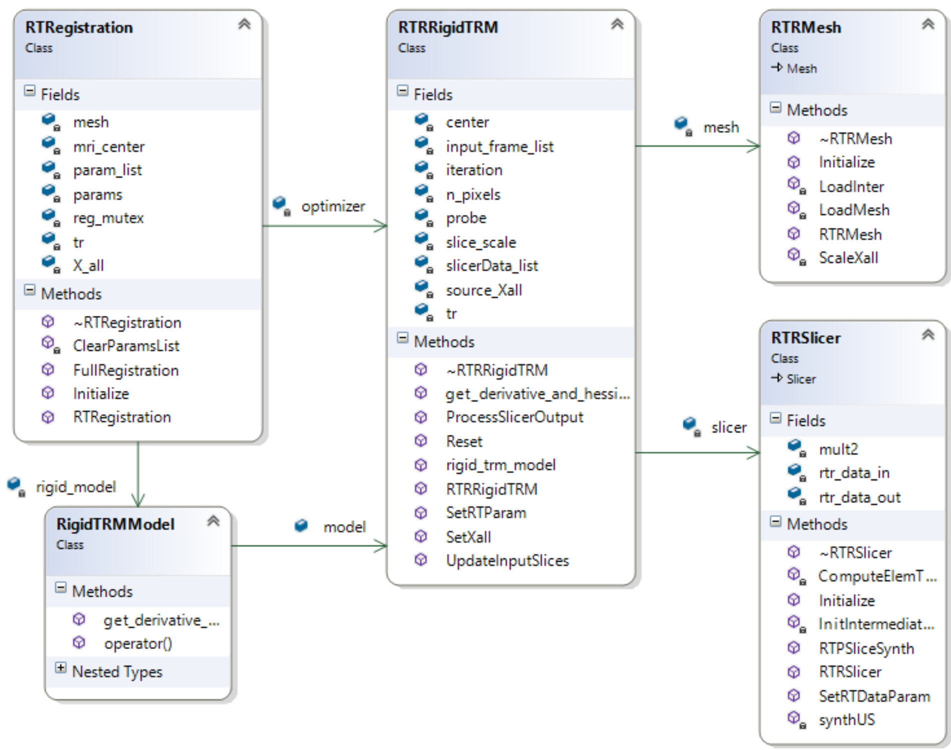


Figure 7.1: UML diagram of real-time registration library

Registration Class

`RTRegistration` is the main class responsible for the initialization, start and stop of the real-time registration. This class is initialized by providing it with the filepaths of the tetrahedral prostate mesh generated after the registration process. The class creates the optimizer object along the optimizer model object required by the dlib library. The real-time registration is started by providing the class with a list of ultrasound images and their respective angles. The class will pass these parameters to the model and begin the optimization. When the optimization is complete, the class returns the transformed mesh as a list of points.

Function Model Object

`RigidTRMModel` is a "function model" which is used in the `find_min_trust_region` routine of the dlib library. In particular, this object represents a function and its associated derivative and hessian. This object contains a reference to `RTRRigidTRM` and calls its function implementations.

Rigid Trust Region Minimizer Object

`RTRRigidTRM` is an instance object which contains the data and references to objects required to performed the optimization. This object keeps track of the slicer and mesh objects responsible for the sampling of the ultrasound images. Most significantly, this class contains the concrete implementations of the `rigid_trm_model` and `get_derivative_and_hessian` functions which are used by the `RigidTRMModel` when called by the optimizer.

7.4.2 Visualization Application Integration

The real-time registration library is integrated into the main visualization application through a single class, `DVPRegistration`, which handles the registration logic. The class structure can be seen in Figure 7.1 where the class is created by the Document object. Once created, the class begins a thread which, when the registration is enabled, begins a loop which constantly calls the optimizer by passing it the current ultrasound image and motor angle.

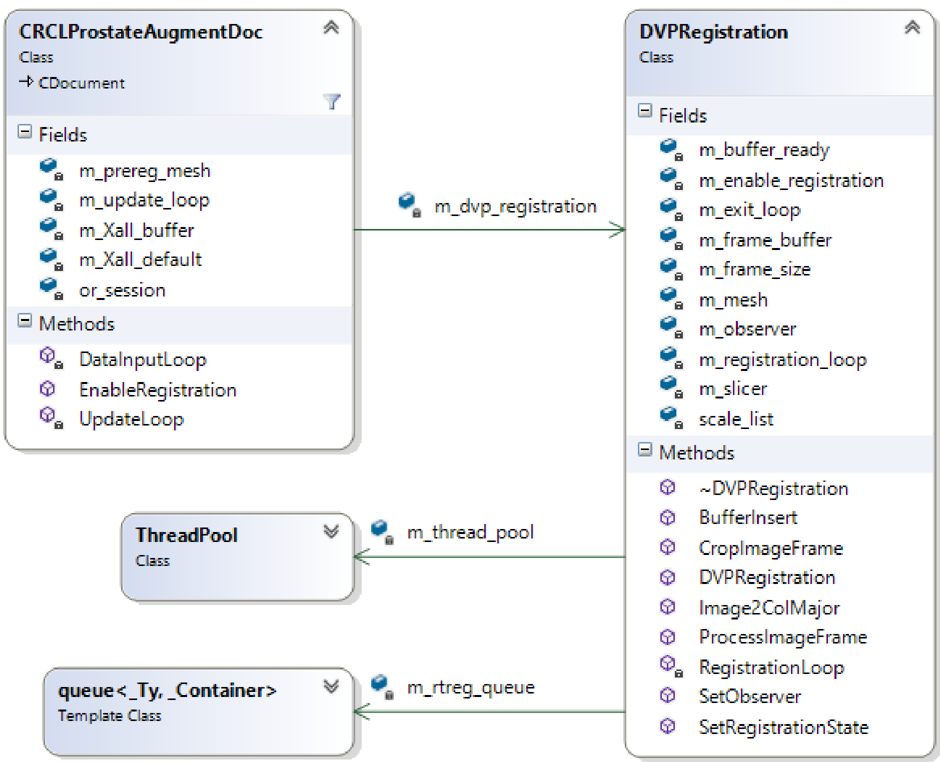


Figure 7.2: Simplified UML diagram of real-time registration library integration into visualization application

In order to test the registration algorithm, the da Vinci robot data simulator developed in Chapter 6 was extended with additional functionality. The simulator was modified to send the current timestamp and frame number associated with the current da Vinci instrument coordinate data to all subscribing applications. This allows the recorded ultrasound video file to be correctly synchronized with the robot data. The ultrasound video file is loaded in the visualization application and displays the frame specified by the simulator. If necessary, a frame offset is introduced to compensate for the delays introduced during the data recording process. Video frames obtained from the video file are then added to the frame buffer which would be used by the registration algorithm.

The main error metric for evaluating the performance of the registration algorithm is registration accuracy. The runtime of the algorithm is also of concern as it could be correlated to the accuracy of the registration. This runtime metric can be largely divided into latency and time between registrations or update frequency. Latency serves to measure the delay between when a new ultrasound image is obtained and the registration model is updated. The longer the delay, the greater the discrepancy between current ultrasound image and model. The update frequency is also a contributor to accuracy as each subsequent registration is based on the previous one. As the frequency between model updates decreases, the chances of the algorithm failing increases.

During the implementation of the registration algorithm, a problem was quickly noticed on how the model update rate affects the registration robustness. As each model update step relies on the results of the previous step, a bigger delay between update results in more significant differences between each update. These differences are particularly problematic when there is significant motion in the ultrasound image between updates. This problem is particularly apparent during specific stages of the surgical procedure, where the surgeon makes rapid movements. This rapid instrument motion directly translates to the angular velocity of the ultrasound probe as its configured to track the instrument position. These rapid motions often cause the registration algorithm to fail as the algorithm does not have the update frequency to match the sampling rate needed to properly track the ultrasound motion. In order to alleviate these problems, two solutions were introduced in the form of a low pass filter and the introduction of a thread pool.

In order to compensate for these problems, a low pass filter was introduced in the form of a moving average applied to the prostate mesh coordinates. This filter serves the dual purpose of smoothing the abrupt model motion whenever the algorithm updates the model as well as help smooth any irregularities which may arise when the surgeon moves the instruments too quickly.

A thread pool is a software design pattern used for the parallel execution of a task. The thread pool maintains a list of threads which can be used to perform a task specified by the program. As shown in Figure 7.2, the `DVPRegistration` object contains a queue of `RTRegistration` objects along with a `ThreadPool` object. The threads are pre-allocated along with the registration object which reduces the latency of the task. No resources are required for the creation and deallocation of threads. The number of threads can also be configured to be specific to the resources available to the application which may vary depending on the number of parallel processors, cores or memory available. For the current application specifically, the thread pool allows for multiple registration optimizations to be performed simultaneously. Each thread task will start when sufficient new frames are provided. This will allow for the maximum amount of overlap between threads and thus reduce the amount of time between model updates as much as possible.

7.5 Results

7.5.1 Experimental Setup

There are a few specific parameters which significantly impact the performance of the algorithm which may translate in improvements to accuracy. These parameters include image scale, buffer size and thread pool size. Image scale is a parameter which determines the scaling factor applied to the ultrasound image. While a smaller image improves the performance of the system, it could also remove features in the image necessary for a proper convergence. Buffer size is a parameter which determines the number of frames collected for the registration algorithm. A larger buffer increases system stability and accuracy but significantly impacts performance by increasing overall runtime of the algorithm. The last parameter is thread pool size which largely only affects algorithm update frequency. While too many threads may have an adverse effect on overall system performance, increasing the number of threads greatly reduces the delay between each model update.

From an analysis of the individual parameters, it could be determined that the image scaling factor is inversely proportional to the runtime of the registration algorithm. The relationship is linear with each decrease in scaling ratio increasing performance by 23 ms. An analysis of buffer size shows an even more significant effect on system performance. Buffer size is directly proportional to runtime increasing runtime by approximately 50ms with each buffer element added. From these results, it was determined that a scaling factor of 3 provides the best accuracy-to-runtime ratio. Too small of an image results in the algorithm being unable to function well; too big of an image requires too much time to process the images. As for buffer size, it was determined that a size of 3 provides the best stability to performance ratio. With too few frames causing the algorithm to fail more frequency and with too many frames requiring too much time to process the data.

Evaluating the accuracy of updating registration on a moving prostate registered model is a difficult process as there is no baseline available. Initially it was hoped that ultrasound phantoms could be used to test the accuracy of simple translation and rotational movements. Unfortunately, the procedure of testing with ultrasound phantoms will not work for this application as phantoms lack the image distinction needed for the algorithm to perform adequately. Past surgical cases would be simulated and reconstructed to test the registration algorithm. As mentioned prior in Chapter 3, a recording of the da Vinci instrument data along with ultrasound images has been created for multiple studies. Using this simulation, two separate image similarity metrics were used to determine the overall accuracy of the system.

Dice Similarity Coefficient

The Sørensen-Dice Similarity Coefficient (DSC) is commonly used metric used to determine the similarity between two image segmentations. The discrete form of the DSC can be seen in Equation 7.1.

$$DSC(A,B) = \frac{2|A \cap B|}{|A| + |B|} \tag{7.1}$$

In order to obtain the binary mask required for this metric, a binary mask of the current prostate model is obtained by sampling the mask volume with the transformation obtained by the registration algorithm. This provides a single image slice obtained from the same angle as the current ultrasound image. This current ultrasound frame is manually segmented to obtain the second binary mask. The DSC metric between the two given masks was then calculated using a separate script in order in order to not affect the run-time performance of the registration algorithm. To compare the results of the experiment, a control case where the registration algorithm is not used was also sampled with the DSC calculated. This was repeated with 200 frames collected sampled at 30Hz.

Normalized Cross Coefficient

The Normalized Cross Correlation (NCC) is a metric used to measure the similarity two images. Unlike DSC which is primarily designed for discrete data, NCC is used to compare the similarity of image intensities and can be seen in Equation 7.2.

$$NCC(A,B) = \frac{\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(A(i,j)-\bar{A}\right)\left(B(i,j)-\bar{B}\right)}{\sqrt{\left[\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(A(i,j)-\bar{A}\right)^2\right]\left[\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(B(i,j)-\bar{B}\right)^2\right]}} \tag{7.2}$$

Where $A(i,j)$ and $B(i,j)$ are the reference images of size $n \times m$ and $\bar{A}$ and $\bar{B}$ are the mean intensities of A and B respectively. In order to obtain the reference images necessary for the NCC, the application was modified to record the current ultrasound video frame along with the sampled ultrasound image slice. The NCC metric was calculated from the recorded image files using a separate script in order to affect the performance of the registration algorithm. The data collected consists of approximately 1600 frames collected while the registration algorithm was running.

7.5.2 Experimental Results

The results of the DSC values over the course of the experiment can be seen in Figure 7.4. From a quick glance, it can be seen that in general, the experimental results show a greater similarity score over the control case.

The results of the NCC results can be seen in Figure 7.5. A general trend can be observed where over time, the correlation between the two sets of images begin to decrease steadily.

7.5.3 Discussion

From the DSC experiment, it can seen that overall, the registration algorithm performs better than the baseline control case. The first highlighted region shows an area where the experimental case performs much better than the control case. This region of the data corresponds to a significant movement of the prostate. As the registration algorithm was able to adjust the model accordingly, it was able to match up with the segmented results much better than the static control case. In the second region, a case where the static control case performed better than the experimental case can be seen. In this region of the results, the surgeon has lifted the prostate to operate on a region beneath it as seen in Figure 7.3. This sudden motion likely resulted in the registration algorithm unable to compensate for the deformation. By the end of the region, the registration has likely determined a better fit and accounted for the rotation in the image.

