2-8-2017 12:00 AM

A Framework for Tumor Localization in Robot-Assisted Minimally Invasive Surgery

The University of Western Ontario

Supervisor
Dr. Rajni V. Patel
The University of Western Ontario

Graduate Program in Electrical and Computer Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Engineering Science
© Nikita Chopra 2017

Abstract

Manual palpation of tissue is frequently used in open surgery, e.g., for localization of tumors and buried vessels and for tissue characterization. The overall objective of this work is to explore how tissue palpation can be performed in Robot-Assisted Minimally Invasive Surgery (RAMIS) using laparoscopic instruments conventionally used in RAMIS. This thesis presents a framework where a surgical tool is moved teleoperatively in a manner analogous to the repetitive pressing motion of a finger during manual palpation. We interpret the changes in parameters due to this motion such as the applied force and the resulting indentation depth to accurately determine the variation in tissue stiffness. This approach requires the sensorization of the laparoscopic tool for force sensing. In our work, we have used a da Vinci® needle driver which has been sensorized in our lab at CSTAR for force sensing using Fiber Bragg Grating (FBG). A computer vision algorithm has been developed for 3D surgical tool-tip tracking using the da Vinci®'s stereo endoscope. This enables us to measure changes in surface indentation resulting from pressing the needle driver on the tissue. The proposed palpation framework is based on the hypothesis that the indentation depth is inversely proportional to the tissue stiffness when a constant pressing force is applied. This was validated in a telemanipulated setup using the da Vinci® surgical system with a phantom in which artificial tumors were embedded to represent areas of different stiffnesses. The region with high stiffness representing tumor and region with low stiffness representing healthy tissue showed an average indentation depth change of 5.19 mm and 10.09 mm respectively while maintaining a maximum force of 8N during robot-assisted palpation. These indentation depth variations were then distinguished using

the k-means clustering algorithm to classify groups of low and high stiffnesses. The results were presented in a colour-coded map. The unique feature of this framework is its use of a conventional laparoscopic tool and minimal re-design of the existing da Vinci® surgical setup. Additional work includes a vision-based algorithm for tracking the motion of the tissue surface such as that of the lung resulting from respiratory and cardiac motion. The extracted motion information was analyzed to characterize the lung tissue stiffness based on the lateral strain variations as the surface inflates and deflates.

Acknowledgments

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviations

BMFLC Band Limited Fourier Linear Combiner

BRISK Binary Robust Invariant Scalable Keypoints

CSTAR Canadian Surgical Technologies and Advanced Robotics

CT Computed Tomography

DOF Degrees of Freedom

DVI Digital Visual Interface

EBRT External Beam Radiation Therapy

EKF Extended Kalman Filter

FBG Fiber Bragg Grating

FL Feature Localization

FLC Fourier Linear Combiner

GUI Graphical User Interface

GPU Graphics Processing Unit

HD Hausdorff Distance

HC Hunt and Crossley

IFME Inter Frame Motion Estimation

KF Kalman Filter

KV Kelvin Voigt

LCI Left Camera Input

LCSI Left Camera Split Input

LED Light Emitting Diode

LHSC London Health Science Center

LMS Least Means Square

MAD Mean Absolute Difference

MC Master Console

MI Mutual Information

MIS Minimally Invasive Surgery

MRI Magnetic Resonance Imaging

MTM Master Tool Manipulator

NCC Normalized Cross Correlation

NMSE Normalized Mean Square Error

NRMSE Normalized Root Mean Square Error

NSCLC Non Small Cell Lung Cancer

OSP Open Surgical Procedure

PSC Patient Side Cart

PSM Patient Side Manipulator

RAMIS Robot Assisted Minimally Invasive Surgery

RCI Right Camera Input

RCIP Right Camera Input Processed

RSME Root Mean Square Error

SCLC Small Cell Lung Cancer

SEM Stereo Endoscope Manipulator

SIFT Scale Invariant Feature Transform

SSD Sum of Squared Difference

SURF Speeded Up Robust Features

TRID Tissue Resonator Indenter Device

TSI Tactile Sensing Instrument

US Ultrasound

USB Universal Serial Bus

VATS Video Assisted Thoracoscopic Surgery

VC Vision Cart

Chapter 1

Introduction

According to the world cancer report, lung cancer is one of the most commonly diagnosed cancers among both men and women [1]. Lung cancer is categorized as either Non-Small Cell Lung Cancer (NSCLC) or Small-Cell Lung Cancer (SCLC). SCLC can spread quickly to other parts of the body. It is important to contain the growth of cancer by either removing, ablating or annihilating the cancerous cells. Otherwise it can spread to other parts of the body which can eventually lead to death. Lung cancer is diagnosed using preoperative imaging such as Computed Tomography (CT). The treatment for the lung cancer depends upon the type of cancer, size, and spread of cancer. There are several ways to treat lung cancer that include surgery, radiation therapy, chemotherapy, cryogenic treatment, and brachytherapy. The patient may receive a combination of treatments or a single treatment depending on the severity of the spread of cancer. Surgical options are preferable for early stage cancer that has not spread to other parts of the body. Chemotherapy is considered a preferred solution for the last stage due to high severity and metastasis. Chemotherapy uses drugs and targeted molecular agents to kill cancerous cells. Radiation therapy uses high-energy radiation to shrink cancerous cells and can be differentiated into two techniques: External Beam Radiation Therapy (EBRT) and Brachytherapy. EBRT is the preferred solution if the tumor is present in a very difficult location that is unreachable or too complicated to reach through surgery. Also, if a patient is physically

weak or unable to withstand the trauma of surgery, this serves as an effective therapy. Since EBRT focuses radiation from outside the body, it is necessary to compensate for the motion of tumors due to breathing so as to minimize the harm from the radiation to the surrounding non-malignant tissue. Therefore, various measures are taken to map the tumor location precisely, and the radiation beams are then targeted from an appropriate angle to cover the tumor accurately [2]. Brachytherapy is a new variation of radiation therapy where radioactive seeds or pellets are directly inserted into cancerous cells using hollow needles. Surgical resection is the primary treatment method for patients suffering from early stage NSLC. The resection involves removing cancerous tissue along with a section of healthy tissue to prevent the spread of cancer. Video Assisted Thoracoscopic Surgery (VATS) is a minimally invasive surgical resection performed on lung cancer with visual feedback from a miniaturized endoscope called a thoracoscope [3]. The target lung to be excised is completely deflated at the beginning of surgery to perform excision. Surgeons can use real-time visual feedback and intraoperative Ultrasound (US) for localizing the tumor that has shifted in position due to lung deflation and motion of the heart and the contralateral lung as the patient lies in a lateral recumbent position. The tumorous region along with a margin of healthy tissue is excised, and the excised tumor is sealed in a bag for removal from the keyhole port [4]. VATS is used as a stand-alone procedure or in combination with other procedures for treating lung cancer. Intraoperative US is not very reliable due to the presence of some air in the alveoli that interferes with the reflected US signals [5]. Therefore, it is essential to develop an assistive framework that can help the surgeons to localize the tumors in absence of the ability of the surgeons to touch the lung with their hands. The framework to be developed should also be able to compensate for the shift in tumor location due to the motion of the deflated lung acquired from the contralateral lung and the heart.

1.1 Motivation

In Open Surgical Procedures (OSP) such as thoracotomy [6], prostatectomy [7], the surgeons usually use the fingers to press on the tissue surface with a light pressure for physical diagnosis. Apart from tissue texture assessment, manual palpation also assists in the detection of anatomical landmarks and other abnormal lumps. Over the last few decades, most of the conventional OSPs have moved towards Minimally Invasive Surgery (MIS) due to numerous benefits such as faster healing, less scarring and reduced pain. In MIS, small incisions (5-12mm) are made for passing the surgical tools and an endoscopic camera into the patient's body. These instruments provide surgeons distorted kinaesthetic force feedback due to the friction between the tool and the trocar port and also due to the deformation of the tissue surface. Although MIS is relatively more beneficial, it requires an intensive learning curve and training [8] to adjust to the inverted hand-eye coordination, lack of depth perception due to a limited 2D laparoscopic view and corrupted force feedback from the tools [9] [10]. The shortcoming associated with laparoscopic procedure such as reduced Degrees of Freedom (DOF) from the rigid tools makes it difficult to maneuver the tool to perform suturing and ligation tasks. Also, poor hand-eye coordination due to 2D display was countered with the advent of RAMIS [11]. RAMIS employs the master-slave design that enables the surgeon to remotely maneuver the tools within the patient's body. The surgeon's action on the master side is mimicked by a robotic arm using teleoperative communication. The motion can be amplified or stabilized depending on the application. The intuitive user interface, highly immersive 3D vision and various DOF of the tool with no inversion makes RAMIS a preferred choice. It has been used in a large number of surgical procedures, e.g, see [12–15]. Although the current technology for RAMIS provides several benefits such as better ergonomics, 3D immersive surgical view, dexterity, reduced trauma and speedy recovery for patients, it does not provide the surgeons a sense of touch. Manual palpation enables the surgeons to determine various tissue properties such as shape, size, temperature, orientation and stiffness. The perceived signal after contact with a

surface during OSP is a combination of the input signals from various neuro-receptors. The information obtained from manual palpation can be considered as a combination of multiple inputs. To replicate a similar perception in RAMIS, the framework should be able to analyze multisensory data along with the mechanical haptic information at the master side. Current robot-assisted systems such as the da Vinci® do not have integrated haptic sensing [16]. Also, specifically in minimally invasive thoracic surgery, the target lung is deflated during the procedure for accurate resection. It results in a shift in the location of the tumor as interpreted from the preoperative CT scans. So, it is important for a thoracic surgeon to obtain an intraoperative feedback for the tumor localization. In the next section, we review the state-of-the-art intraoperative tumor localization techniques in the literature to provide an overview of the progress in this field in the context of the proposed work.

1.2   Literature Review

In order to develop a list of challenges and propose solutions for addressing them, we have reviewed the pertinent literature and techniques. Here, we present a literature survey of recent techniques based on their ability to localize tumors intraoperatively and characterize tissue surface properties. The techniques proposed in the literature can be broadly categorized as force-based sensing, tactile-based sensing, aspiration devices and intraoperative medical imaging-based visualization techniques. An ideal tumor localization device should possess the following properties

1. It should be small enough to pass through a small incision, a trocar port of diameter $5mm$ to $12mm$ [17].

2. It should be compatible with the current setup of RAMIS to ease integration and reduce any additional cost of re-design.