From the results of the NCC experiment, a general downwards correlation trend can be seen. This is likely caused by the increasing amount of movement and de-

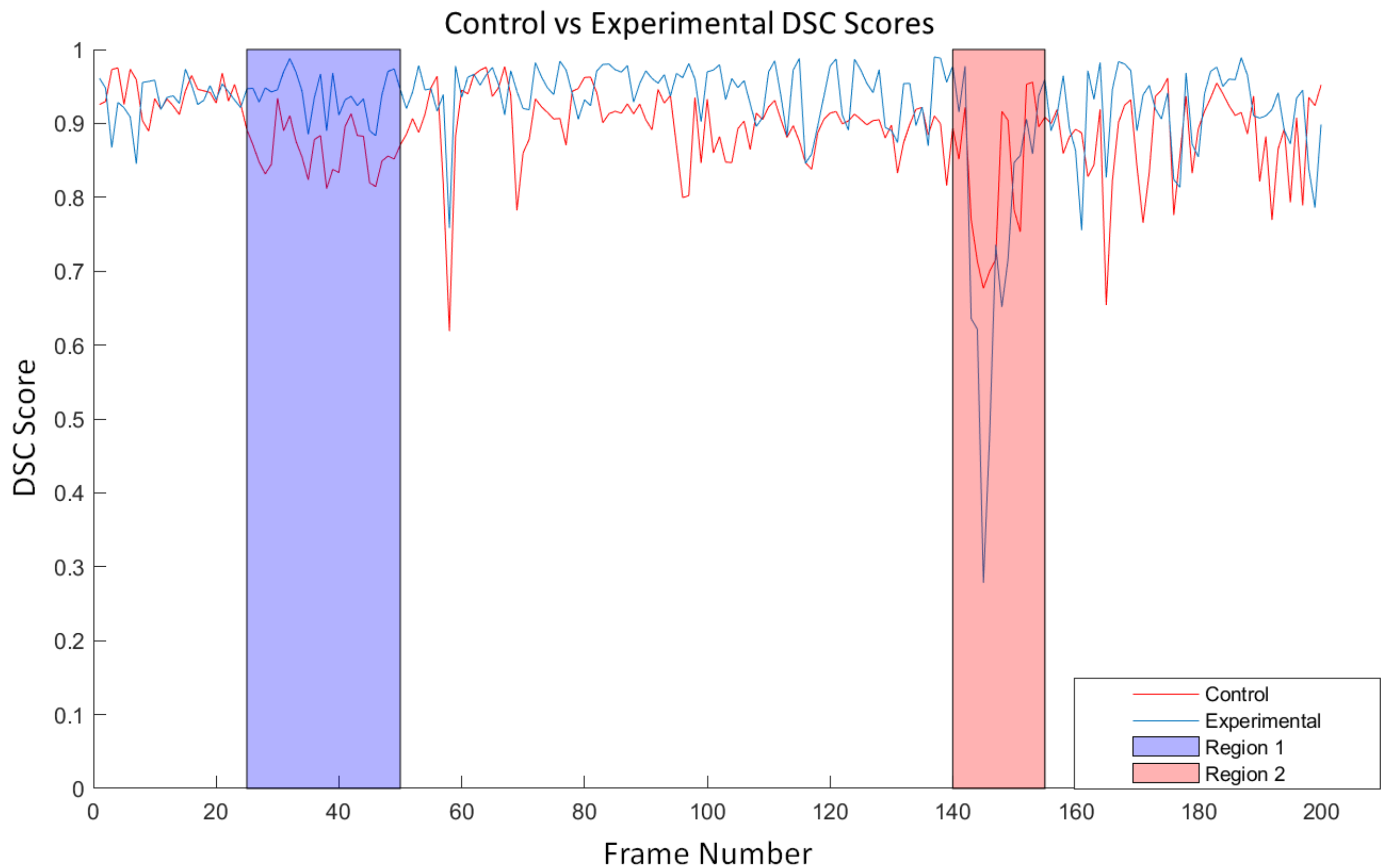


Figure 7.4: Experimental results using Sørensen-Dice Similarity Coefficient

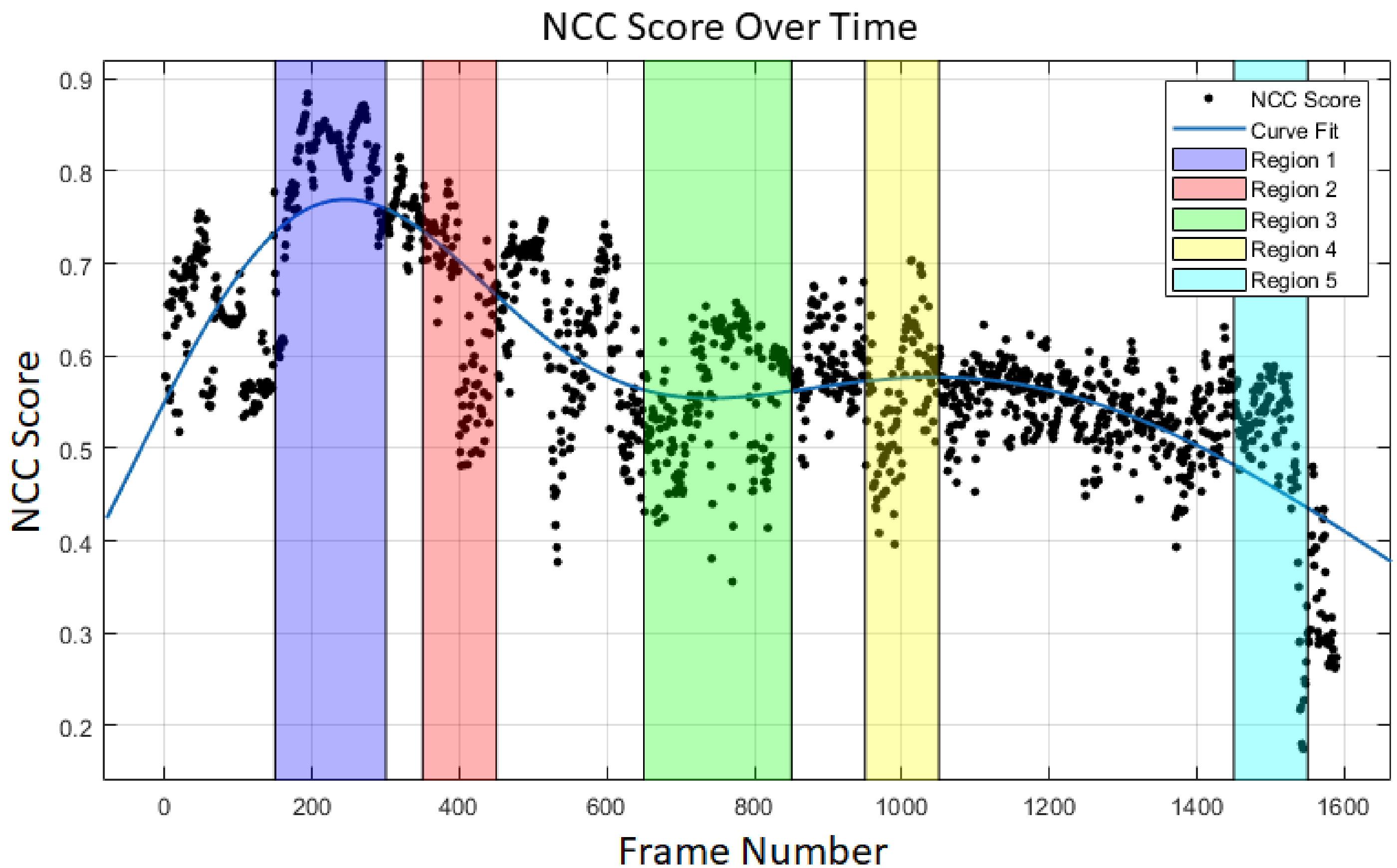


Figure 7.5: Experimental results using Normalized Cross Correlation

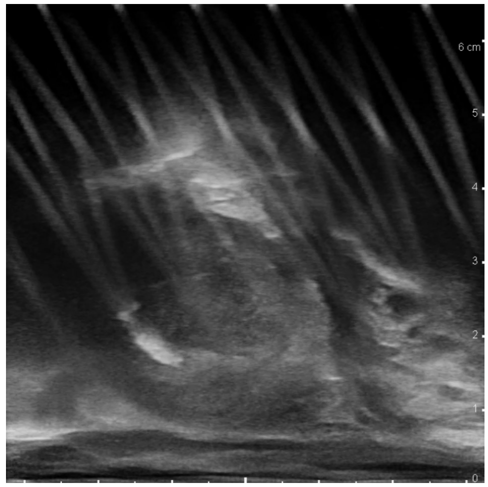


Figure 7.3: Significant deformation in prostate from surgeon motions

formation of the prostate causing it to differ from the initial reference volume. This is also one of the potential causes on the degrading performance of the algorithm the longer it runs. Another observation of note is the trend of high correlation followed by low correlation scores. One possible reason for this behaviour could be the latency in the registration algorithm. Whenever there is a significant deformation and movement of the prostate, some time is required for the algorithm to process the change and adjust the model accordingly. In the figure, a few particular regions of interest can be seen.

In the first region, it can be seen that the correlation score is very high. Viewing the ultrasound images in Figure 7.6, it can be seen that the algorithm is currently capable of compensating for a small rotation in the model.

In the second region, a drop in correlation is quickly followed up with a recovery. Examining the figures seen in Figure 7.7, a noise pattern can be seen in the reference image. This noise artifact is caused when the surgeon uses the cauter-

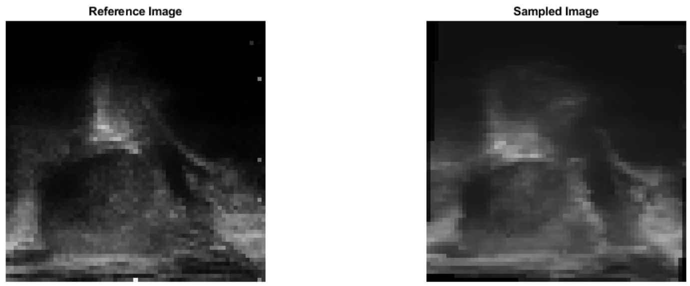


Figure 7.6: Strong correlation between reference and sample

izer tool which likely produces noise which are then picked up by the transducer. This cauterizer noise artifact may be one cause of the decreased similarity score between the two images as it introduces a pattern which cannot be matched in the sampled image. Along with this noise pattern, it can be seen that the algorithm has yet to account for the translation in the prostate. By the end of the region, with the figures seen in Figure 7.8, it can be seen that the algorithm has largely determined the correct transformation and the correlation score increased again.

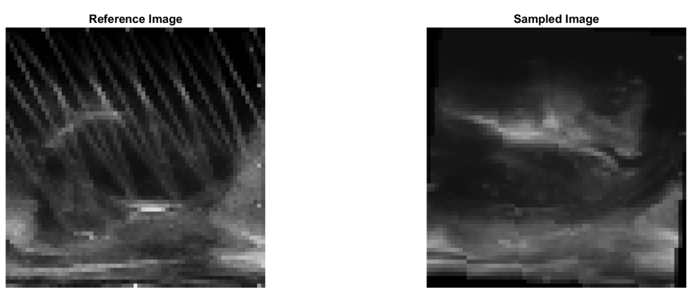


Figure 7.7: Cauterizer noise pattern in reference image

In region three, the surgeon has begun operating in the posterior region of the prostate and has lifted it up thus introducing a large translation and rotation to the image as seen in Figure 7.9. The registration algorithm appears to be initially

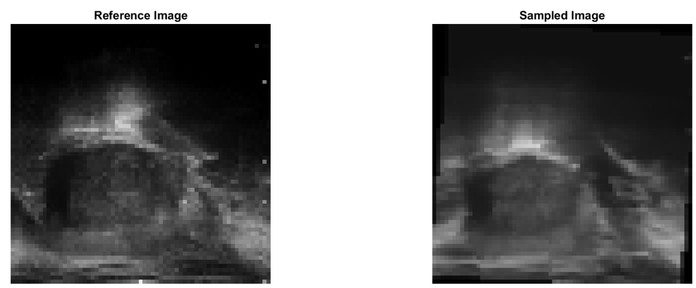


Figure 7.8: Registration recovery after cauterizer noise

unable to account for the deformation of the scene, but eventually manages to compensate. In Figure 7.10, it can be seen that the algorithm eventually determines a transformation capable of accounting for the rotation.

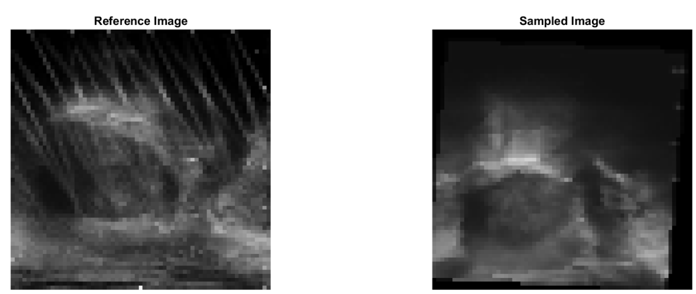


Figure 7.9: Rotation introduced in reference

In region four, another observable sample can be seen in Figure 7.11. In this sample, the registration algorithm has compensated for the prostate deformation by introducing a significant amount of rotation. This rotation has also introduced a significant amount of empty pixels into the image which can be seen as black pixels surrounding the sampled image. Not only will this negatively impact the NCC metric, but could potentially be a source of error in the registration algorithm.

In the last region seen in Figure 7.12, it can be seen that the algorithm is start-

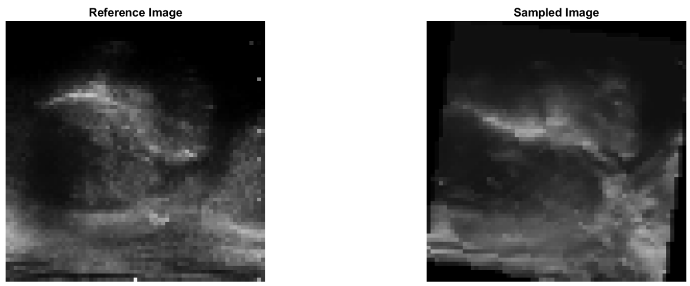


Figure 7.10: Registration algorithm account for rotation

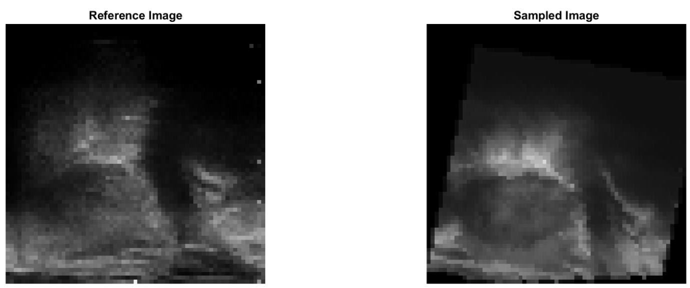


Figure 7.11: Gaps in image introduced from large rotation

ing to have difficulty compensating for both the deformation and rotation in the reference image. During this part of the surgery, the surgeon has likely lifted the prostate again for another task. Unfortunately, this deformation and rotation in the scene is too much for the algorithm to handle and resulted in it losing track of the prostate quickly after.