3. It should possess fast sensing speed to analyze the stiffness in minimal time.

4. It should be sterilizable or disposable.

5. It should provide accurate repeatable results.

1.2.1 Medical Imaging Techniques

Preoperative imaging technologies such as CT, MRI and US are used for physical diagnosis and tumor identification prior to a surgical procedure. The location of the tumor obtained preoperatively needs to be registered to the intraoperative setup accounting for the changes in viewpoint, orientation, imaging data-type and the change due to tissue deformation. Non-rigid registration of preoperative and intraoperative images due to deflation of the lung during the VATS procedure can be very time-consuming and cannot compensate for real-time motions due to breathing and respiration [18–20]. Deformable tissue models have been developed from preoperative image data using interpolation in [20], and using nonlinear finite-element model in [21]. Preoperative imaging provides high definition multidimensional imaging data whereas intraoperative imaging modalities are often not of comparable quality. US is the most favored intraoperative imaging modality as it is safe, portable and compatible with metals. US uses high frequency $(1 - 10MHz)$ sound waves to view tissue for diagnosis, guidance, and therapy. Radiologists are trained to manipulate US transducers, and read and analyze the data displayed on the screen. Though most of the procedures still use conventional 2D US imaging, there have been several development in 3D US imaging with significant advancement in computing and rendering power. The US probe is inserted through one of the minimally invasive ports into the thoracic cavity for localizing a tumor. US is sensitive to the amount of air present in the lung alveoli even after the lung has been deflated. Therefore, the image acquired using US has artifacts and needs enhancement for tumor localization [22]. Also to improve the survival rates, it may be helpful to localize tumors that are less than $50mm$ in size, and it is tough to localize tumors of smaller size with US. To summarize, US is not the best option for tumor localization intraoperatively during VATS.

1.2.2 Aspiration Devices

These are devices that do not compare the relative stiffness of the sampled points but measure the Young's Modulus of the soft tissue aspired or sucked by the pipette device. The approach works on determining the slope of the stress-strain (Young's modulus) linear variation. The tissue sample is considered to be isotropic and homogenous. This procedure is deemed to be very accurate and hence it is used during a biopsy for diagnosing the presence of a tumor [23]. The basic implementation works by decreasing the pressure by sucking air out of a pipette. Because of this pressure change, there is a change in the level of soft tissue sucked in the pipette. In [24], a light sterilizable aspiration device was developed for *in-vivo* estimation of tissue stiffness. The basic mechanism is the same where the pressure is decreased to measure the rise in the tissue being aspirated. Figure 1.1 shows the mechanism for sensing the rise in the tissue surface. It is achieved by generating a mirrored profile of the tissue surface and then using a recording camera to evaluate the change in depth. The procedure is reliable and has been performed on human liver [25], kidney and spleen [26] and also has been used for ex-vivo prostate cancer diagnosis [27]. In [23], the effects of parameters such as friction between the tissue pipette boundary, the thickness of the pipette, the diameter of the pipette and inhomogeneity of tissue were studied using numerical simulations. Although this process is reliable and measures the absolute Young's modulus for characterization, it is very time consuming to aspirate and release multiple points intraoperatively. Also, its fabrication and implementation are dependent on multiple features such as axial orientation of the pipette during the experiment, the thickness and diameter of the pipette wall [23] .
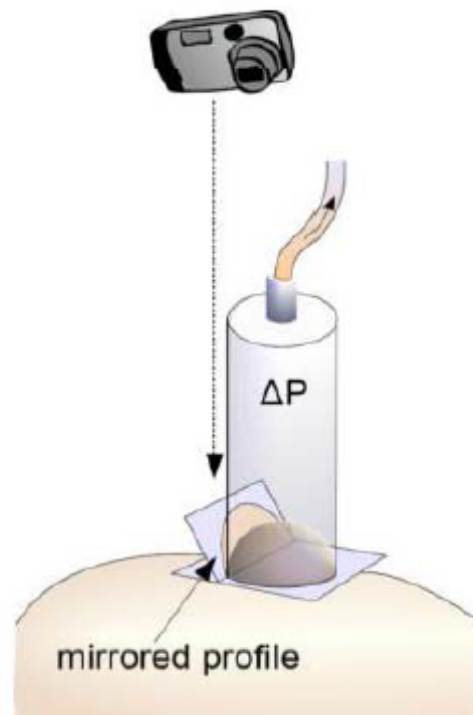
Figure 1.1: Aspiration device designed in [24] where Δp corresponds to the aspiration pressure change and the rise in the tissue level depicted from a mirrored profile with the use of a vision-based analysis [24].

1.2.3 Probing devices.

1.2.3.1 Force-sensorized indentation probing devices

Commercially available sensors are accurate but they are not small enough to be compatible for use in RAMIS. NANO 17 (ATI Industrial Automation Inc., USA) is the smallest commercial sensor, and its diameter is 17mm and therefore, it is not viable for force sensing in RAMIS since the diameter of the trocar port is 5-12 mm. In [28], a hollow strain gauge sensor is placed on the tip of the probe for the force sensing during an arthroscopic surgical procedure. In [29–31], multiple FBG sensors are used to develop a sensorized da Vinci probe for force sensing. FBG sensors work on the variation of the refractive index and the wavelength to determine the force. Unlike strain gauges, FBG sensors are resistant to temperature change and can be

used for a prolonged time.

McCreery *et al.* [32] designed an experiment with a vertical probe on which was mounted a commercial Gamma Sensor (ATI Industrial Automation Inc., USA) at the proximal end for obtaining the kinaesthetic feedback to localize a tumor within an ex-vivo porcine lung as depicted in Figure 1.2. A Mitsubishi PA10-7C robotic arm was programmed to move down in pre-specified increments to a fixed depth, and the force response was recorded at that depth for sampled points. It was the first probing experiment, designed to examine the tissue response by varying parameters such as indentation depth and force applied at variable and constant tool velocity. The design can be directly replicated in RAMIS, and also the work provided a statistical validation that kinaesthetic feedback at controlled depths can characterize tissue stiffness.
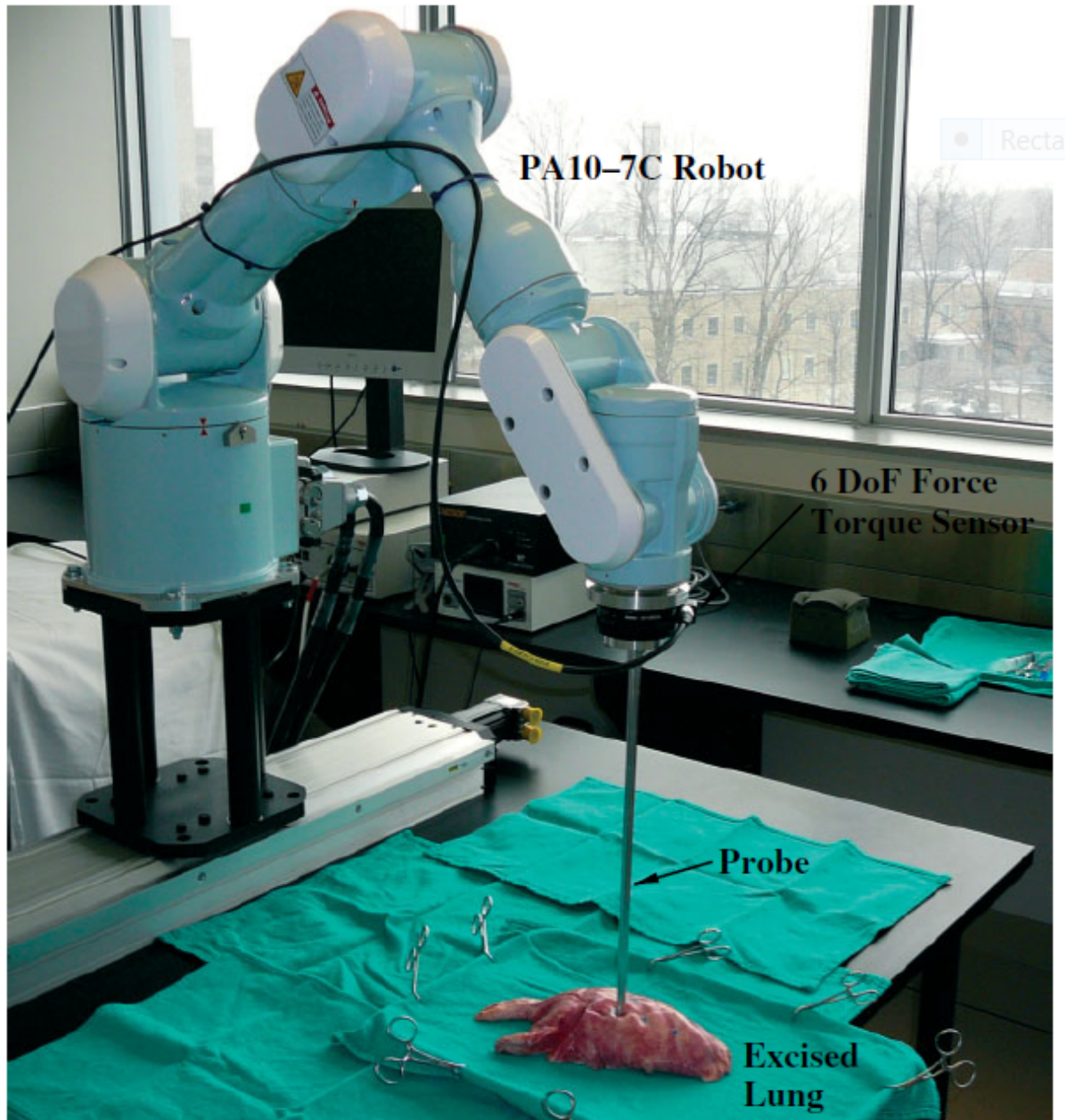
Figure 1.2: Mitsubishi PA10-7C robotic arm performing a probing experiment on sampled points on an ex-vivo porcine lung [32].

In [33], an augmented reality and haptic interface is presented that uses a master-slave prototype setup for robot-assisted palpation. Figure 1.3 shows the experimental testbed design with a stereo camera for 3D point cloud data construction for a visual overlay. It also contains a slave manipulator that works as probing device with an attached commercial NANO-sensor

for force sensing. This work establishes that force sensing and depth sensing together can be used for the characterization of a tissue surface. Yamamoto *et al.* [33] also evaluated the performance of contact models in characterizing tissue property in a similar surgical scenario and reported that the Hunt and Crossley (HC) model best characterizes tissue property by estimating the stiffness and the damping parameters [34]. This work serves as a prototype but it is not designed to be implemented during RAMIS due to the size of the commercial sensor which is too large for a trocar port in RAMIS. Also, the probing arm is a two-arm manipulator which makes it easier to perform position sensing. In an actual RAMIS procedure using the da Vinci® master-slave teleoperation robot, involves a stereo endoscope with a small baseline variable focus and 18 joints between the manipulator and the base.
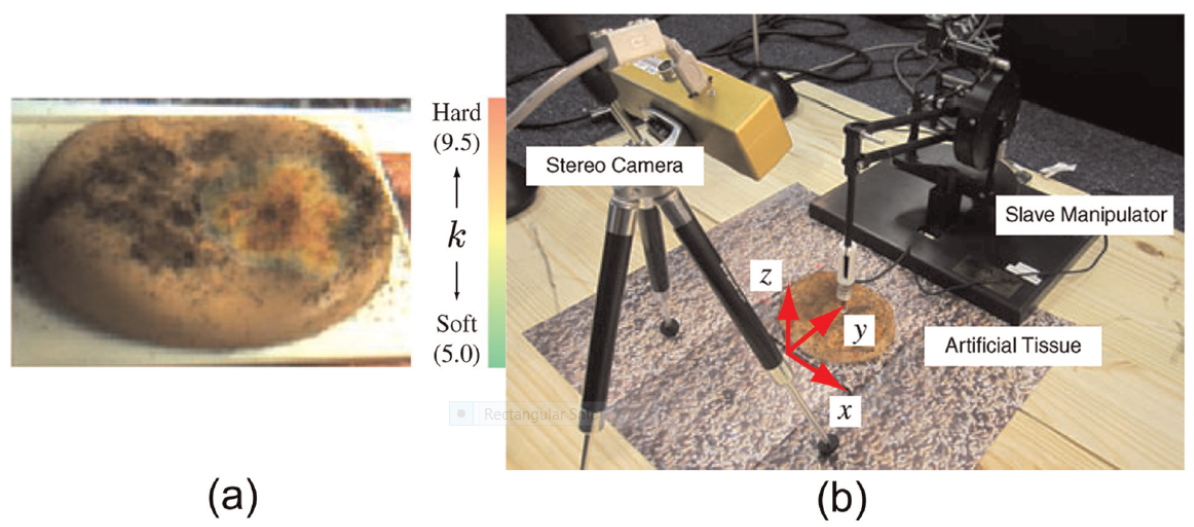
Figure 1.3: A color-coded stiffness map shown in (a) obtained from position sensing and force sensing using a master-slave hardware prototype testbed constructed in (b) [33].

Figure 1.4 shows the design of a portable Tissue Resonator Indenter Device (TRID) that can measure local mechanical properties of soft tissue [35]. The system is designed with two known spring constants, masses, and an indenter tip. The underlying principle is that the device when comes in contact with a surface will show a shift in its natural frequency and increase in the damping ratios. The system uses a linear Kelvin-Voigt (KV) contact dynamics model for

estimating static and dynamic stiffnesses as the probe is pressed on the surface. The probe is not miniaturized to be compatible with MIS, and also its accuracy is dependent on the orientation at which probe is pressed on the surface.
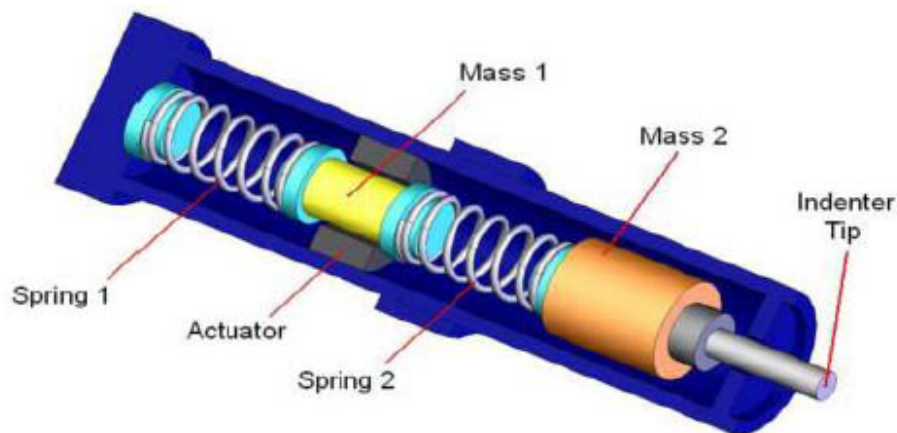
Figure 1.4: An indentation device that measures tissue stiffness based on shift in resonance-frequency [35].

In [36], a multi-axis device with four linear indenters connected to springs with different spring constants is designed that produces a varied displacement on interaction with different surfaces. The device has an installed USB camera that enables measurement of the movement of the indenters as it is pressed on the surface, using image processing to characterize the surface. Figure 1.5 shows the device design and indentation technique. The tool design and mechanism are independent of the orientation with which the surface is pressed. Even the sensing range and resolution of stiffness identification are modifiable, and the design is suitable for diagnosis based on external palpation. The miniaturized version of this tool can be mounted on an endoscopic camera to assist in tissue stiffness characterization during RAMIS.
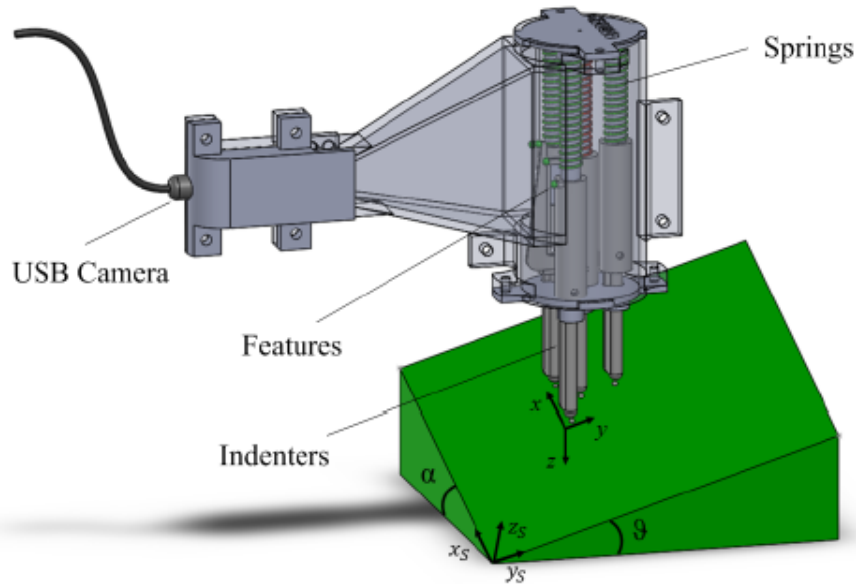
Figure 1.5: A multi-axis stiffness sensing probe pressing on a surface oriented at an angle for mechanical property estimation of the surface [36].

1.2.3.2   Rolling indention probing devices

Another procedure that is different from uni-axial compression is the rolling indentation technique.This technique is analogous to the lateral movement of a finger over a surface. This technique is based on the hypothesis that normal forces will vary for different materials if the indentation depth is kept the same while rolling the tool over the surface. One such work is described in [37] where a rolling indentation probe is designed that can measure the indentation depth and the tissue reaction force at a different orientation. The designed probe shown in Figure 1.6 is not miniaturized to be compatible with MIS but the prototype can construct a stiffness map of the surface as shown in the figure. The arrangement of the fiber force sensor on the tool enables only the measurement of the axial forces. Therefore, the probe was kept perpendicular to the surface for tissue characterization.
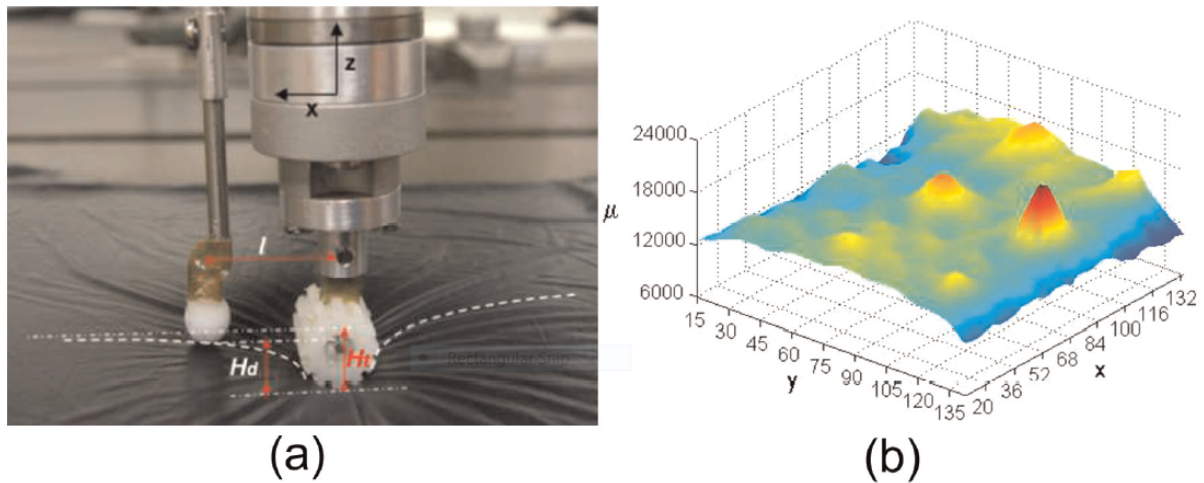
Figure 1.6: Rolling indentation probe design and its motion as it glides over the surface while maintaining the same depth (a) and (b) shows the stiffness map [37].