From the experiments performed, observations could be made about the performance of the registration algorithm. Overall, the registration algorithm performs adequately and can track rigid non-deformable motion of the prostate fairly well. Unfortunately, the algorithm lacks robustness and has a tendency to fail in certain circumstances. This is largely due to the nature of the algorithm depending on re-

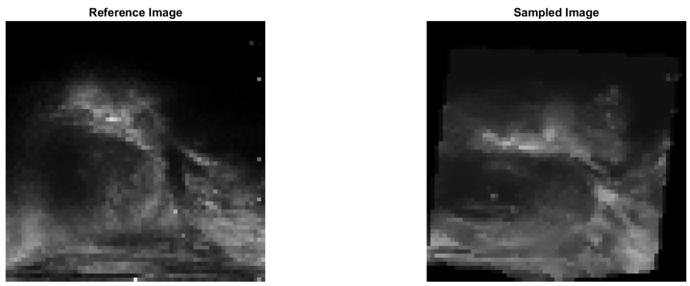


Figure 7.12: Registration algorithm unable to compensate for deformation

sults from previous iterations where any inaccuracies from previous cycles quickly compound and quickly lead to the registration becoming unstable and failing. At the same time, there are also cases where the algorithm is capable of recovering from such instabilities. The algorithm was able to slowly, over a series of frames locate the prostate and re-establish a reasonably accurate registration.

There are a few major failure cases which were observed during the experiment. The first problem is the latency of the algorithm where rapid motions of the ultrasound results in the algorithm unable to register quickly enough to account for the motion. This problem can potentially be prevented by further optimizing the performance of the algorithm and reducing response time as much as possible. Possible steps could include performing calculations using the Graphical Processor Unit (GPU) instead through Compute Unified Device Architecture (CUDA). The second major point of failure lies in the quality of the ultrasound image. As the transducer probe is set to track the surgeon's instrument, actions near the centre of the prostate provide a good view and high quality image of the prostate with many and distinct features. However, when the surgeon operates near the edges of the prostate, the ultrasound image loses the majority of its features as the prostate is no longer in view. These images often cause the optimizer to behave erratically and generate an incorrect transformation. This problem could also be potentially avoided by introducing checks which disable the optimization near the boundaries of the prostate volume. Alternatively, a mechanism could also be introduced to evaluate the quality of the current ultrasound image and remove it from the buffer if its determined to be too dark with insufficient features. Another problem noticed during the experiments was how the algorithm current reacts to non-rigid deformation. Often, the algorithm attempts to compensate for deformation through rotation which can quickly cause the registration to fail. To fix this problem, the non-rigid deformable registration algorithm should be introduced to account for this type of deformation.

7.6 Conclusion

Overall, the new registration algorithm was able to perform better than the baseline system currently being used. Depending on how parameters are configured and the computer hardware available, the algorithm can update the model at a frequency of 25 Hz with latency values below 100 ms. This results in a model which updates quickly with a reasonable degree of accuracy. The registration algorithm performs best when the TRUS volume used is similar to the current prostate anatomy. As the surgeon moves as operates on the region, changes to current ultrasound image begin to degrade the accuracy of the algorithm. Currently, the algorithm begins to fail when significant deformation takes place. This instability could potentially be alleviated when the non-rigid deformable registration algorithm is introduced to the system.

Chapter 8

Surgical Studies

Over the course of the project, an ongoing experimental study has been conducted to test the efficacy of the TRUS-MRI guidance system. The surgical studies are performed by two surgeons at the Department of Urological Sciences at Vancouver General Hospital. The studies detailed in this thesis specifically were conducted by Dr. Peter Black.

8.1 Study Workflow

In order to setup and use the image guidance system, a operating protocol has been developed. A high level overview of the current workflow is as follows:

1. Establish connection and stream data from da Vinci robot
2. Calibrate system by registering robot and ultrasound coordinate frames to track instrument position
3. Obtain 3D TRUS volume
4. Perform manual registration between 3D TRUS volume and preoperative

MRI

5. Enable guidance software system
   (a) Load registration and calibration
   (b) Start visualization output to TilePro system

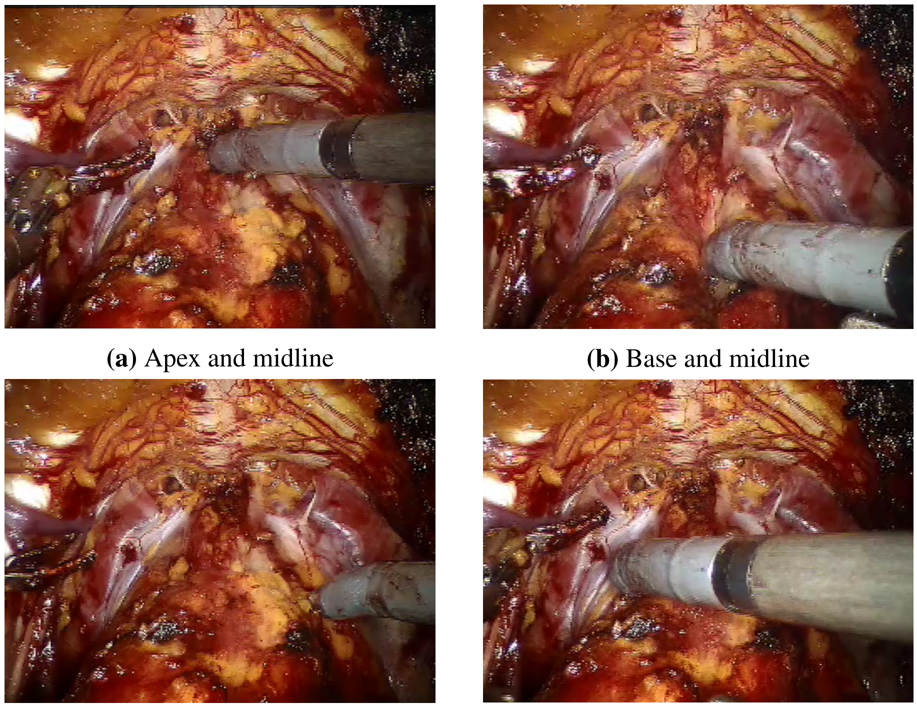
A more detailed and in-depth operating manual can be found in [Appendix].

Establishing Connection

The first step in the system workflow consists of establishing connection the da Vinci robot. In order to do so, an Ethernet cable must be connected from the desktop computer to the da Vinci robot. In order to establish connection, the desktop computer's IP address must be switched to the address block 10.0.0.X. This address block is reserved for local communications within a private network. Once this is completed, the robot control application can be started up to establish communications with the robot.

Calibration Procedure

The calibration procedure begins with the surgeon palpating the surface of the prostate using the da Vinci instrument tool tip. This palpation introduces motion to the ultrasound image which makes it easier to differentiate the tool tip from the surrounding anatomy. Using the motor control application, the ultrasound probe is manually rotated until the plane at which the instrument tool tip artifact can be visible with the highest intensity is determined. When located, the coordinates of both the da Vinci instrument and the location of the artefact are saved. The 3-dimensional coordinates of the da Vinci instrument are obtained through the robot control application. This process is completed for a minimum of four times at



(a) Apex and midline
(b) Base and midline
(c) Right and midgland
(d) Left and midgland

Figure 8.1: Calibration procedure for determining transformation between da Vinci robot and ultrasound coordinate system

different locations on the prostate surface which can be seen in Figure 8.1. Once all four points are obtained, a rigid transformation between the US and the da Vinci coordinate system is calculated and saved.

3D TRUS Volume

In order to obtain a 3D TRUS volume, the Motor Control application is used to rotate the probe for the entire span of its possible rotation. When obtaining this 3D volume, a certain consideration must be made where the number of frames must be balanced against the time used. While more frames will provide a more accurate and better visualization, it will also take more time. While in a controlled laboratory setting, the time taken may not be of great concern, it must be considered during an OR study. A longer movement time increases the chances of the imaged tissue moving as the surgeon operates. If the surgeon is to stop while the scan takes place, this time will delay the operation. Thus, a balance must between image quality and time taken must be ensured.

The robot system is programmed to have a maximum rotation angle of 50 degrees positive and negative. As a compromise between accuracy and speed, it was determined that 200 slices at 0.5 degrees per slice provides a high resolution volume sufficient for the system. During this sweep process, an additional delay must be introduced in between rotations. At each angle, the frame rate of the both the US acquisition and US screen must be considered. The system is delayed for the maximum time of both frame rates to ensure the US machine has enough time to update before the frame is grabbed. Once all the frames have been acquired, the resulting volume is written to the disk for the other applications to use.

TRUS-MRI Registration

The TRUS-MRI registration is performed using the registration application by a resident during the surgery. The preoperative MRI and the previously obtained 3D TRUS volume are both loaded and once loaded, the MRI volume boundary is projected on top of the TRUS image. Using the application, the clinician can rigidly and non-rigidly deform the projection to match the prostate boundary in the ultrasound image. The manual registration process can be seen in Figure 8.2 where the prostate outline in blue is rotated and translated until it fits best the prostate boundary. Once completed, the registration between the two volumes is computed and saved.

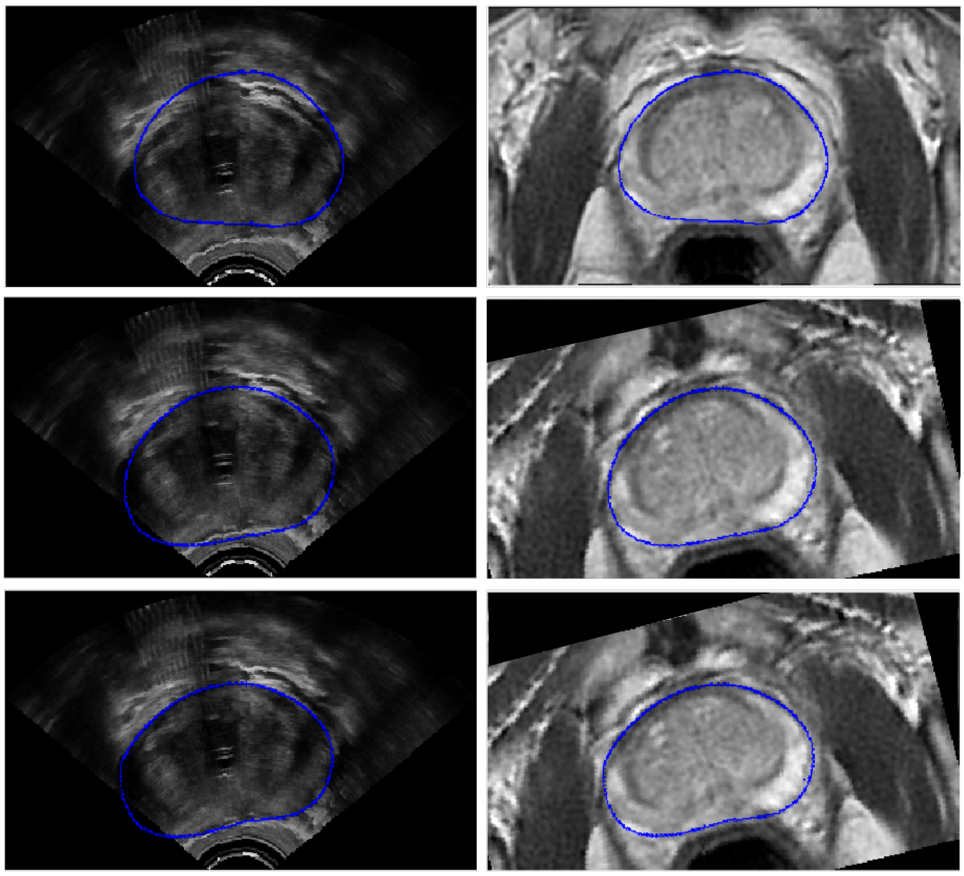


Figure 8.2: MRI to TRUS volume registration process

Guidance Software

The last step of the workflow consists of running the actual image guidance system. In order for the surgeon to be capable of viewing the virtual surgical scene, two SDI cables are connected from the DeckLink framegrabber card output to the TilePro input ports. The visualization application is started and a patient mesh is loaded along with the calibration and registration results. Once everything is loaded, the virtual scene is constructed and sent to the da Vinci console.

8.2 Study Results

Throughout the course of the work completed in this thesis, a total of 15 surgical studies have been completed. Samples from the recorded video from each of these cases can been seen in the following:

Patient 26

Patient 26 was the first surgical case tested with a few samples of the case shown in Figure 8.3. As the first case, it went surprisingly well with the calibration process performing well. As see in the Figure, it can noted that the ultrasound image is of very good quality with good contrasts between prostate and the surrounding tissue. One particular event of note was a data corruption problem during this case. It was discovered that the activation of cauterizer tool was causing troubling with the Epiphan Pearl recorder which resulted in parts of the data lost.