McKinley *et al.* [38] designed a disposable haptic palpation probe shown in Figure 1.7. It is based on the principle of probe tip deflection mapping as the spherical probe tip is moved over the surface in an autonomous mode. The designed disposable probe of size $15mm \times 10mm$ is mounted on the tip of an $8mm$ diameter da Vinci$^{\circledR}$ needle driver tool. The probe uses a Hall effect sensor that measures the change in electrical potential produced by the magnetic flux. This flux change is encoded as the tip displacement while the tool glides over the surface at a speed of $(0.5 - 21)mm/s$. The deflection of the probe tip attached to a spring with a fixed spring constant determines the force profile along the path. The autonomous palpation routine is designed to maintain a constant indentation depth for the set of experiments so that the force profile can characterize the points. The designed probe is miniaturized but still cannot pass through the trocar. Also, the autonomous setup glides the tool along the phantom while maintaining a constant depth. Such a gliding motion at fixed depth is difficult to perform through a trocar with a fixed RCM. The magnetic flux from the surroundings may reduce or enhance the field that the Hall probe detects, hence rendering the results inaccurate.
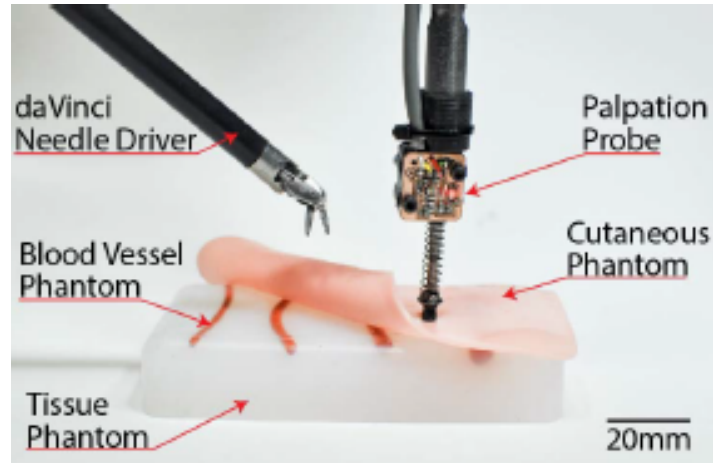
Figure 1.7: Palpation probe mounted at the tip of a da Vinci® needle driver designed to sense the deflection of the tip as the tool moves over the phantom maintaining the same depth [38].

.

1.2.3.3   Tactile sensing.

A tactile sensor provides one or more of the sensing functions available through the human sense of touch such as shape, stiffness, temperature, vibration and normal forces [39]. Tactile sensing is actively used in service robots, surgical procedures, rehabilitation, food processing and robotic automation but in this section, we only discuss the work and advancement related to surgical procedures. In our application, we will consider a tactile sensor to be made up of an array of multiple discrete sensing elements that determine the normal forces or pressure applied during a physical contact. This information is translated into a digital image $P(x, y)$ for the pressure surface map visualization and interpretation. Tactile sensing allows surgeons to analyze a region entirely instead of a point as in the case of force-sensorized probing devices. The sensing technologies employed in designing a tactile sensor are capacitive, piezoresistive, thermosensitive, inductive, piezoelectric, magnetic, elastomeric and optical [39]. The transduction technique is dependent on a parameter that is modulated which should ideally generate a high sensing response and a good spatial resolution. Pressure Profile Systems, Inc. and Tekscan, Inc. produce commercial tactile sensors that show a pressure distribution map over a tissue surface. In a capacitive tactile sensor, the variation in capacitance is indicative of the

pressure. Similarly, parameters such as change in resistance, magnetic coupling, current fluctuation and a shift in resonant frequency are studied for different transduction techniques. The tactile sensor is a layered structure designed with a sensing layer, an electronic layer, a protective layer and a support layer. The sensing elements are connected via wires in case of an electrical component or manufactured as thin pressure sensing films. The ideal properties of a tactile sensor that is designed to be small enough to be compatible with RAMIS are: high sensing range, sterilizable or disposable and also it should show a good resistance to changes in temperature and humidity. The probes in the literature for use in MIS can be broadly categorized as grasping probes, non-grasping probes and those based on catheter tip sensing.

Schostek *et al.* [40] designed a laparoscopic grasper with a substrate material that had 32 spherical electrodes on top of it and a layer of conductive polymer. The change in the contact area between the conductive layer and the electrode due to the application of pressure results in a change in the resistance at that electrode. This resistance change results in a change in electrical current from the electrodes across a constant potential difference. Figure 1.8 shows the design of the laparoscopic grasper and the fabrication layer structure of the designed tactile sensor. This designed grasper is cost efficient, disposable and is able to withstand a high grasping force. Cohn *et al.* [41] constructed a laparoscopic tactile device that resembles the structure of a hand with fingers enabled for tactile sensing using a strain sensor array.
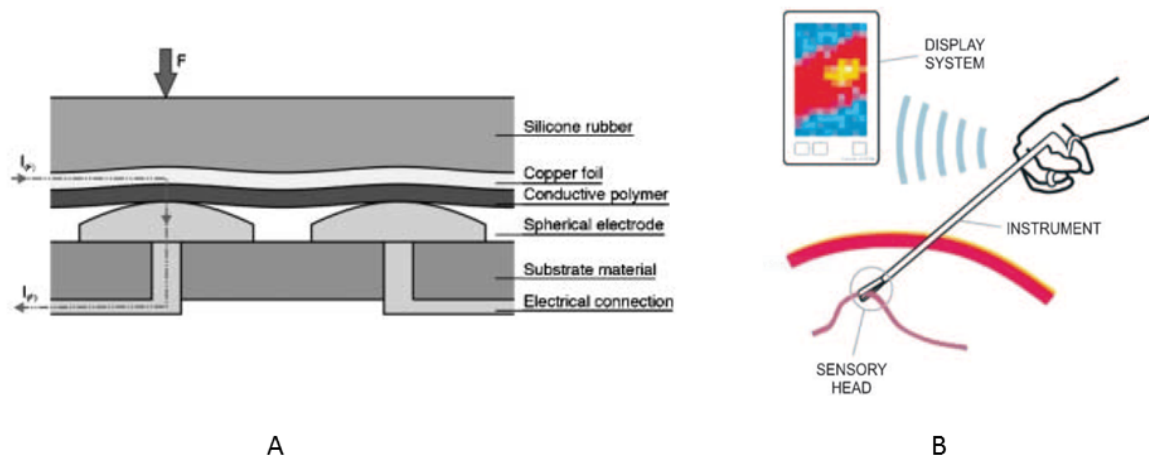
Figure 1.8: Schematic of the grasping tactile sensor design and the digital image display for the pressure map [40].

Trejos *et al.* [42] developed a Tactile Sensing Instrument (TSI) that used an industrial TactArray sensor from Pressure Profile Systems for pressure sensing as the tool is pressed on the surface. This research demonstrated that robot-assisted palpation results in the application of less force and higher accuracy in localizing tumors when compared to manual palpation. Bilateral control is the most basic method used to combine direct force feedback into robot-assisted surgery [43] . Talasaz *et al.* [44] proposed an integration of force feedback with tactile sensing by implementing a hybrid impedance control method to improve tumor localization. Tactile sensors suffer from long-term loading issues, and this can result in inaccuracies in sensor output over time. Tactile sensors also result in inaccuracies in the presence of tissue motion and change in surface properties due to the presence of fluids and blood during a procedure. Naidu *et al.* [45] designed a three DOF, miniaturized, inexpensive, sterilizable tactile sensor that can measure contact pressure with $1kPa$ resolution. The instrument designed is a hybrid of two different sensing modalities, i.e., US sensing on one side and tactile sensing on the other side of the tool. It can be easily switched between US sensing and tactile sensing during MIS palpation, hence increasing the reliability and accuracy of tumor localization. We conclude that incorporating multi-modality sensing in these techniques can result in higher confidence in tu-

mor localization.

Intraoperative imaging, kinaesthetic feedback, and tactile sensing are the categories of feed-back that can be made available to surgeons during RAMIS. Visual feedback from an intraop-erative imaging modality such as the US is distorted due to the presence of air in the alveoli of a lung, and hence cannot be used by itself to localize tumors accurately. We have reviewed state-of-the-art probing techniques for intraoperative tumor localization, their design princi-ples, limitations, functioning and integration challenges in current surgical setups. Most of the work in the literature is in the development phase where the proposed design is not small enough for use in RAMIS. A few RAMIS compatible probes that are suggested in the literature depend on a rolling indentation mechanism in which a gliding motion is performed at a fixed depth. It is challenging to perform such a motion at a fixed depth through a small port and a fixed RCM. Therefore, we can conclude from our extensive literature review that the palpation procedure should mimic manual palpation and the probe design should be RAMIS compatible. Since we are working with the da Vinci® surgical system, the probe should have a functionality similar to that of a da Vinci® endowrist tool, i.e., it should have 7 DOF and should also ex-hibit properties such as motion scaling, tremor reduction, and intuitive finger tip-control. The major limitations for the probing devices are their speed of sensing and their accuracy. There-fore, if the framework is accompanied by a probabilistic model which can highlight the region that is more likely to contain a tumor, it can significantly improve the computational time and confidence in tumor localization.

1.3   Contributions

The main contributions of this thesis are related to the design and development of a teleoper-ated framework that mimics the manual palpation in a master-slave experimental setup. The framework combines the sensing capabilities of an FBG sensorized da Vinci® tool and a vision-based depth sensing of a tool-tip using a miniaturized da Vinci® stereo endoscope.

We developed a vision-based tool-tip tracking algorithm that can assist a surgeon during a robot-assisted palpation in determining the 3D position of the tool-tip as it is moved teleoperatively. The 3D tracking algorithm is designed to work and adapt with minimal upfront information of the initial parameters, and later the algorithm adapts on-the-fly to the scale and rotation changes in the tool-tip template as the tool moves closer or farther from the camera or rotates during robot-assisted palpation.

The framework design works on the principle of decoding the stiffness property based on the variation in indentation depth as the force-sensorized tool presses on the phantom surface minimally invasively. The user has a continuous force-feedback through a visual cue while maneuvering the tool-tip teleoperatively that enables him to press on the phantom surface with the same amount of force for a set of sampled points. This feature eliminates the variability associated with pressing hard or pressing softly that can generate inconsistent stiffness comparisons. We validated that with a constant amount of pressing force the indentation depth is sufficient to characterize tissue stiffness.

The analysis is followed by the segmentation of points with lower relative stiffness using the unsupervised K-means clustering algorithm. One of the major highlights of the framework is its ease of integration in the current robot-assisted minimally invasive surgical setup because the vision-based tracking is implemented using the inbuilt da Vinci® stereo endoscope, and a force-sensorized endowrist da Vinci® tool is used for palpation which is commonly used in surgical tasks such as suturing. This ensures a minimal redesign for the current robotic surgical setup, and the tool being a standard da Vinci® tool is small enough to pass through a conventional trocar.

We also developed a vision-based algorithm for lung surface tracking that incorporates the presence of a distinctive patterned biological marker on a lung surface, and the quasi-periodic nature of its motion. An actual surgical scene would involve lung tissue surface that has some motion even after it has been deflated. This quasi-periodic motion results from the motion of the contralateral lung and the heart. The integration of surface tracking with the framework

presented in Chapter 2, will enable us to compute the indentation relative to the surface motion. Also, the repetitive pressing motion can be performed either at the rising part of the breathing cycle or at the falling part of the breathing cycle to eliminate the source of variability in the results due to the presence of motion in the palpated lung surface. Another significant contribution from tracking the lung surface is that we were able to validate the hypothesis that a tumorous surface will stretch less than non-tumorous tissue surface due to a continuous change in the volume of the lung. This hypothesis can be extensively used to develop a probabilistic model that can characterize tissue stiffness based on the inherent motion of the lung surface. The development of this probabilistic model and its integration in the framework developed in Chapter 2 will make this work a suitable candidate for testing it in an *in-vivo* robot-assisted palpation setup.

1.4 Thesis Overview

In this Section, we present a brief overview of the four chapters of the thesis. The first chapter illustrates the challenges associated with RAMIS and state-of-the-art solutions suggested in the literature. It provides an overview of the various mechanisms developed over time and their functionalities. In Chapter 2 and Chapter 3, we present the major work of this thesis, i.e., robot-assisted palpation and motion tracking of a lung surface to characterize tissue stiffness. The last chapter summarizes the thesis work with concluding remarks and possible future work.

1.4.1 Introduction

This chapter contains the introduction and the motivation behind pursuing this research work. We briefly discuss the current scenario of lung cancer treatments and several benefits, challenges introduced due to the advent of RAMIS. To address one such significant challenge, i.e.,

the lack of a sense of touch, we present a review of the existent solutions proposed in the literature, their advantages and limitations. Based on the analysis of these works, we inferred the preferred characteristics of a robot-assisted palpation framework design and developed these further in subsequent chapters. This chapter also presents the thesis outline containing the novel contributions of this thesis and an overview of the entire thesis.

1.4.2 Tissue Stiffness Characterization Based on Robot-Assisted Palpation

This chapter presents the theory and design principle behind the framework that can characterize the stiffness of a set of points palpated teleoperatively. The principle behind the robot-assisted palpation is to analyze the parameters that vary while pressing on a tissue surface such as force and indentation depth. The combined analysis of these parameters depends on the synchronized acquisition of force and digital visual stereo data. Therefore, in this chapter we discuss the system architecture explaining the interfacing of data acquisition devices and the software implementation and the Graphical User Interface (GUI) design. We have used an FBG sensorized da Vinci® tool with a resolution of $0.05N$ for force sensing as the tool presses on the phantom surface. For the purpose of measurement of tissue indentation depth, we have designed a computer vision-based 3D tracking algorithm utilizing the data from da Vinci®'s stereo endoscope. We show that the depth can characterize the stiffness while controlling the amount of force applied by a force-sensorized tool. We also present a segmentation approach using the K-means clustering algorithm for color coding the points with lower stiffness. The chapter concludes by presenting the results of a validation experiment followed by remarks on the performance of the robot-assisted palpation framework.

1.4.3 Tissue Stiffness Characterization Based on Lung Surface Tracking

In this chapter, a vision-based tracking method for computing the 2D motion of multiple distinctive landmarks on the surface of a lung is presented. In Chapter 2, we propose a framework that can be easily integrated into the current da Vinci® surgical setup. The framework's design principle and the working mechanism are dependent on evaluating the indentation depth while pressing on the tissue surface. In an actual surgical procedure, the target lung acquires motion from the contralateral lung and the heart. Therefore, to make the framework compatible with surgical procedure, we need to extend its capability to track the tissue surface motion as well as the tool-tip motion as it presses on the tissue surface in a telemanipulated manner. The motion of a deflated lung surface is quasi-periodic in nature because it is acquired from the low-frequency periodic motion of the contralateral lung and the high frequency periodic motion of the heart. Therefore, the motion signal for the lung surface is modeled as a truncated Fourier series with unknown amplitudes and a defined bandwidth corresponding to the motion of the contralateral lung and the heart. We used a Kalman filter for the adaptive computation of unknown weights and the prediction of the motion in the next frame. The prediction results in more accurate initialization and determination of the search space, and hence improves the robustness and speed of our tracking algorithm. The designed algorithm is implemented on videos obtained from an *in-vivo* VATS. The quasi-periodic change in the lung's volume is indicative of a lateral strain on its surface. We measured this strain variation by sampling pairs of templates along a scan line and measuring the relative motion of those templates using the tracking algorithm. An *in-vivo* VATS video of a lung surface with a tumor was used to validate the hypothesis that the lateral strain variation on a malignant surface is damped relative to the strain variation for points sampled on the non-malignant surface. This hypothesis has the potential to be formulated as a probabilistic model for detecting the region with tumor.

1.4.4 Conclusions and Future Work

This chapter summarizes all the work presented in the thesis and proposes various extensions to the current framework as possible future work, the major future work being the integration of the work presented in Chapters 2 and 3 and validation of the outcome.

Chapter 2

Tissue Stiffness Characterization Based on Robot-Assisted Palpation

2.1 Introduction

This chapter presents an approach for analyzing tissue stiffness property by the integration of force sensing using a RAMIS-compatible da Vinci® tool which is sensorized for force sensing using FBG sensors [29] and vision-based tool-tip tracking using a da Vinci® stereo endoscope manipulator (SEM). The framework design works on the same principle as manual palpation where a finger presses on the tissue surface for assessment and the mechanoreceptors present on the fingertips enables the surgeon to sense tissue texture and also assist in characterizing any abnormal lump's presence [46]. The characterizing parameters for tissue stiffness estimation are indentation depth, rate of depth change and force applied. The FBG sensorized da Vinci® needle driver tool provides the real-time force feedback to the surgeon during a tool-tissue interaction. This functionality can be employed to reduce a variable force application while palpating the sampled points and also to avoid the application of damaging force on delicate tissue surface. Since, we can control the amount of force applied while pressing on tissue,

we show that the indentation depth variation alone is sufficient to characterize tissue stiffness. An experiment is designed to mimic the finger's pressing motion using a teleoperated force sensorized tool that repetitively presses on a phantom surface. The proposed framework studies the tool-tissue interaction at a set of sampled points on a phantom by analyzing the depth change upon indentation while maintaining a constant force.

This framework is designed as a pipeline as shown in Figure 2.1. It consists of three stages: data acquisition, data abstraction, and data interpretation. The pipeline's first element performs a synchronized data acquisition of left frame, right frame and force values. In section 2.2, we explain the first element, i.e. data acquisition which is a system integration step that combines the architecture and buffering rates of various devices with data acquisition software design. The major challenge in this stage is synchronization of the acquired data from hardware with accurate real-time computations. Therefore, this section gives a detailed description of all hardware used, and its deployment as well as the software implementation of threading techniques and the design of the user interface for data acquisition. These acquired values are then processed to determine the location of the tool-tip in the world coordinates as robot-assisted teleoperated palpation is conducted. To be able to differentiate the depth change at points with different stiffnesses, we need an algorithm that computes the depth change of the tool while pressing to sub-millimeter accuracy. The aim of pipeline's second element is to design an accurate, robust and tool agnostic localization method. In Section 2.3 we discuss the second step of the pipeline, i.e., data abstraction. In this Section, we present a detailed explanation of the vision-based algorithm designed to localize the da Vinci® tool-tip to a subpixel accuracy as it presses on the phantom. It also contains the design and results of a validation experiment performed to report the accuracy in depth measurement. In Section 2.4 we present the description of experimental setup to perform robot-assisted palpation telemanipulatively using a master-slave robotic setup. Pipeline's final element involves data interpretation and grouping techniques to analyze the points that respond differently. In Section 2.5, the final element of pipeline, i.e., data interpretation step that describes the machine learning algorithm used for

data characterization. This whole framework can serve as an assistive probabilistic model that determines the points that behave differently. After multiple iterations, the points are classified into two groups, differentiated based on their stiffnesses. In Section 2.6, we present the results obtained from the validation experiment and the conclusions.
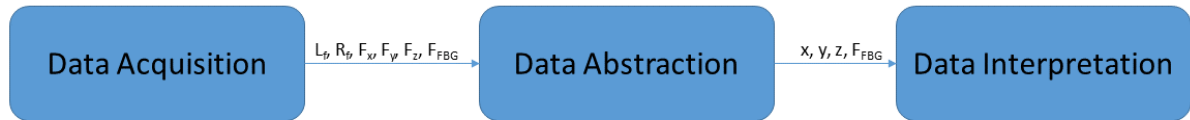
Figure 2.1: Overall framework pipeline.