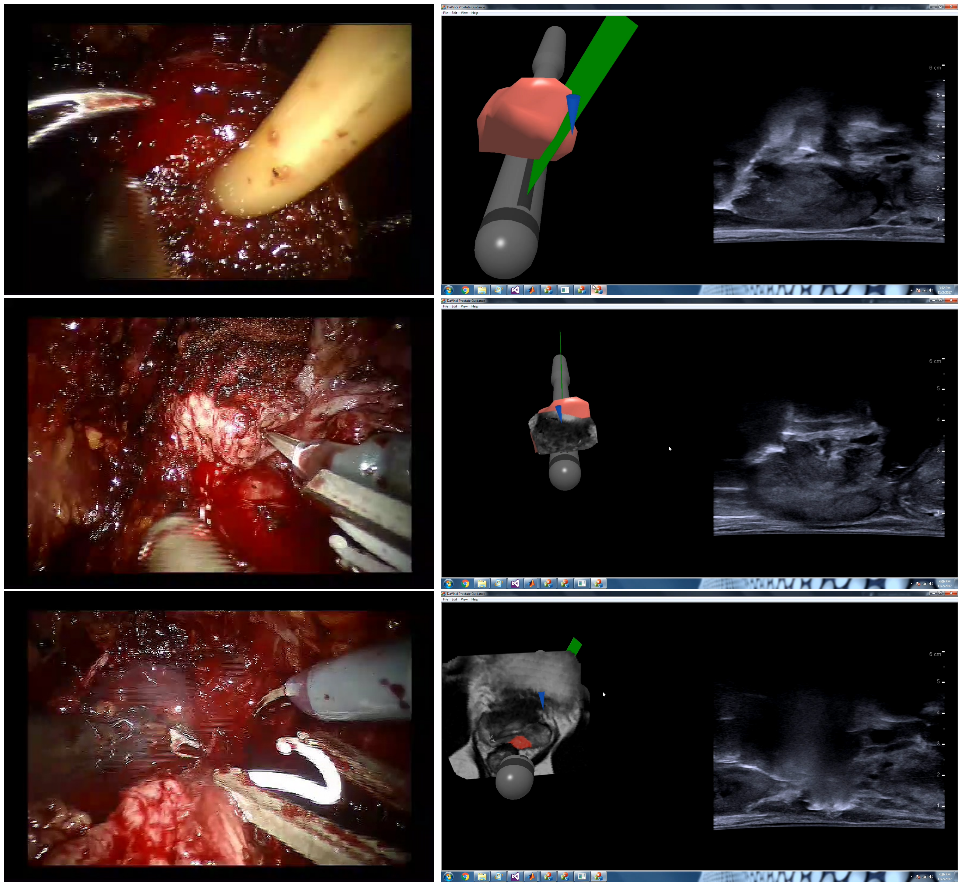


Figure 8.3: Samples from patient 26

Patient 27

Unlike the first case which went fairly smoothly, there were a few more difficulties in this case. The main problem was the ultrasound image which, as shown in Figure 8.4, contains a reflection artefact. This artefact is most likely caused by air remaining in the ultrasound probe cover. While this artefact is unfortunate, the resulting calibration appears to remain fine. During this case, the data corruption issue was resolved by connecting the power supply of the data recorder to another socket isolated from the cauterizer tool.

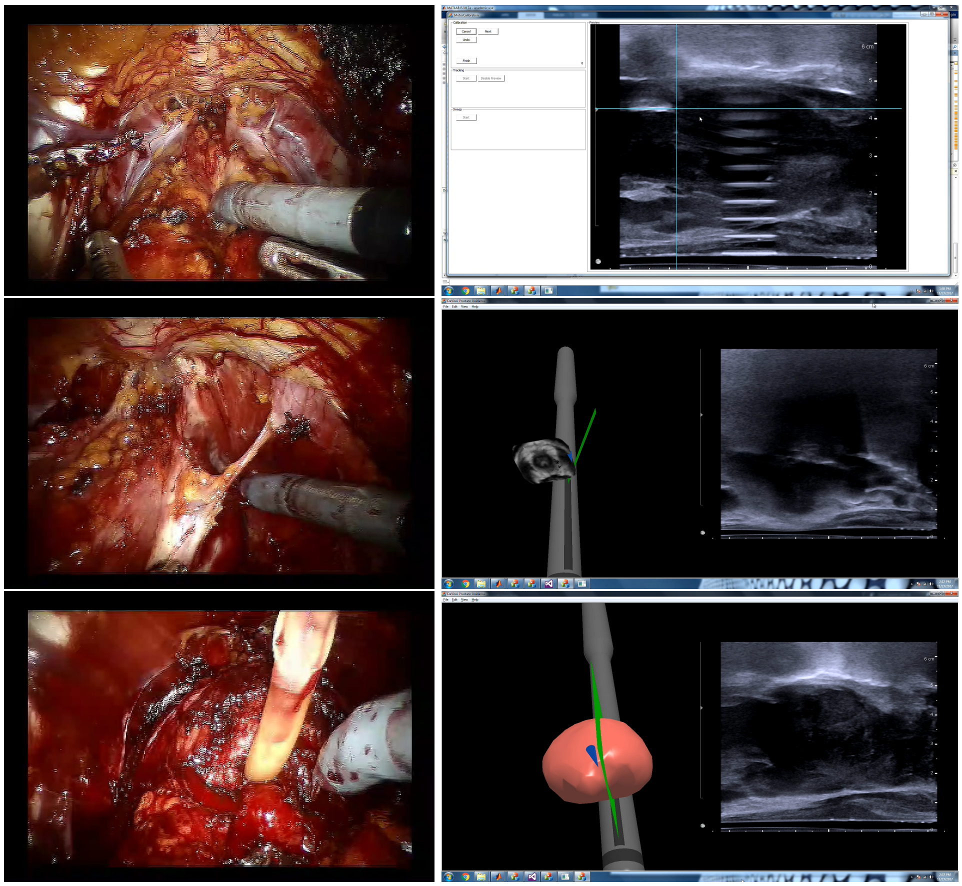


Figure 8.4: Samples from patient 27

Patient 28

Patient 28 was an interesting case where a problem with the MRI-TRUS registration occurred rather than during the calibration. As seen in Figure 8.5, the ultrasound quality was good which resulted in a good transformation from the calibration. However, in both the first two images, the surgeon was unable to view the slices of the MRI volume. This was most likely caused by a problem in the registration application.

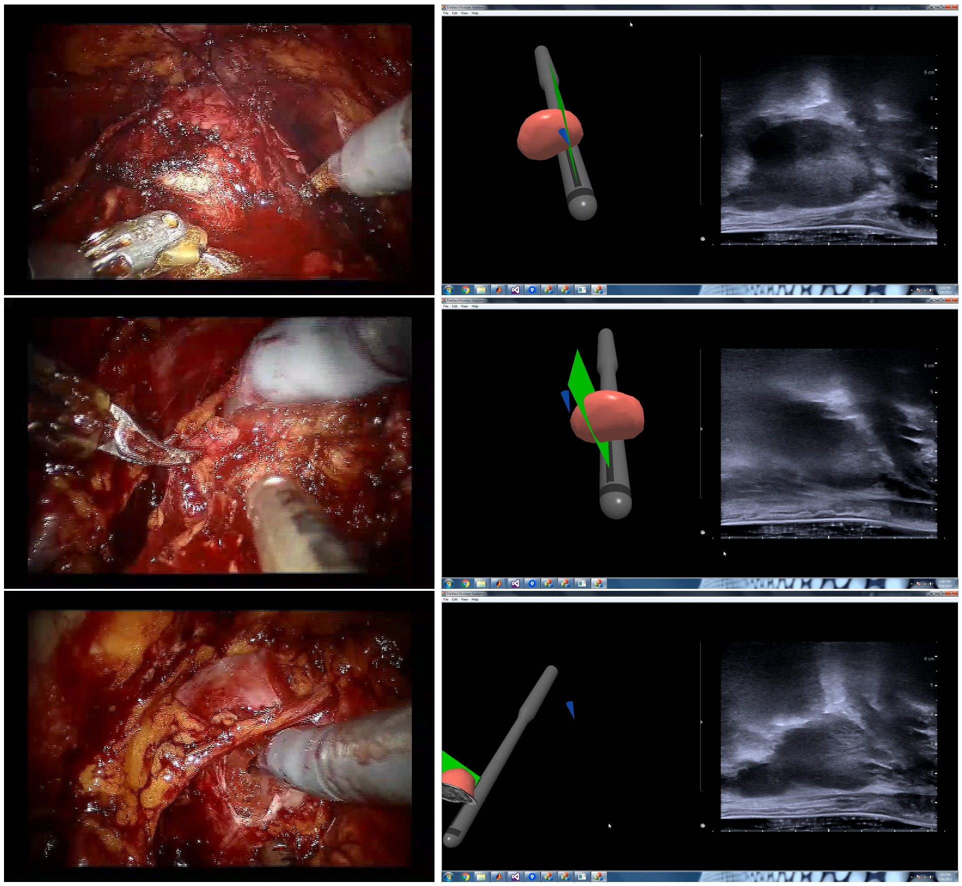


Figure 8.5: Samples from patient 28

Patient 29

Patient 29 was a routine case with very positive results. During the case, the highlighted tumour segmentation was introduced as shown in Figure 8.6. The tumour is seen highlighted bright in red in two of the images.

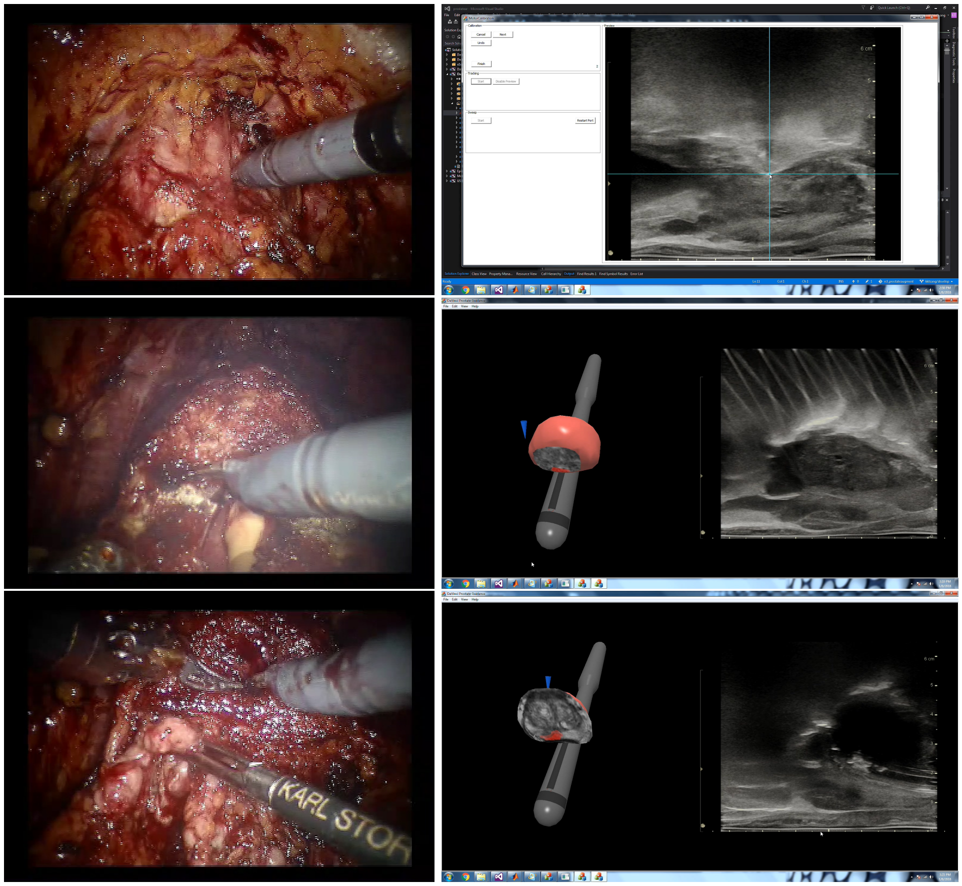


Figure 8.6: Samples from patient 29

Patient 30

Patient 30 was the first major failure in which an unfortunate software bug caused the system to fail. During the case, the calibration and registration processes occurred without any problems. However, a bug in the software resulted in the transformation file no loading properly. As seen in Figure 8.7, this resulted in the virtual scene not being visible.

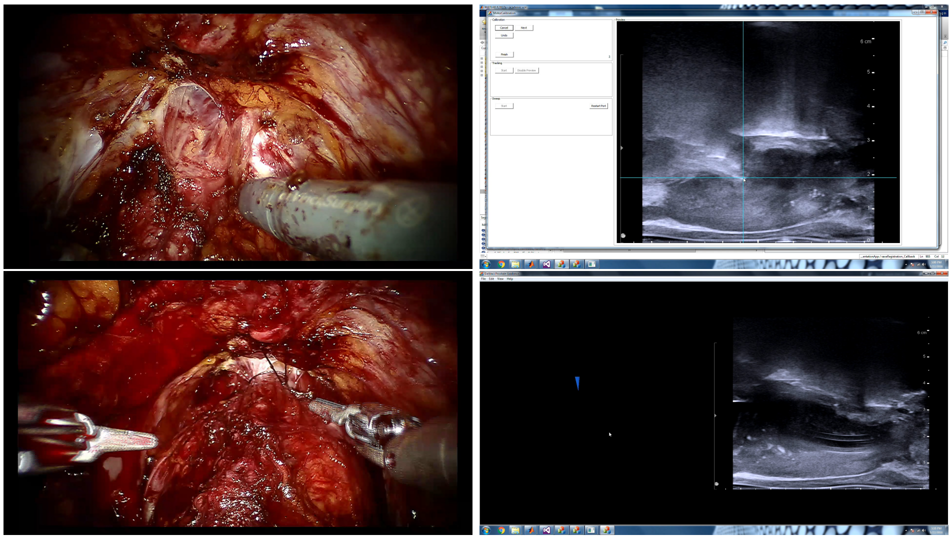


Figure 8.7: Samples from patient 30

Patient 33

Patient 33 was another case which unfortunately did not go well. Shown in Figure 8.8, it can easily seen that the ultrasound quality was terrible for this case. While it is difficult to determine the exact cause, the distortions in the ultrasound are either caused by stool in the patient or air in the probe cover. In either case, this resulting in an inaccurate calibration with the resulting view rotated.

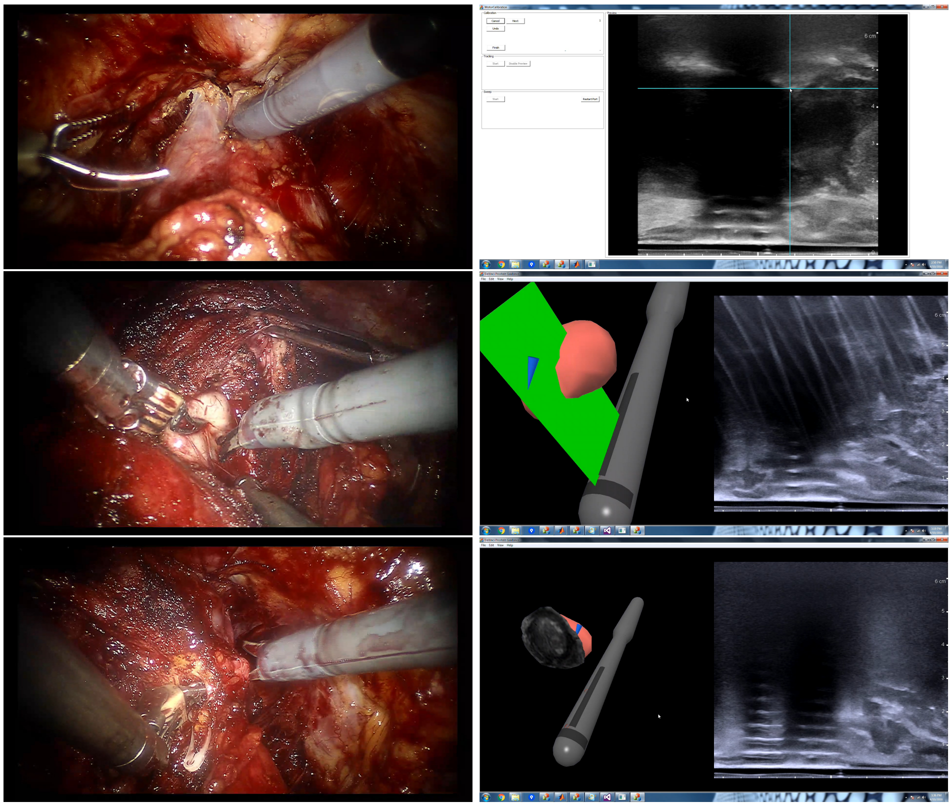


Figure 8.8: Samples from patient 33

Patient 36

In order to not continue the failures from the previous two cases, a major software rework and refactor was performed. As seen in Figure 8.9, the virtual scene is much refined with the ultrasound image more cleanly cropped. The ultrasound plane is recoloured to blue which makes it less glaring. Unfortunately, this case was also a failure which is most likely caused by a improperly positioned ultrasound problem. As seen in the ultrasound images, the prostate image is only half visible.