2.2 Data Acquisition

2.2.1 System Integration

This experiment relies on the integration of various devices such as the robotic platform, the classic da Vinci® running under the daVinci® Reseach Kit (dVRK) which is used to create motions analogous to the pressing motion of a fingertip during manual palpation. Also, a FBG sensorized da Vinci® instrument is used to which are attached four FBG sensors arranged in a pattern for lateral force sensing [29]. An optical interrogator (Micron Optics; Atlanta, US) is used for data buffering and encoding wavelength shifts for all four FBG sensors. For the purpose of simultaneous digital data acquisition from the stereo camera we have used a pair of external frame grabber manufactured by Epiphan, Ottawa, Canada. Figure 2.2 shows the hardware architecture and the flow of data during the experiment. This section focuses on data flow, device integration, and data acquisition.
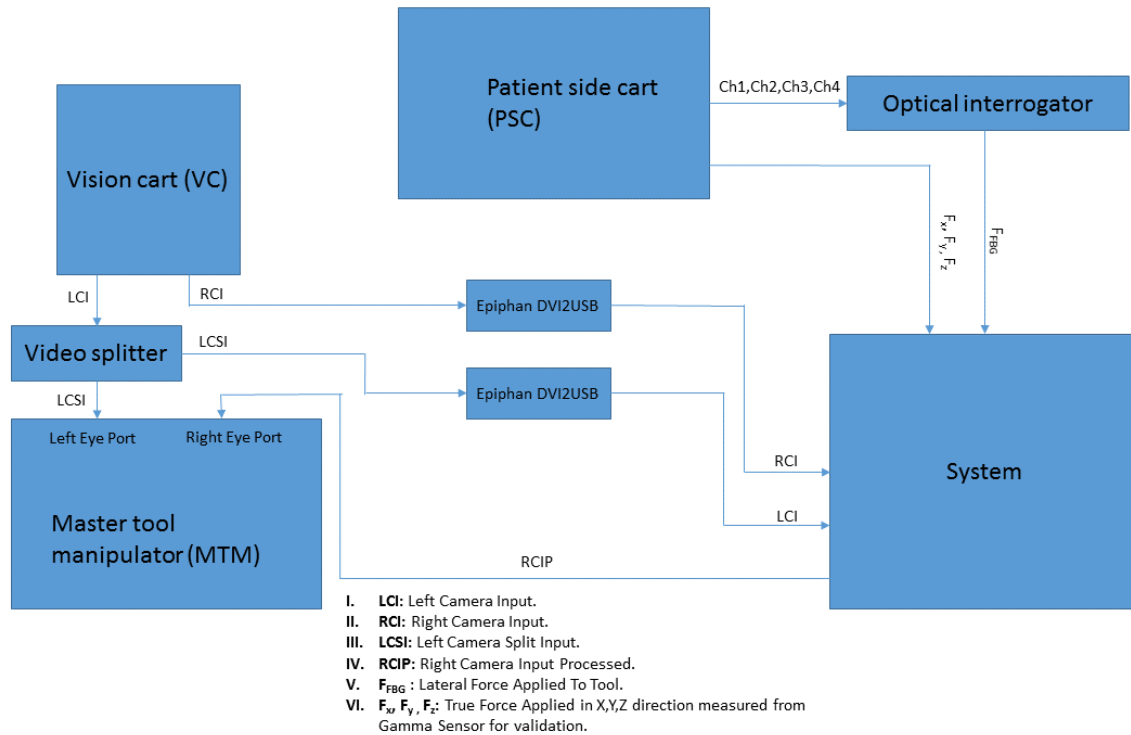
Figure 2.2: Deployment Diagram.

2.2.1.1 Hardware Architecture

Robot-assisted teleoperated surgical procedures performed using the da Vinci® surgical robot (Intuitive Surgical, Sunnyvale, CA) mainly depend on the feedback obtained from stereo vision for navigation, surgical resection, and other vision-based operations. As we have discussed in the previous chapter, although MIS procedures have helped surgeons attain higher degrees of freedom, precision, and accessibility, they have removed the sense of touch. daVinci® is a master-slave robotic system consisting of a Master Console (MC) and a Patient Side Cart (PSC) that enables surgeons to perform surgical operations teleoperatively. The MC design consists of a pair of Master Tool Manipulators (MTM), a stereoscopic head rest and a set of foot-controlled switches (clutch, camera control, camera focus, etc.). MTM allows surgeons to control the patient side robotic arms while getting continuous visual feedback from a binocular head rest that provides a 3D view of the surgical site. This ergonomically designed console

allows a surgeon to manipulate the tools with high dexterity and with appropriate scaling and control. On the PSC are mounted three Patient Side Manipulators (PSM) that consist of the left and the right robotic arms and a SEM. The SEM is a thin tubular streo endoscope with two optical channels connected to a pair of cameras. It also contains a circular light source at its circumference. It provides continuous visual feedback to the surgeon at the head rest display while the PSM mimics the surgeon's motion. The separation between the two cameras gives an immersive 3D view of the operation site enabling the surgeon to analyze the tool's orientation and navigation. Figures 2.3 and 2.5 obtained from [47] show the components of MC and PSC respectively. Figure 2.4 shows the binocular view displayed on the head rest to the surgeon for an immersive 3D view and the MTM design and its manipulation with finger's motion.
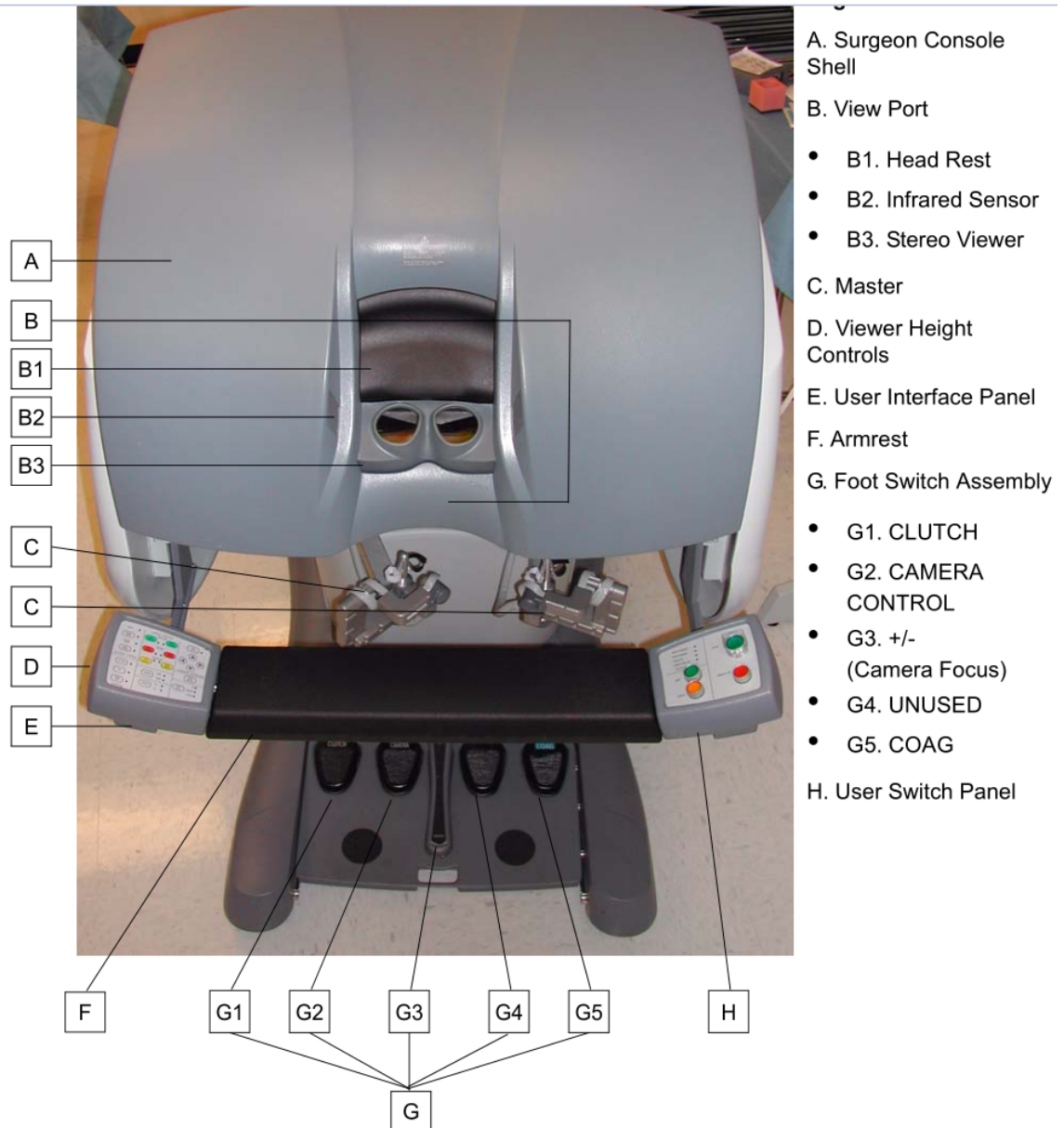
Figure 2.3: Components of the da Vinci® MC [47].

(a) Binocular head rest on MC [47].
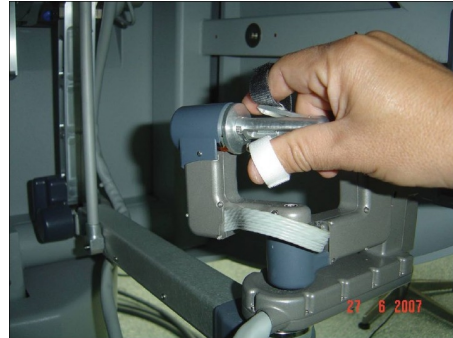
(b) Master tool manipulator [47].