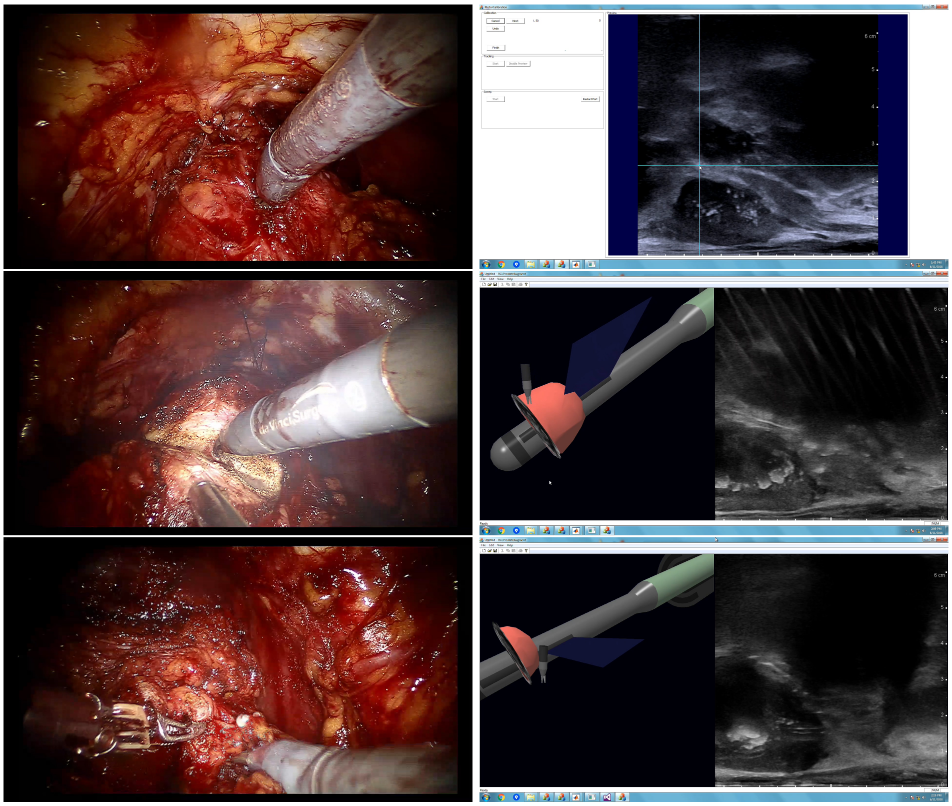


Figure 8.9: Samples from patient 36

Patient 38

Patient 38 was the first major success after a series of failed cases. As seen in Figure 8.10, the ultrasound image quality is very high which resulted in a very accurate calibration. This can be seen in the ultrasound images where the da Vinci instrument tooltip can be clearly seen.

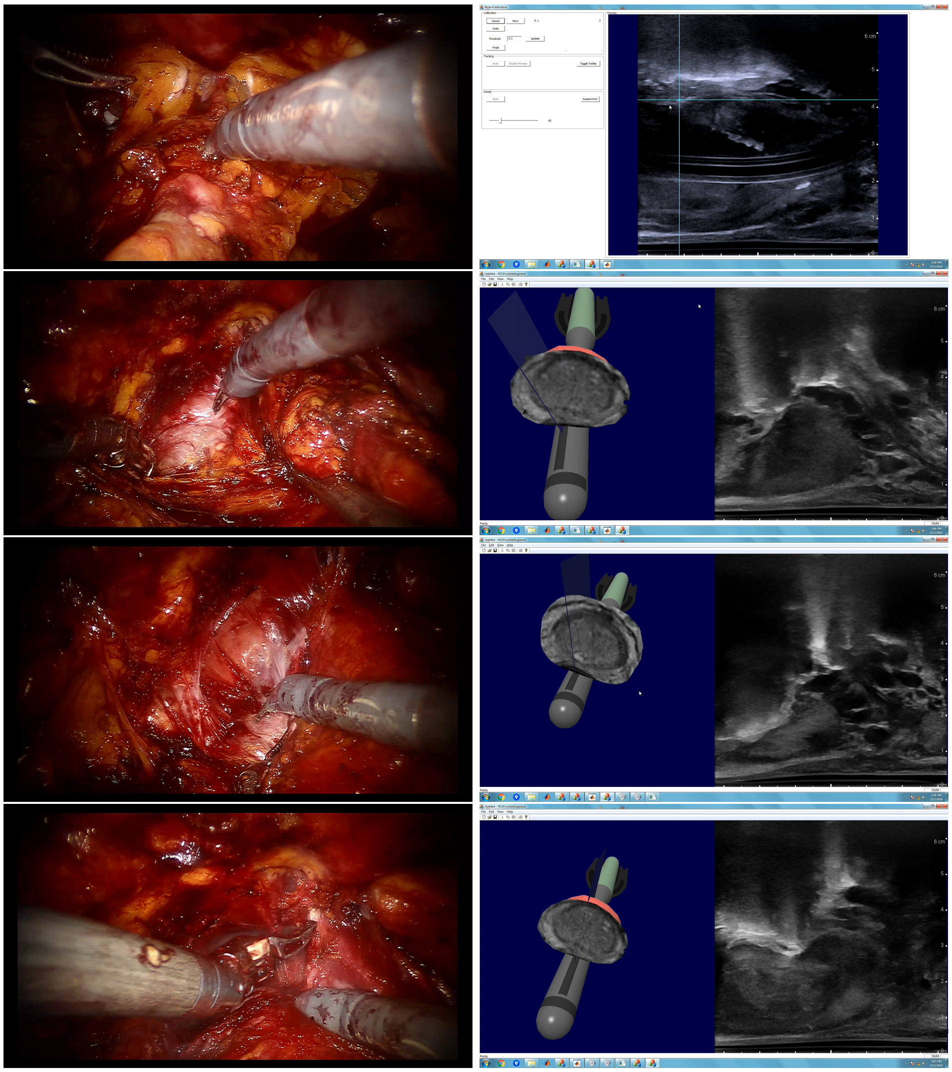


Figure 8.10: Samples from patient 38

Patient 39

Patient 39 was another unfortunate failure where once again, the ultrasound probe was improperly positioned. As seen in Figure 8.11, during the calibration process, the points palpitated by the surgeon could not be visible in the ultrasound image.

This resulted in a highly inaccurate transformation. During this case, Dr. Black also requested the ultrasound to be retracted early as the ultrasound probe was pushing up on the prostate.

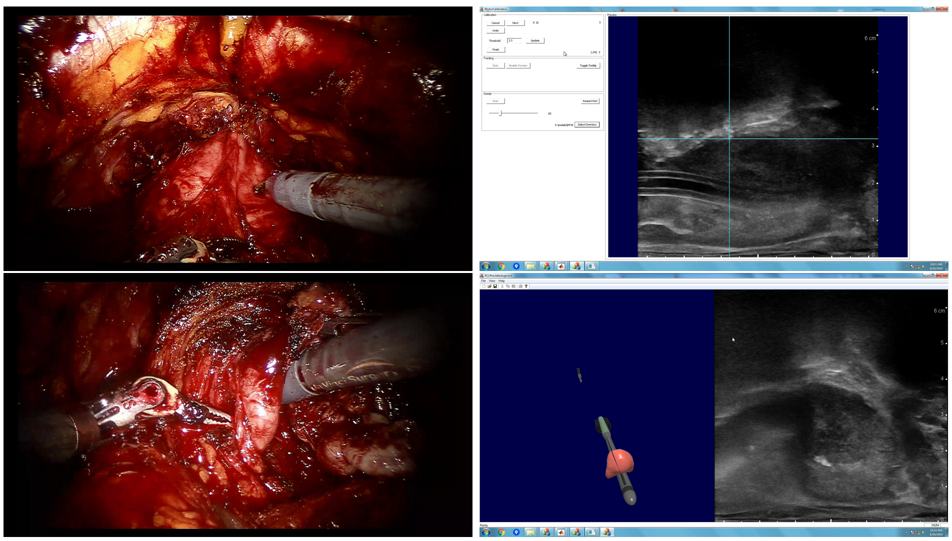


Figure 8.11: Samples from patient 39

Patient 40

Patient 40 was a successful routine case. The ultrasound image quality was great throughout the case as seen in Figure 8.12.

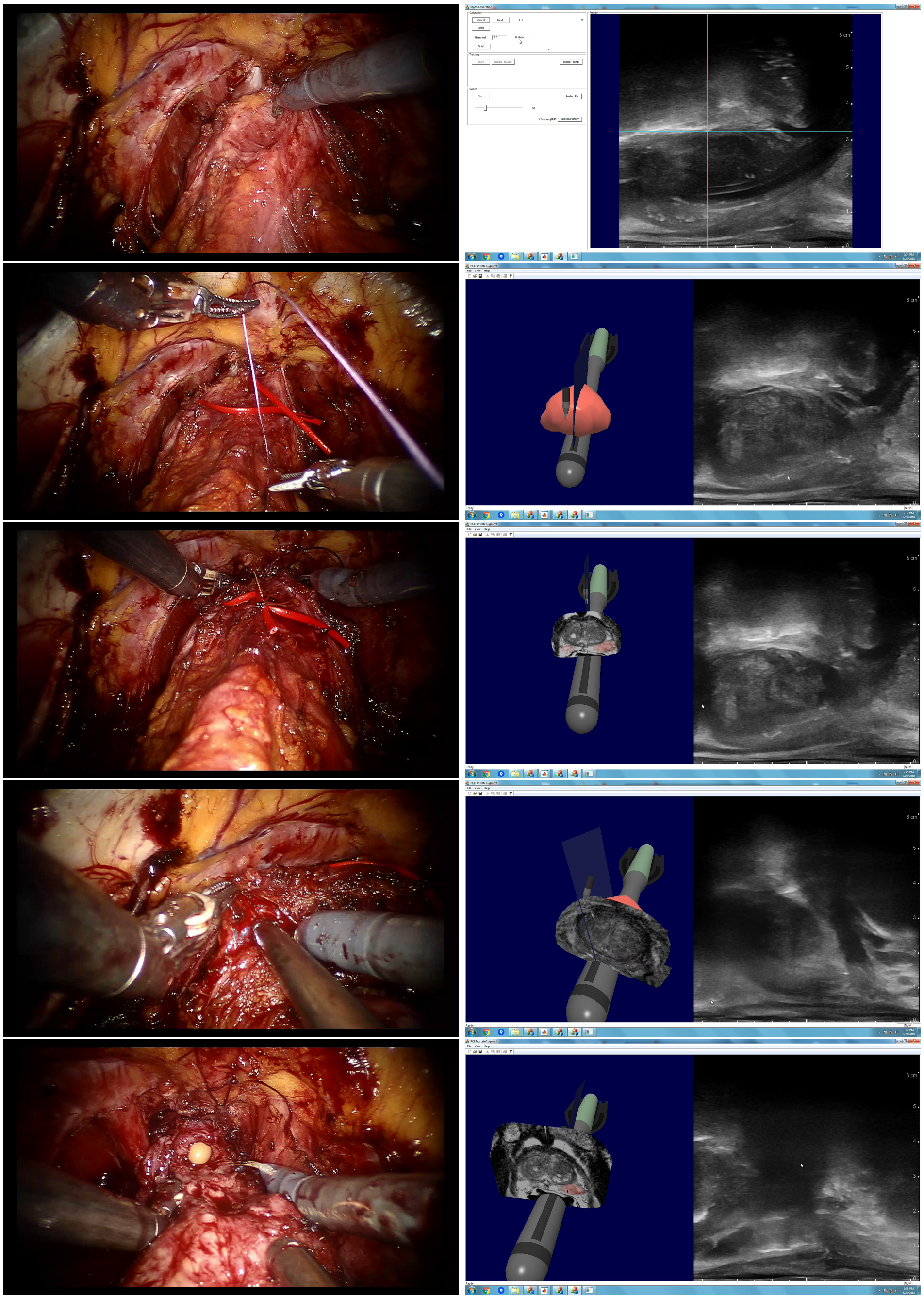


Figure 8.12: Samples from patient 40

Patient 41

Patient 41 was a partially successful case. Just like the previous times, the ultrasound probe was not inserted properly with only part of the prostate visible as seen in Figure 8.13. In order to compensate, we attempted to perform the calibration on only the visible parts of the prostate. The resulting transformation wasn't the best with the virtual scene only visible at a distance. During this case, the ultrasound overlay was first tested. Dr. Black did not like the solid colours as it obstructed the ultrasound image too much.

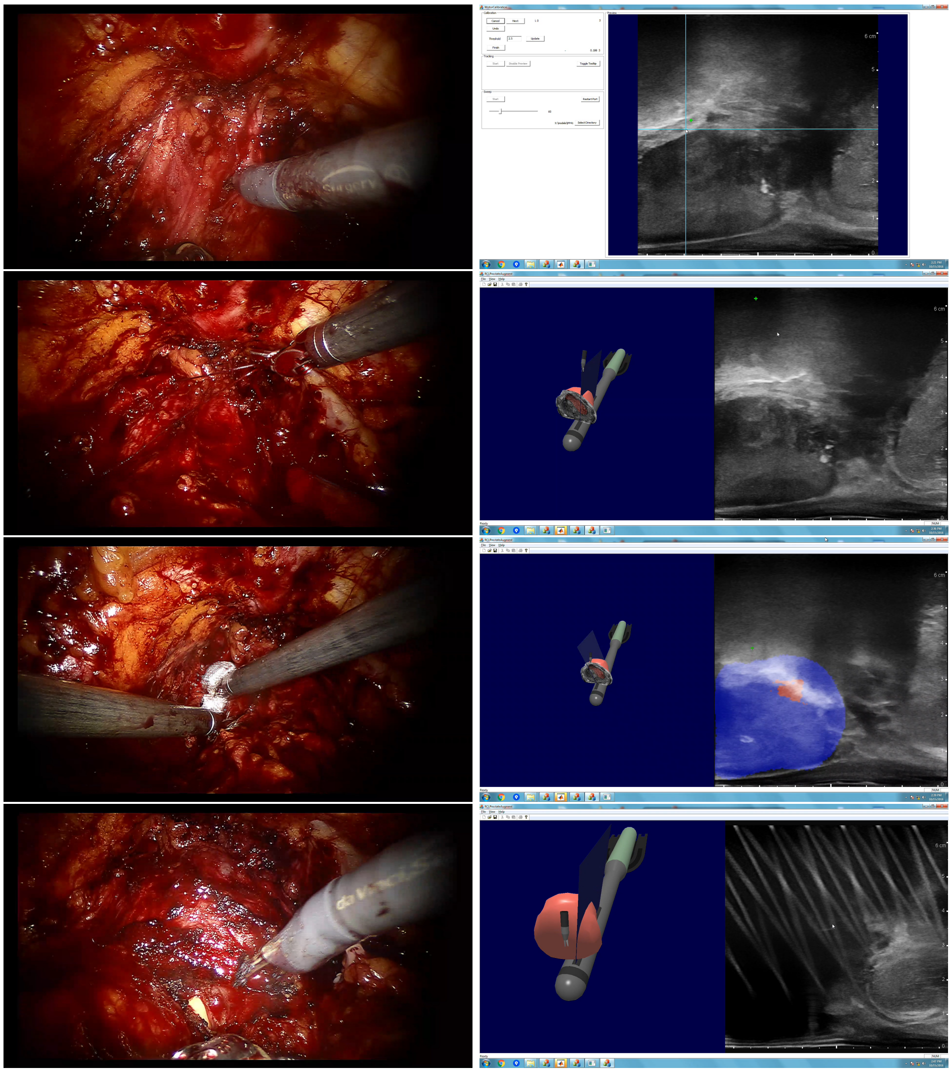


Figure 8.13: Samples from patient 41

Patient 42

Patient 42 was a successful routine case. During this case, the solid ultrasound overlay was replaced with just the prostate boundary outlined which can be seen in Figure 8.14. Dr. Black preferred this variant more found the overlay to be quite helpful for navigating the system.

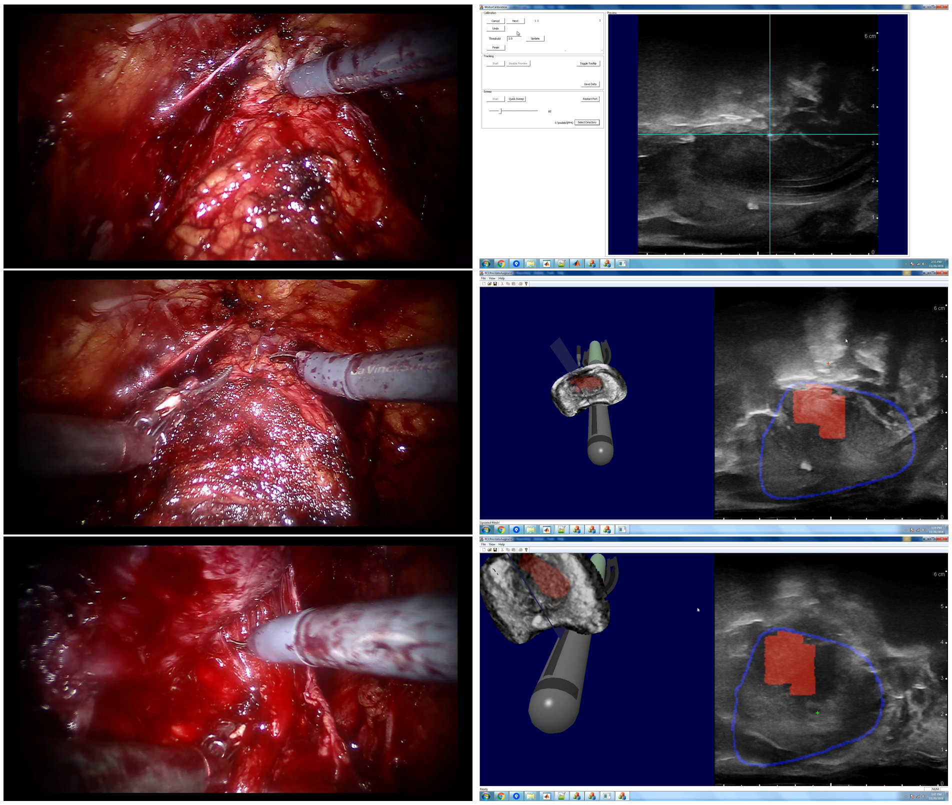


Figure 8.14: Samples from patient 42

Patient 43

Patient 43 was a more outlier case in which the patient had a very large prostate. During this case, the ultrasound image quality also wasn't very good which is likely due to the size of the prostate. While this made the calibration process more difficult, the resulting transformation wasn't terrible. Due to the outlier nature of the prostate, the boundary overlay didn't work properly as seen in Figure 8.15. During this case, Dr. Black requested the ultrasound probe to be retracted early again as it was pushing up on the prostate.

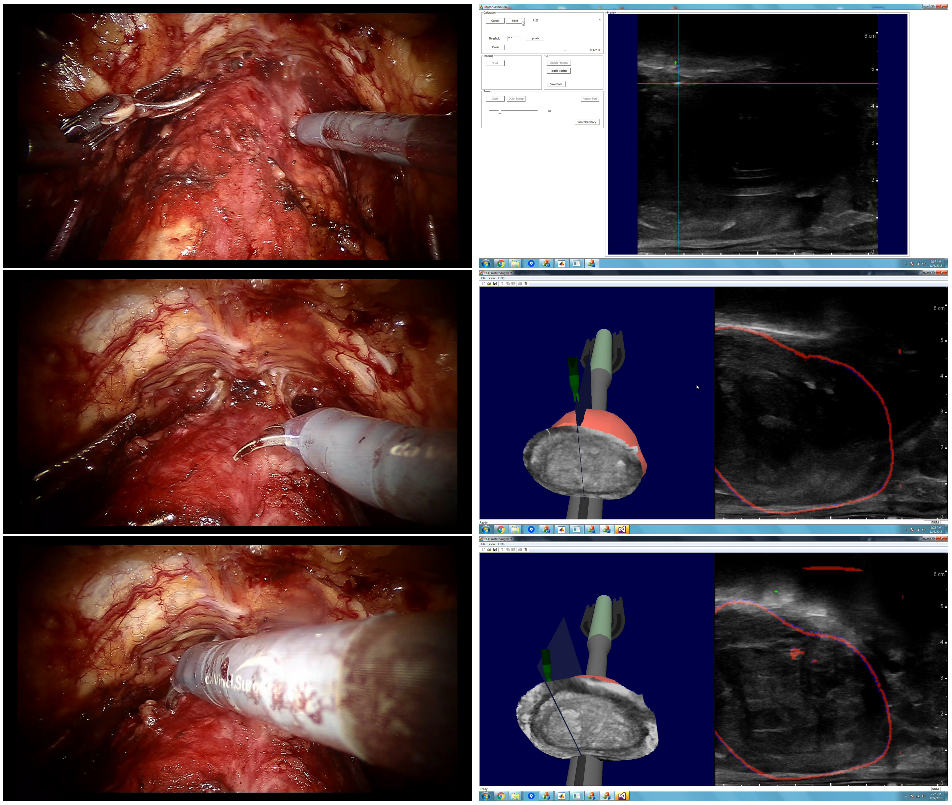


Figure 8.15: Samples from patient 43

Patient 44

Patient 44 was a fairly routine case with a lower than average ultrasound image quality. The ultrasound overlay was changed to also show the tumour as a boundary rather than a solid colour as seen in Figure 8.16. Unfortunately, this did not work properly with both the prostate and tumour sharing the same colour.

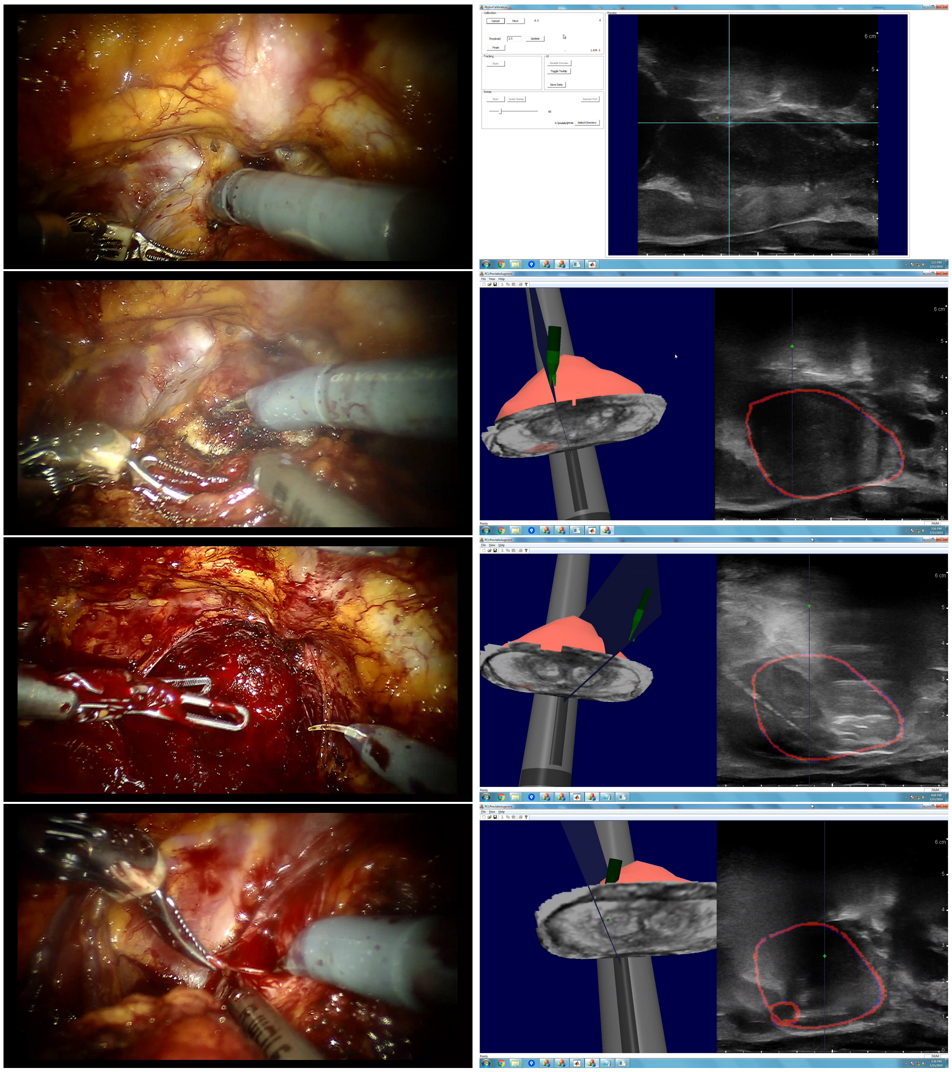


Figure 8.16: Samples from patient 44

Patient 46

Patient 46 was the last performed surgical case during the writing of this thesis. Unfortunately, the ultrasound quality of this case wasn't very good either. We were however successful in testing the fixed ultrasound overlay system with it working properly this time which can be seen in Figure 8.17.

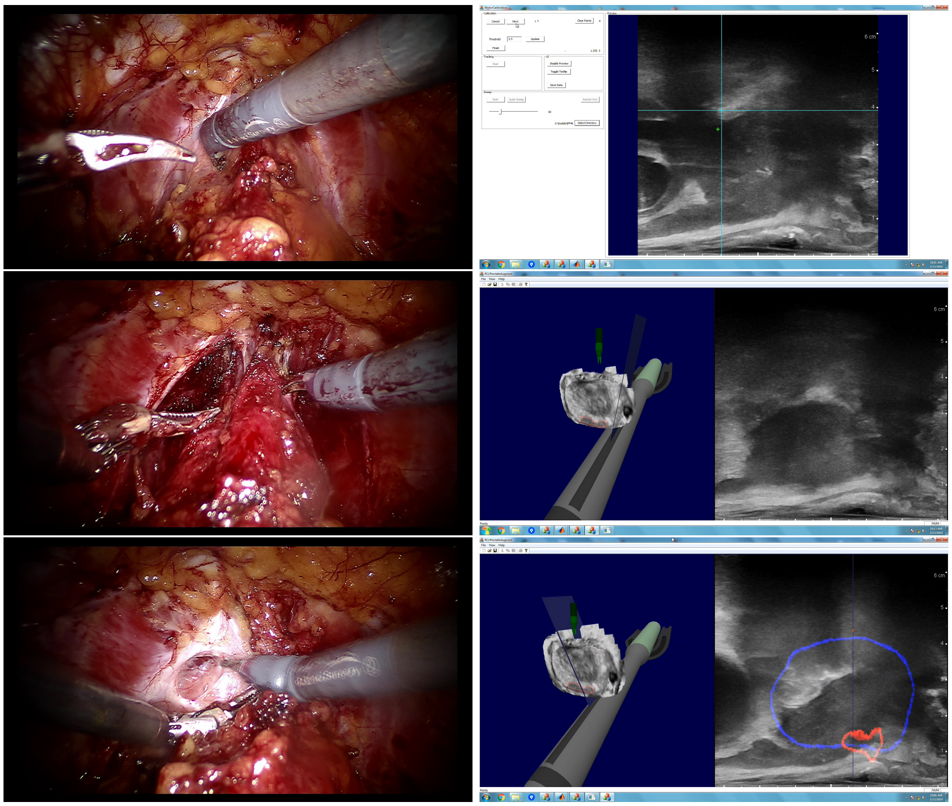


Figure 8.17: Samples from patient 46

8.3 Discussion

Currently, as the studies are yet to be completed, a quantitative review on the system performance have yet to be conducted. Instead, most of the evaluation of the system has been conducted through the comments made by the surgeon using the system, Dr. Black. Looking at the snapshots from the recorded surgical videos and image guidance system, the progression of the software system can be easily seen. Patient 26 shows the earliest version of the software where only the framegrabber image capturing pipeline was integrated, while from Patient 38 onwards one can see the refactored and redesigned rendering framework. During Patient 41's case, the first version of the ultrasound augmented reality overlay was being tested with the completed version appearing in Patient 46's study. The first version of the overlay seen in Figure 8.13 overplayed the entire prostate region on the ultrasound image. Dr. Black commented that this overlay was obstructive and blocked his view of the ultrasound image. As a result, a boundary based overlay was tested over the next few cases with the final version appearing in Figure 8.17. In this version, Dr. Black like the overlay as it allowed him to see the tumour region in both the ultrasound and MRI.