Figure 2.4: (a) Binocular head rest update the surgeon with visual feedback while manipulating; (b) MTM to maneuver the robotic arms on PSC with appropriate scaling and control [47].
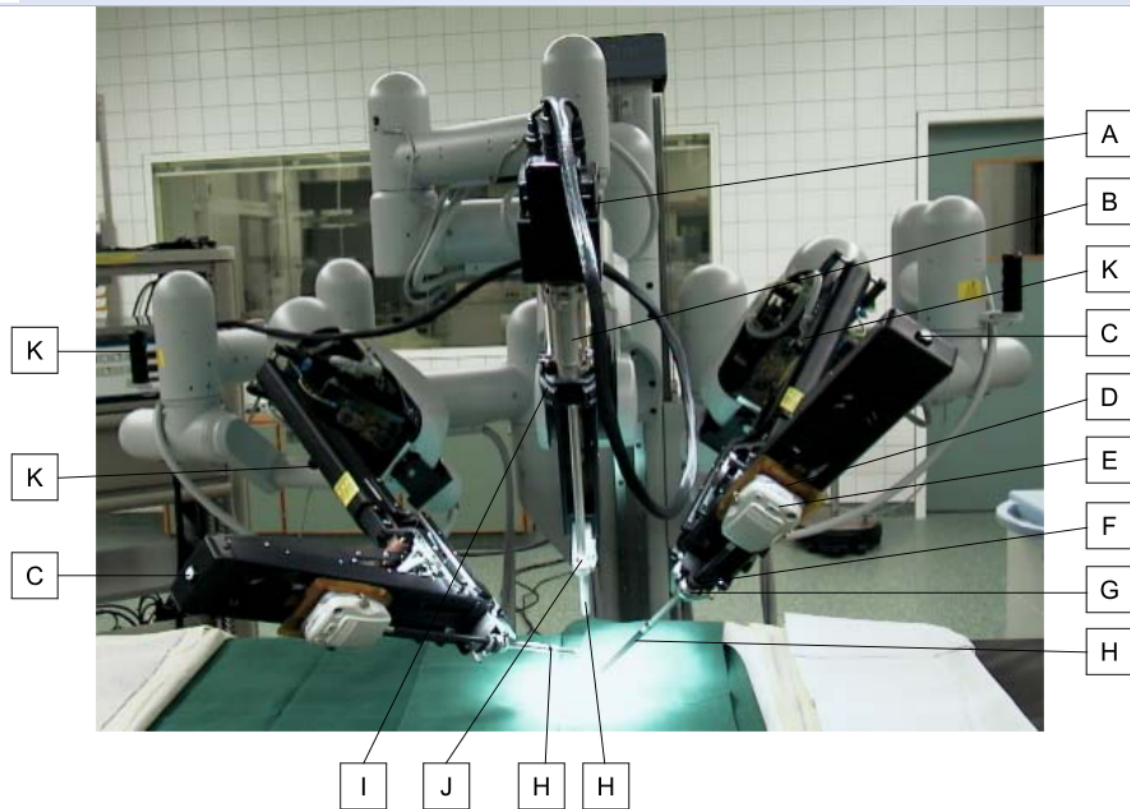
A. Camera Head (attached to Camera Arm)   G. Instrument Arm Cannula

B. Endoscope (Scope)   H. Remote Center

C. Instrument Arm Clutch Button   I. Camera Arm Sterile Adapter

D. Instrument Arm Sterile Adapter   J. Camera Arm Cannula Mount

E. Intuitive Surgical Instrument   K. Set-up Joint Release

F. Instrument Arm Cannula Mount

Figure 2.5: Components of the da Vinci® PSC [47].

2.2.1.1.1 Digital Visual Data Acquisition Hardware Forward kinematics can localize the tool tip of the end effector by retrieving the value from the encoders at the joints, but the value obtained by these encoders may have an error which accumulates when the forward kinematics are computed relating the base to the end-effector. Therefore, the value obtained from the kinematics calculation could have a high error as reported in [34]. Therefore, for proper evaluation, it is necessary to localize the tool-tip in the field of view of the camera using vision-based tool tracking. The frame grabber used for the purpose of acquiring the

digital visual signal from the left and right camera is Epiphan DVI2USB (Epiphan Systems, Ottawa, Canada). Epiphan is a self-powered external frame grabber that captures the high-resolution digital visual data at 60 Hz. As shown in Figure 2.2, input to the left monitor is split into two copies using a video splitter, so the same input is passed to the left screen of the binocular head rest and the computer system for processing. The right eye, on the other hand, is directly input to the computer system for processing. The processed right frame data containing superimposed force value at that instant is input to the right screen of the binocular head rest. Till now we have discussed the flow of digital visual data during the experiment and now we present the details of acquired digital visual data and its rectification process. We use the SEM on PSC for data acquisition from the actual surgical site, this stereo endoscopic camera has a resolution of 640×480 and a sampling rate of 30 Hz. The surgical scene is first captured using the SEM and this captured data requires correction for radial and tangential distortion and also stereo rectification. Stereo rectification is the process in which a stereo image-pair is rotated onto the same plane for correct triangulation of the localized tool-tip in both frames; this results in accurate depth estimation. We explain the procedure used for stereo calibration of the SEM in detail in the next paragraph.

Stereo Calibration and Coordinate Transformation   The SEM has multiple focus settings which can be altered using the $+/-$(G3) foot controlled focus switches on the MC. The focus of the SEM can also be externally controlled using a pair of switches present on the Vision Cart (VC) that changes the focus in and out of SEM. The variability in focus creates a problem while conducting the experiment as a change in focus also changes the field of view, lens radial, and tangential distortion parameters as well as the principal point. The problem associated with focus change and calibration parameters variability can be resolved by creating a data-set of parameters at all eight discrete focus settings as suggested in this work [48]. Therefore, the calibration parameters corresponding to the instantaneous value of focus setting can be retrieved during the experiment. But for the purpose of consistency and com-

putational efficiency, instead of creating a separate data-set, we have kept the focus setting

constant throughout the experimental procedure. The SEM is calibrated using the standard

stereo camera calibration techniques illustrated in [49]. The explanation for the 3D stereo cali-

bration and the transformation equations are adopted from [50]. The stereo calibration process

involves empirically determining intrinsic, extrinsic and distortion parameters of each camera

separately as well as finding the transformation matrix of one camera's positioning relative

to the other camera. The relative orientation of both cameras should remain constant during

the experiment and since the SEM is a rigid endoscope with a fixed baseline between the two

cameras, this constraint is satisfied. Stereo calibration involves the use of a calibration grid

to compute the radial, tangential distortion and spatial transformation parameter, so we have

used a flat 7×5 chessboard for this purpose as it is easy to determine the corner position using

OpenCV library functions [50]. First, each camera is separately calibrated using the OpenCV

function (cv::calibrateCamera) for the computation of intrinsic and extrinsic parameters from

multiple non-planar views of chessboard spanning the whole working space. The transforma-

tion from 3D space to image space and radial and tangential distortion correction is defined

using the Equations (2.1) and (2.2) respectively.

$$s \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic parameters}} \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} & T_1 \\ R_{21} & R_{22} & R_{23} & T_2 \\ R_{31} & R_{32} & R_{33} & T_3 \end{bmatrix}}_{\text{Extrinsic parameters}} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (2.1)$$

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \underbrace{(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)}_{\text{Radial distortion correction}} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \underbrace{\begin{bmatrix} 2\rho_1 x_i y_i + \rho_2 (r^2 + 2x_i^2) \\ 2\rho_2 x_i y_i + \rho_1 (r^2 + 2y_i^2) \end{bmatrix}}_{\text{Tangential distortion correction}} \qquad (2.2)$$

Figure 2.6 shows the stereo vision setup illustrating the 3D point's projection on the left and

right camera planes. The spatial position of left and right cameras in 3D world coordinates can

be computed using the OpenCV function (cv::stereoCalibrate). The rotation R and translation T matrices corresponding to the whole stereo system can be computed in terms of R_l, T_l and R_r, T_r using Equation 2.4. The algorithm developed by Zhang *et al.* [49] uses a closed-form solution that is followed by nonlinear refinement using the Levenberg-Marquardt algorithm for the computation of the transformation parameters.
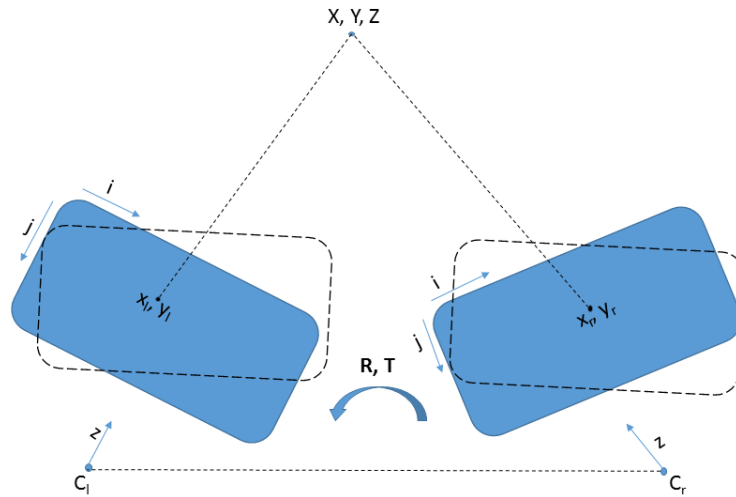
Figure 2.6: Typical stereo vision setup with two cameras and acquired images for the same scene. 3D point (X, Y, Z) is represented as (x_l, y_l) and (x_r, y_r) in stereo images.

$$\begin{bmatrix} x_l \\ y_l \end{bmatrix} = R_l \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + T_l, \quad \begin{bmatrix} x_r \\ y_r \end{bmatrix} = R_r \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + T_r \tag{2.3}$$

$$R = R_r R_l^T, T = T_r + RT_l \tag{2.4}$$

Stereo rectification means to align both left and right images onto a common plane as illustrated in Figure 2.6 with the dotted rectified image pair moved onto the same plane so as that the

epipolar lines are parallel to the horizontal axis. Triangulation of the two corresponding points in a pair of stereo rectified image results in depth estimation of the point. After the images are corrected to remove the presence of radial and tangential distortion, the image can be remapped onto a common plane using a linear transformation. A direct mapping from the left image coordinate frame to the right image coordinate frame can be expressed using the fundamental matrix F. The matrix that relates the shift in the corresponding points to the depth is denoted by Q. These computed calibration parameters are later used in Section 2.3.1.3 for distortion correction and stereo rectification of the acquired stereo images.

2.2.1.1.2 Optical Force Sensing Hardware FBG sensors are well suited for use in MIS instruments due to their high sensitivity, bio-compatibility and miniaturized size. We have used an FBG-sensorized da Vinci®, large needle grasper instrument that has been developed at Canadian Surgical Technologies, and Advanced Robotics (CSTAR) lab for lateral force estimation during RAMIS [29]. The sensorized tool has high resolution and sensitivity, and is designed to measure tool-tissue interaction force during RAMIS. The configurational design of the four FBG sensors mounted on a daVinci® needle driver, are explained in [29]. The wavelength shift for all four FBG sensors attached to the sensorized tool is encoded into lateral force value based on a mathematical force-strain model. The sensors are placed in such a manner that they enable location agnostic force measurements while compensating for axial and torsional disturbances. The developed tool has a force sensing resolution of $0.05N$ at 1KHz, it is because of this fine resolution that we have chosen force as the variable whose maximum value is controlled during the experiment. The superimposed force values are fed to the user's right screen on the head rest while performing the experiment in real-time. This real-time force feedback helps the user in pressing the sampled points with the same maximum force applied on all the sampled points. Later, in this chapter, we show that if the maximum force applied to each point can be controlled, then the maximum depth change becomes the characterizing parameter to indicate the stiffness variation. We have also used a commercial force-torque sen-

sor (Gamma F/T, ATI, North Carolina, US) for validating the best orientation of the tool while

palpating for lateral force estimation.