Of all the cases, 4 cases were considered failures where either the calibration was incorrect OR a software bug caused problems. Patient 30, seen in Figure 8.7, was a case where a software problem caused the transformation to not be calculated properly and resulted in the image guidance system not working. Patient 33 and 36 seen in Figure 8.8 and Figure 8.9 also failed from a poor calibration but likely due to another reason. During both these cases, another member of the project began collecting elastography data for their own study. In order to do so, the cradle for the ultrasound probe was changed to allow for a vibration motor to be installed for obtaining elastography data. This cradle however, has a shorter length than the previous cradle used. It is very likely that this shorter cradle prevented the ultrasound probe from properly reaching the prostate region. This was corrected by modifying the cradle for the next surgical case and from the results seen in the next case in Figure 8.10, the problem appears to have be resolved.

In 2 cases, seen in Figure 8.11 and Figure 8.15, Dr. Black had requested the ultrasound probe to be removed OR retracted. In both cases, Dr. Black indicated that the probe was pushing up on the prostate. It should be noted that one difference between the current study and Mohareri Omid et al. [33]'s study is the current resident had not undergone any training on brachytherapy ultrasound probe placement. This is of some concern as there have been a few cases where the ultrasound image was too dark likely due to poor contact with the prostate. This contact is largely dependent on how the brachytherapy mount is placed and not controlled by any of the software systems.

Overall throughout all the studies, whenever the calibration was successful, Dr. Black would use the visualization on an average of two times per case. During some more difficult surgeries, Dr. Black did not have time to test the software system, but during regular cases he uses it at least 2 to 3 times. During these times, Dr. Black has had positive comments on the system. However, he does constantly point out the key problem which is without a tracking mechanism, he cannot trust the image guidance system to be accurate when he uses it. It was through these comments that the system has been continuously improved with additional features and enhancements.

8.4 Conclusion

Over the course of the work completed in this thesis, an ongoing study was in progress in order to determine the effectiveness of the image guidance system. The system was tested on 15 patients in total and in general, the feedback from the surgeon has been positive. Dr. Black would refer to the image guidance system on an average of 2 times per surgical case. Due to the ongoing state of the study, a qualitative evaluation has yet to be performed. While the current results are positive, much work still needs to be completed in order to address the remaining concerns of the surgeon as well improving the robustness of the system. Additional steps need to be taken in order to improve the calibration process and reduce the chances of a poor transformation. The real-time registration algorithm from Chapter 7 also needs to be tested.

Chapter 9

Conclusions

The overall objective of this project is the improvement of patient outcomes in RALRP through the introduction of an augmented reality image guidance system. To achieve this goal, four main objectives were established and accomplished.

In Chapter 4, a imaging pipeline framework was developed as a method for strengthening depth cues for surgeons using the augmented reality system. This framework includes both hardware using a framegrabber for outputting video signals to the da Vinci surgeon console as well a software graphics framework for rendering scenes from multiple angles. The current feedback from the surgeon has largely been positive. The stereo effect does works as intended and helps provide the surgeon with a depth cues when traversing the visualization software.

In Chapter 5, a new motor control application was developed by integrating the previously developed motor libraries with a new user interface along with the DeckLink framework for ultrasound images. Improvements to the calibration protocol resulted in a more robust and accurate workflow. First, the da Vinci to ultrasound registration algorithm was changed to allow for better outlier rejection.

This greatly increases accuracy and robustness of the calibration as it allows for more samples to be taken. Next, workflow feedback mechanisms were introduced to help determine when a calibration was successful. When used in current studies, the new revised calibration protocol has demonstrated greater success rates than before. Unfortunately, the current calibration protocol is still heavily reliant on user skill on recognizing image in the ultrasound image. Much work could still be done in improving the reliability of the system.

In Chapter 6, the visualization application was developed to fulfill the requirements established by the project. This includes the rendering of the prostate mesh along with the real-time ultrasound images. An ultrasound augmentation system was introduced in the form of boundary overlays on top of the ultrasound image. This framework allows for the easier visualization and localization of tumours in the virtual prostate scene. With instrument tip locations indicated in both sagittal and transverse views, the surgeon can easily determine their instrument position relative to the prostate. The current overlay system has received positive feedback from the surgeon.

In Chapter 7, prostate motion compensation was introduced in the form of a real-time registration algorithm which updates the current model based on live ultrasound images. This allows the visualization to maintain the MRI to ultrasound registration over the course of the surgery. The current implementation is only capable of tracking rigid non-deformable prostate motion. While current experiments show positive results in performance, the system is still rather early in development and currently not ready for intraoperative testing. The current algorithm lacks robustness especially when dealing with rapid motion and poor ultrasound images.

In Chapter 8, the results of the surgical studies were analyzed qualitatively through an analysis of failures as well as comments from the surgeon using the system. Overall, the surgeon's feedback has been positive with him using the system at least twice in each case. Of the surgical cases where the software system failed, the majority of them were caused by a poor calibration often caused by a poor ultrasound image. The other significant comment made by the surgeon was the request for a tracking system. Steps towards addressing these concerns were taken during the work in Chapter 5 as well as Chapter 7.

While the current system has been improved significantly since its previous state, there remain many more improvements that could be made. First, while the current calibration process does function adequately, it is not robust enough to guarantee an accurate calibration. The current system relies heavily on the user's ability to determine the location of the instrument tool tip in the ultrasound image. When the ultrasound image quality is high, locating the instrument tool tip artifact is not very difficult. Unfortunately, during patient studies, the ultrasound image quality is rarely ideal. Frameworks such as the automatic tool tip detection system developed in Mohareri et al. [32] could be introduced to help eliminate some of the uncertainty during this calibration process. Another area of improvement is in the real-time registration system. While the current framework works sufficiently well for rigid, non-deformable motion, this is hardly the only type of motion present during a surgery. As such, the non-rigid deformable update should also be introduced to better handle prostate motion. Any extension to the registration algorithm is likely to also increase runtime and latency. As such, additional work should also be done in optimizing the algorithms used through systems such as CUDA GPU-based processing. While there remains much work which could be done, this project has largely fulfill the goals it was set out to accomplish.

Bibliography

[1] T. Adebar, S. Salcudean, S. Mahdavi, M. Moradi, C. Nguan, and L. Goldenberg. A Robotic System for Intra-operative Trans-Rectal Ultrasound and Ultrasound Elastography in Radical Prostatectomy. In R. H. Taylor and G.-Z. Yang, editors, *Information Processing in Computer-Assisted Interventions*, Lecture Notes in Computer Science, pages 79–89. Springer Berlin Heidelberg, 2011. ISBN 978-3-642-21504-9. → pages 27, 28, 31, 61, 65

[2] 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, Apr. 2017. ISSN 1361-8415. doi:10.1016/j.media.2017.01.007. URL http://www.sciencedirect.com/science/article/pii/S1361841517300178. → page 16

[3] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6):394–424, 2018. ISSN 1542-4863. doi:10.3322/caac.21492. URL https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21492. → page 1

[4] J. Cepek, B. A. Chronik, U. Lindner, J. Trachtenberg, S. R. H. Davidson, J. Bax, and A. Fenster. A system for MRI-guided transperineal delivery of needles to the prostate for focal therapy. *Medical Physics*, 40(1):012304, 2013. ISSN 2473-4209. doi:10.1118/1.4773043. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4773043. → page 6

[5] C. Conrad, M. Fusaglia, M. Peterhans, H. Lu, S. Weber, and B. Gayet. Augmented Reality Navigation Surgery Facilitates Laparoscopic Rescue of Failed Portal Vein Embolization. *Journal of the American College of*

Surgeons, 223(4):e31–e34, Oct. 2016. ISSN 1072-7515, 1879-1190. doi:10.1016/j.jamcollsurg.2016.06.392. URL https://www.journalacs.org/article/S1072-7515(16)30681-0/abstract. → page 18

[6] J. W. Davis, U. S. Kreaden, J. Gabbert, and R. Thomas. Learning Curve Assessment of Robot-Assisted Radical Prostatectomy Compared with Open-Surgery Controls from the Premier Perspective Database. *Journal of Endourology*, 28(5):560–566, May 2014. ISSN 0892-7790. doi:10.1089/end.2013.0534. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3995359/. → page 4

[7] J. M. Fitzpatrick. Fiducial registration error and target registration error are uncorrelated. In *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, volume 7261, page 726102. International Society for Optics and Photonics, Mar. 2009. doi:10.1117/12.813601. URL https://www.spiedigitallibrary.org/conference-proceedings-of-spie/7261/726102/Fiducial-registration-error-and-target-registration-error-are-uncorrelated/10.1117/12.813601.short. → page 69

[8] S. Fuchs and S. May. Calibration and registration for precise surface reconstruction with Time-Of-Flight cameras. *IJISTA*, 5:274–284, Jan. 2008. doi:10.1504/IJISTA.2008.021290. → page 20

[9] J. J. Fütterer, A. Briganti, P. De Visschere, M. Emberton, G. Giannarini, A. Kirkham, S. S. Taneja, H. Thoeny, G. Villeirs, and A. Villers. Can Clinically Significant Prostate Cancer Be Detected with Multiparametric Magnetic Resonance Imaging? A Systematic Review of the Literature. *European Urology*, 68(6):1045–1053, Dec. 2015. ISSN 0302-2838. doi:10.1016/j.eururo.2015.01.013. URL http://www.sciencedirect.com/science/article/pii/S0302283815000366. → page 5

[10] O. Goksel, S. E. Salcudean, S. P. DiMaio, R. Rohling, and J. Morris. 3d Needle-Tissue Interaction Simulation for Prostate Brachytherapy. In J. S. Duncan and G. Gerig, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2005*, Lecture Notes in Computer Science, pages 827–834. Springer Berlin Heidelberg, 2005. ISBN 978-3-540-32094-4. → pages 10, 11, 29, 40, 89

[11] S. A. Gudmundsson, H. Aanaes, and R. Larsen. Fusion of stereo vision and Time-Of-Flight imaging for improved 3d estimation. *International Journal*

of Intelligent Systems Technologies and Applications, 5(3/4):425, 2008. ISSN 1740-8865, 1740-8873. doi:10.1504/IJISTA.2008.021305. URL http://www.inderscience.com/link.php?id=21305. → page 20

[12] M. Han, C. Kim, P. Mozer, F. Schfer, S. Badaan, B. Vigaru, K. Tseng, D. Petrisor, B. Trock, and D. Stoianovici. Tandem-robot Assisted Laparoscopic Radical Prostatectomy to Improve the Neurovascular Bundle Visualization: A Feasibility Study. *Urology*, 77(2):502–506, Feb. 2011. ISSN 0090-4295. doi:10.1016/j.urology.2010.06.064. URL http://www.sciencedirect.com/science/article/pii/S0090429510014950. → page 7

[13] J. V. Hegde, R. V. Mulkern, L. P. Panych, F. M. Fennessy, A. Fedorov, S. E. Maier, and C. M. C. Tempany. Multiparametric MRI of prostate cancer: An update on state-of-the-art techniques and their performance in detecting and localizing prostate cancer. *Journal of Magnetic Resonance Imaging*, 37(5): 1035–1054, 2013. ISSN 1522-2586. doi:10.1002/jmri.23860. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.23860. → page 5

[14] A. Heidenreich, P. J. Bastian, J. Bellmunt, M. Bolla, S. Joniau, T. v. d. Kwast, M. Mason, V. Matveev, T. Wiegel, F. Zattoni, and N. Mottet. EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent Update 2013. *European Urology*, 65(1):124 – 137, 2014. ISSN 0302-2838. doi:https://doi.org/10.1016/j.eururo.2013.09.046. URL http://www.sciencedirect.com/science/article/pii/S0302283813010403. → page 1

[15] A. J. Hung, A. L. D. C. Abreu, S. Shoji, A. C. Goh, A. K. Berger, M. M. Desai, M. Aron, I. S. Gill, and O. Ukimura. Robotic Transrectal Ultrasonography During Robot-Assisted Radical Prostatectomy. *European Urology*, 62(2):341–348, Aug. 2012. ISSN 0302-2838. doi:10.1016/j.eururo.2012.04.032. URL http://www.sciencedirect.com/science/article/pii/S0302283812004988. → page 7

[16] M. Innmann, M. Zollhfer, M. Niener, C. Theobalt, and M. Stamminger. VolumeDeform: Real-time Volumetric Non-rigid Reconstruction. *arXiv:1603.08161 [cs]*, Mar. 2016. URL http://arxiv.org/abs/1603.08161. arXiv: 1603.08161. → page 24

[17] M. Kibsgaard and M. Kraus. Measuring the Latency of an Augmented Reality System for Robot-assisted Minimally Invasive Surgery. In *VISIGRAPP*, 2017. doi:10.5220/0006274203210326. → page 94

[18] Y. M. Kim, C. Theobalt, J. Diebel, J. Kosecka, B. Miscusik, and S. Thrun. Multi-view image and ToF sensor fusion for dense 3d reconstruction. In *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pages 1542–1549, Sept. 2009. doi:10.1109/ICCVW.2009.5457430. → page 20

[19] D. E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009. → page 100

[20] A. Kolagunda, S. Sorensen, S. Mehralivand, P. Saponaro, W. Treible, B. Turkbey, P. Pinto, P. Choyke, and C. Kambhamettu. A mixed reality guidance system for robot assisted laparoscopic radical prostatectomy. In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*, pages 164–174. Springer, 2018. → page 16

[21] B. Lin, Y. Sun, X. Qian, D. Goldgof, R. Gitlin, and Y. You. Video-based 3d reconstruction, laparoscope localization and deformation recovery for abdominal minimally invasive surgery: a survey. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 12(2):158–178, June 2016. ISSN 1478-596X. doi:10.1002/rcs.1661. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rcs.1661. → page 16

[22] J. Lin, N. T. Clancy, D. Stoyanov, and D. S. Elson. Tissue Surface Reconstruction Aided by Local Normal Information Using a Self-calibrated Endoscopic Structured Light System. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, pages 405–412. Springer, Cham, Oct. 2015. ISBN 978-3-319-24552-2 978-3-319-24553-9. doi:10.1007/978-3-319-24553-9_50. URL https://link.springer.com/chapter/10.1007/978-3-319-24553-9_50. → page 20

[23] Q. Liu-Yin, R. Yu, L. Agapito, A. Fitzgibbon, and C. Russell. Better Together: Joint Reasoning for Non-rigid 3d Reconstruction with Specularities and Shading. *arXiv:1708.01654 [cs]*, Aug. 2017. URL http://arxiv.org/abs/1708.01654. arXiv: 1708.01654. → page 23

[24] J. K. Logan, S. RaisBahrami, B. Turkbey, A. Gomella, H. Amalou, P. L. Choyke, B. J. Wood, and P. A. Pinto. Current status of magnetic resonance imaging (MRI) and ultrasonography fusion software platforms for guidance of prostate biopsies. *BJU International*, 114(5):641–652, 2014. ISSN 1464-410X. doi:10.1111/bju.12593. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/bju.12593. → page 9

[25] L. Maier-Hein, A. Groch, A. Bartoli, S. Bodenstedt, G. Boissonnat, P.-L. Chang, N. T. Clancy, D. S. Elson, S. Haase, E. Heim, J. Hornegger, P. Jannin, H. Kenngott, T. Kilgus, B. Mller-Stich, D. Oladokun, S. Rhl, T. R. dos Santos, H.-P. Schlemmer, A. Seitel, S. Speidel, M. Wagner, and D. Stoyanov. Comparative Validation of Single-Shot Optical Techniques for Laparoscopic 3-D Surface Reconstruction. *IEEE Transactions on Medical Imaging*, 33(10):1913–1930, Oct. 2014. ISSN 0278-0062, 1558-254X. doi:10.1109/TMI.2014.2325607. → page 21

[26] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur. Beyond Lambert: Reconstructing specular surfaces using color. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 619–626 vol. 2, June 2005. doi:10.1109/CVPR.2005.88. → page 23

[27] A. Malti and A. Bartoli. Combining Conformal Deformation and CookTorrance Shading for 3-D Reconstruction in Laparoscopy. *IEEE Transactions on Biomedical Engineering*, 61(6):1684–1692, June 2014. ISSN 0018-9294, 1558-2531. doi:10.1109/TBME.2014.2300237. → page 20

[28] J. Marescaux, J. Leroy, M. Gagner, F. Rubino, D. Mutter, M. Vix, S. E. Butner, and M. K. Smith. Transatlantic robot-assisted telesurgery. *Nature*, 413(6854):379–380, Sept. 2001. ISSN 1476-4687. doi:10.1038/35096636. URL https://www.nature.com/articles/35096636. → pages 17, 94

[29] T. D. McClure, D. J. A. Margolis, R. E. Reiter, J. W. Sayre, M. A. Thomas, R. Nagarajan, M. Gulati, and S. S. Raman. Use of MR Imaging to Determine Preservation of the Neurovascular Bundles at Robotic-assisted Laparoscopic Prostatectomy. *Radiology*, 262(3):874–883, Mar. 2012. ISSN 0033-8419. doi:10.1148/radiol.11103504. URL https://pubs.rsna.org/doi/full/10.1148/radiol.11103504. → page 6

[30] G. Megali, V. Ferrari, C. Freschi, B. Morabito, F. Cavallo, G. Turini, E. Troia, C. Cappelli, A. Pietrabissa, O. Tonet, A. Cuschieri, P. Dario, and

F. Mosca. EndoCAS navigator platform: a common platform for computer and robotic assistance in minimally invasive surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 4(3):242–251, 2008. ISSN 1478-596X. doi:10.1002/rcs.203. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rcs.203. → page 18

[31] Menon Mani, Shrivastava Alok, Tewari Ashutosh, Sarle Richard, Hemal Ashok, Peabody James O., and Vallancien Guy. Laparoscopic and Robot Assisted Radical Prostatectomy: Establishment of a Structured Program and Preliminary Analysis of Outcomes. *Journal of Urology*, 168(3):945–949, Sept. 2002. doi:10.1016/S0022-5347(05)64548-X. URL https://www.auajournals.org/doi/full/10.1016/S0022-5347%2805%2964548-X. → page 3

[32] O. Mohareri, M. Ramezani, T. K. Adebar, P. Abolmaesumi, and S. E. Salcudean. Automatic Localization of the da Vinci Surgical Instrument Tips in 3-D Transrectal Ultrasound. *IEEE Transactions on Biomedical Engineering*, 60(9):2663–2672, Sept. 2013. ISSN 0018-9294. doi:10.1109/TBME.2013.2262499. → pages 74, 145

[33] Mohareri Omid, Ischia Joseph, Black Peter C., Schneider Caitlin, Lobo Julio, Goldenberg Larry, and Salcudean Septimiu E. Intraoperative Registered Transrectal Ultrasound Guidance for Robot-Assisted Laparoscopic Radical Prostatectomy. *Journal of Urology*, 193(1):302–312, Jan. 2015. doi:10.1016/j.juro.2014.05.124. URL https://www.auajournals.org/doi/full/10.1016/j.juro.2014.05.124. → pages 7, 10, 27, 28, 29, 32, 35, 36, 54, 60, 65, 141

[34] P. Mountney and G.-Z. Yang. Motion Compensated SLAM for Image Guided Surgery. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2010*, Lecture Notes in Computer Science, pages 496–504. Springer, Berlin, Heidelberg, Sept. 2010. ISBN 978-3-642-15744-8 978-3-642-15745-5. doi:10.1007/978-3-642-15745-5_61. URL https://link.springer.com/chapter/10.1007/978-3-642-15745-5_61. → pages 17, 19

[35] F. Mourgues, T. Vieville, V. Falk, and . Coste-Manire. Interactive Guidance by Image Overlay in Robot Assisted Coronary Artery Bypass. In R. E. Ellis and T. M. Peters, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, Lecture Notes in Computer Science, pages

173–181, Berlin, Heidelberg, 2003. Springer. ISBN 978-3-540-39899-8. doi:10.1007/978-3-540-39899-8_22. → page 17

[36] M. Nakamoto, K. Nakada, Y. Sato, K. Konishi, M. Hashizume, and S. Tamura. Intraoperative Magnetic Tracker Calibration Using a Magneto-Optic Hybrid Tracker for 3-D Ultrasound-Based Navigation in Laparoscopic Surgery. *IEEE Transactions on Medical Imaging*, 27(2): 255–270, Feb. 2008. ISSN 0278-0062, 1558-254X. doi:10.1109/TMI.2007.911003. → page 22

[37] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, pages 127–136, Oct. 2011. doi:10.1109/ISMAR.2011.6092378. → page 24

[38] R. A. Newcombe, D. Fox, and S. M. Seitz. DynamicFusion: Reconstruction and Tracking of Non-Rigid Scenes in Real-Time. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 343–352, 2015. URL http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Newcombe_DynamicFusion_Reconstruction_and_2015_CVPR_paper.html. → page 24

[39] S. Parashar, D. Pizarro, and A. Bartoli. Isometric non-rigid shape-from-motion in linear time. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4679–4687, June 2016. doi:10.1109/CVPR.2016.506. → page 19

[40] F. Porpiglia, E. Checcucci, D. Amparore, R. Autorino, A. Piana, A. Bellin, P. Piazzolla, F. Massa, E. Bollito, D. Gned, et al. Augmented-reality robot-assisted radical prostatectomy using hyper-accuracy three-dimensional reconstruction (ha 3d^{TM}) technology: a radiological and pathological study. *BJU international*, 123(5):834–845, 2019. → page 16

[41] G. Samei, O. Goksel, J. Lobo, O. Mohareri, P. Black, R. Rohling, and S. Salcudean. Real-time FEM-based Registration of 3d to 2.5d Transrectal Ultrasound Images. *IEEE Transactions on Medical Imaging*, PP:1–1, Feb. 2018. doi:10.1109/TMI.2018.2810778. → page 24

[42] G. Samei, K. Tsang, C. Kesch, J. Lobo, S. Hor, O. Mohareri, S. Chang, S. L. Goldenberg, P. C. Black, and S. Salcudean. A Partial Augmented Reality System with Live Ultrasound and Registered Preoperative MRI for Guiding

Robot-assisted Radical Prostatectomy. *Medical Image Analysis*, Oct. 2019. ISSN 1361-8415. doi:10.1016/j.media.2019.101588. URL http://www.sciencedirect.com/science/article/pii/S1361841519301288. → page 99

[43] T. Simpfendörfer, M. Baumhauer, M. Müller, C. N. Gutt, H.-P. Meinzer, J. J. Rassweiler, S. Guven, and D. Teber. Augmented reality visualization during laparoscopic radical prostatectomy. *Journal of endourology*, 25(12): 1841–1845, 2011. → page 16

[44] R. Singla, P. Edgcumbe, P. Pratt, C. Nguan, and R. Rohling. Intra-operative ultrasound-based augmented reality guidance for laparoscopic surgery. *Healthcare Technology Letters*, 4(5):204–209, 2017. ISSN 2053-3713. doi:10.1049/htl.2017.0063. → pages 22, 54

[45] D. Stoyanov, M. V. Scarzanella, P. Pratt, and G.-Z. Yang. Real-Time Stereo Reconstruction in Robotically Assisted Minimally Invasive Surgery. In T. Jiang, N. Navab, J. P. W. Pluim, and M. A. Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2010*, Lecture Notes in Computer Science, pages 275–282, Berlin, Heidelberg, 2010. Springer. ISBN 978-3-642-15705-9. doi:10.1007/978-3-642-15705-9_34. → page 19

[46] R. C. Susil, K. Camphausen, P. Choyke, E. R. McVeigh, G. S. Gustafson, H. Ning, R. W. Miller, E. Atalar, C. N. Coleman, and C. Mnard. System for prostate brachytherapy and biopsy in a standard 1.5 T MRI scanner. *Magnetic Resonance in Medicine*, 52(3):683–687, 2004. ISSN 1522-2594. doi:10.1002/mrm.20138. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.20138. → page 6

[47] N. Tan, D. J. A. Margolis, T. D. McClure, A. Thomas, D. S. Finley, R. E. Reiter, J. Huang, and S. S. Raman. Radical prostatectomy: Value of prostate MRI in surgical planning. *Abdominal Imaging*, 37(4):664–674, Aug. 2012. ISSN 1432-0509. doi:10.1007/s00261-011-9805-y. URL https://doi.org/10.1007/s00261-011-9805-y. → page 6

[48] D. Teber, S. Guven, T. Simpfendrfer, M. Baumhauer, E. O. Gven, F. Yencilek, A. S. Gzen, 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, Aug. 2009. ISSN 0302-2838. doi:10.1016/j.eururo.2009.05.017.

URL http://www.sciencedirect.com/science/article/pii/S0302283809005211. → page 22

[49] O. Ukimura, I. S. Gill, M. M. Desai, A. P. Steinberg, M. Kilciler, C. S. Ng, S. C. Abreu, M. Spaliviero, A. P. Ramani, J. H. Kaouk, A. Kawauchi, and T. Miki. Real-time transrectal ultrasonography during laparoscopic radical prostatectomy. *The Journal of Urology*, 172(1):112–118, July 2004. ISSN 0022-5347. doi:10.1097/01.ju.0000128914.21240.c8. → page 7

[50] O. Ukimura, C. Magi-Galluzzi, and I. S. Gill. Real-Time Transrectal Ultrasound Guidance During Laparoscopic Radical Prostatectomy: Impact on Surgical Margins. *The Journal of Urology*, 175(4):1304–1310, Apr. 2006. ISSN 0022-5347. doi:10.1016/S0022-5347(05)00688-9. URL http://www.sciencedirect.com/science/article/pii/S0022534705006889. → page 7

[51] O. Ukimura, M. Aron, M. Nakamoto, S. Shoji, A. L. d. C. Abreu, T. Matsugasumi, A. Berger, M. Desai, and I. S. Gill. Three-dimensional surgical navigation model with tilepro display during robot-assisted radical prostatectomy. *Journal of endourology*, 28(6):625–630, 2014. → page 17

[52] O. Ukimura, M. Aron, M. Nakamoto, S. Shoji, A. L. d. C. Abreu, T. Matsugasumi, A. Berger, M. Desai, and I. S. Gill. Three-dimensional surgical navigation model with TilePro display during robot-assisted radical prostatectomy. *Journal of Endourology*, 28(6):625–630, June 2014. ISSN 1557-900X. doi:10.1089/end.2013.0749. → page 31

[53] B. Vagvolgyi, L.-M. Su, R. Taylor, and G. Hager. Video to CT registration for image overlay on solid organs. *Proc. Augmented Reality in Medical Imaging and Augmented Reality in Computer-Aided Surgery (AMIARCS)*, Jan. 2008. → page 21

[54] G. M. Villeirs and G. O. De Meerleer. Magnetic resonance imaging (MRI) anatomy of the prostate and application of MRI in radiotherapy planning. *European Journal of Radiology*, 63(3):361–368, Sept. 2007. ISSN 0720-048X. doi:10.1016/j.ejrad.2007.06.030. URL http://www.sciencedirect.com/science/article/pii/S0720048X07003269. → page 5

[55] M. Visentini-Scarzanella, T. Sugiura, T. Kaneko, and S. Koto. Deep bonocular 3d reconstruction for assisted navigation in bronchoscopy. *International Journal of Computer Assisted Radiology and Surgery*, 12(7):

1089–1099, July 2017. ISSN 1861-6429. doi:10.1007/s11548-017-1609-2. URL https://doi.org/10.1007/s11548-017-1609-2. → page 20

[56] M. C. Yip, D. G. Lowe, S. E. Salcudean, R. N. Rohling, and C. Y. Nguan. Tissue Tracking and Registration for Image-Guided Surgery. *IEEE Transactions on Medical Imaging*, 31(11):2169–2182, Nov. 2012. ISSN 0278-0062. doi:10.1109/TMI.2012.2212718. → page 22