Figure 2.7: (a) FBG sensorized tool with four FBG sensors for lateral force sensing [29]. (b) Micron Optic's optical interrogator with four channels for the four FBG sensors.

2.2.1.2   Software Framework

The most challenging aspect of data acquisition from various hardware devices is to correlate

the data with respect to time. Error in this can result in inaccuracies in all computational tasks

performed later on this data. The timer is a periodic digital signal that controls the point in

time where the next acquisition cycle starts. To completely understand the acquisition and

synchronization process, it is essential to know the architecture and data buffering mechanism

of the devices. The data acquisition mainly addresses the process of capturing synchronized

data. The data flow and hardware connections are shown in Figure 2.2. For the purpose of data

acquisition, synchronization, and processing of the right frame, we have used the Qt Library for

its efficient threading implementation and support for OpenGL for easier GUI design. The GUI

is devised in a simple manner with control tabs such as " Start Capturing", "Stop Capturing"

and also a split screen for left and right camera eye output. Figure 2.8 shows a snapshot of the

GUI illustrating the design and control tabs.

Figure 2.8: Graphical User Interface.

Figure 2.9 shows the flowchart for the data acquisition software. The main challenge is acquisition and synchronization of the left and right digital data from their respective frame grabbers. As mentioned, Qt has inbuilt functions for the active thread and mutex implementation to manage the access to shared resource among contending threads. We have used two threads for left and right digital data acquire and used a mutex for serialization of these video acquire threads. As soon as the GUI is initialized, a pair of acquisition threads and the timer are constructed. The GUI initialization is the main thread that constructs the timer and the two acquisition threads and after this process threads are present in the memory. When the user clicks the control tab "Start Capturing", the thread initialization method is called, and the threads are ready for execution. As soon as the GUI is initialized, the GUI thread is initialized that evoke the video threads. Both video acquisition threads wait until their state is set to "running" state to begin a new acquisition. The whole process is repeatedly interrupted by the timer after every $30ms$ to trigger a new video capture thread because a new frame is available in the Epiphan buffer after every $30ms$. This process runs until the user clicks the control tab "Stop Capturing" to terminate the acquisition process. The flowchart for the execution of the video capture thread is shown in Figure 2.10.
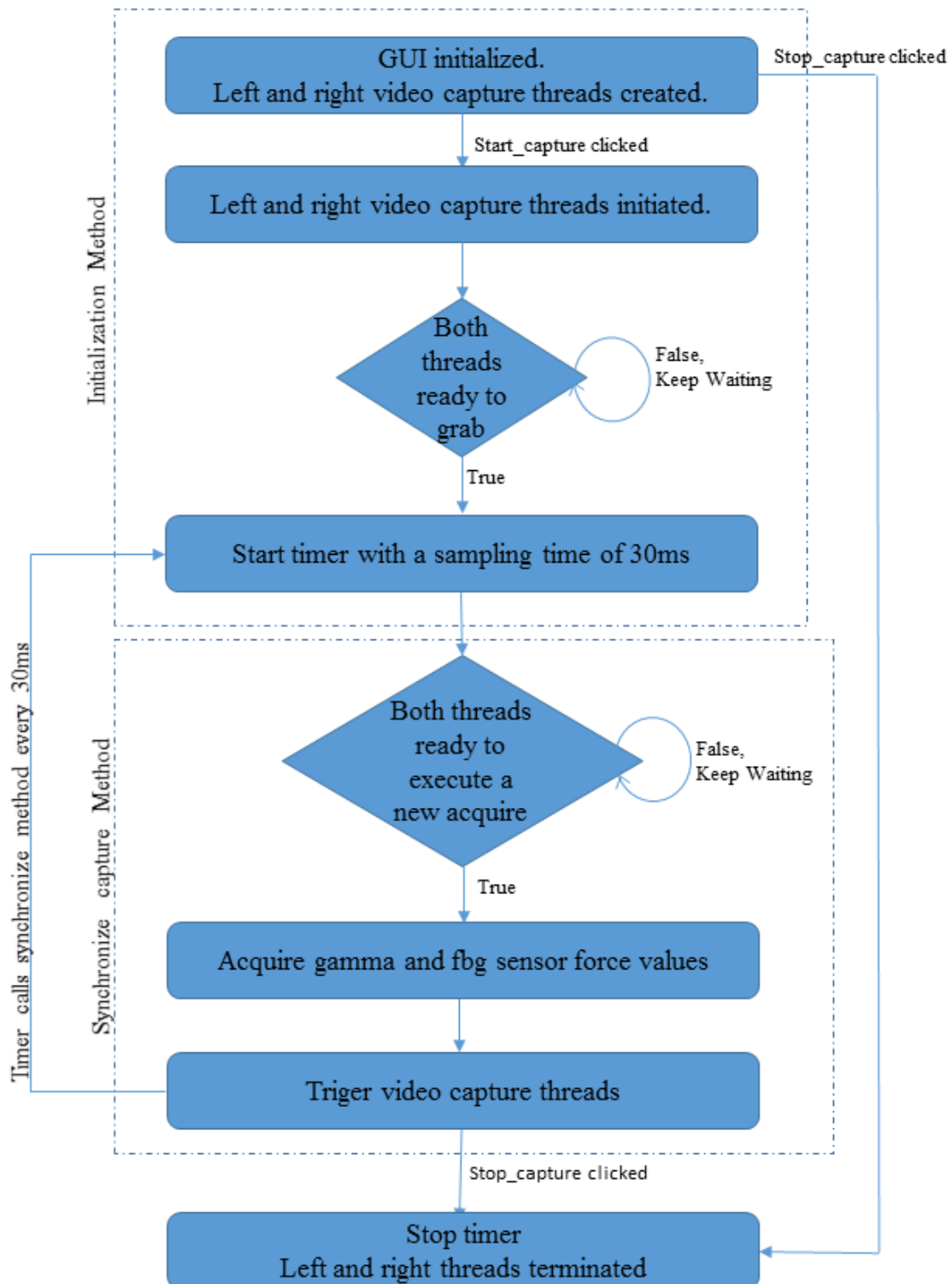
Figure 2.9: Software acquisition flowchart explaining the time-line of the function execution.

The video acquisition threads are present in initialized, ready, running, waiting, blocked and terminated states. Figure 2.11 shows the various states of video capture threads and it also gives a description of an action executed for the transfer from one state to another. These state transitions occur due to mutex locking and unlocking. Mutexes are used in serializing thread executions in case they access any global resource. Figure 2.11 shows different states of the thread and all the mutexes used in our implementation for synchronization and conflict resolution while accessing common resources in the memory.
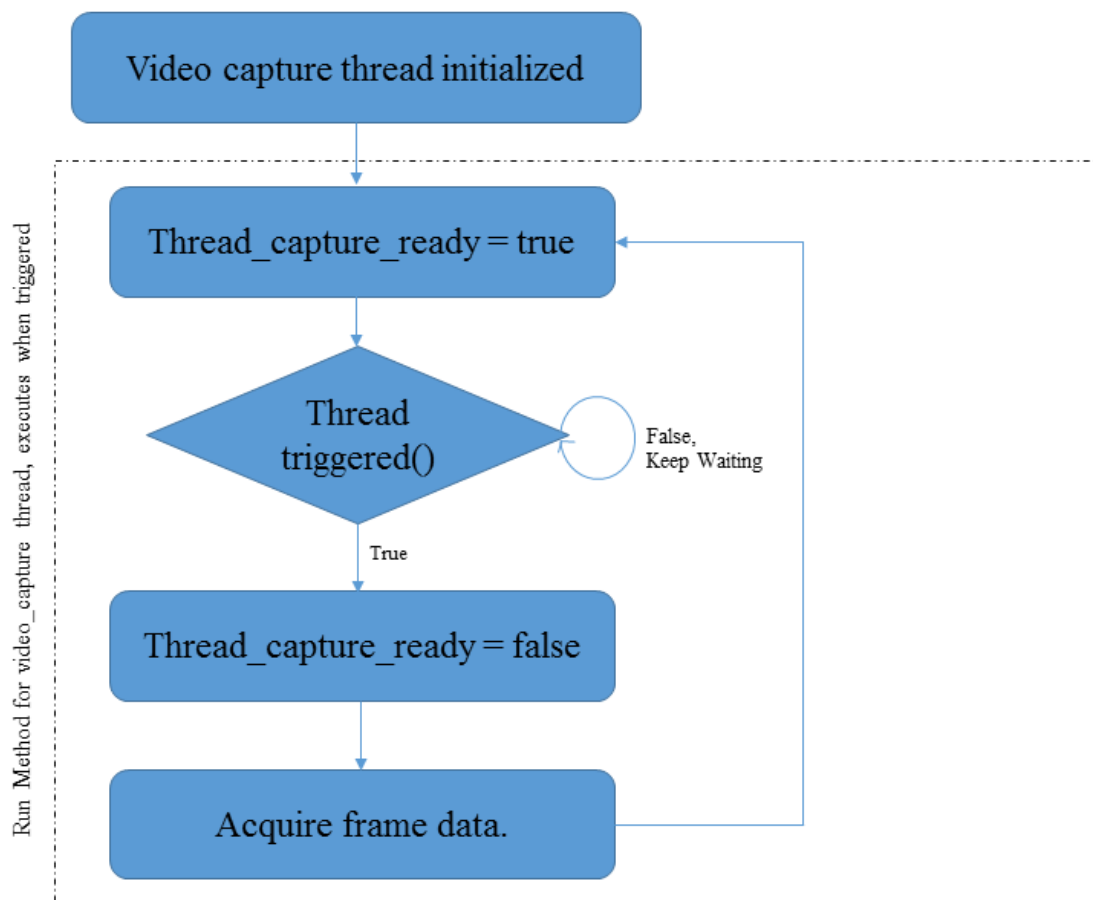
Figure 2.10: Flowchart of the run method of the video acquisition thread.
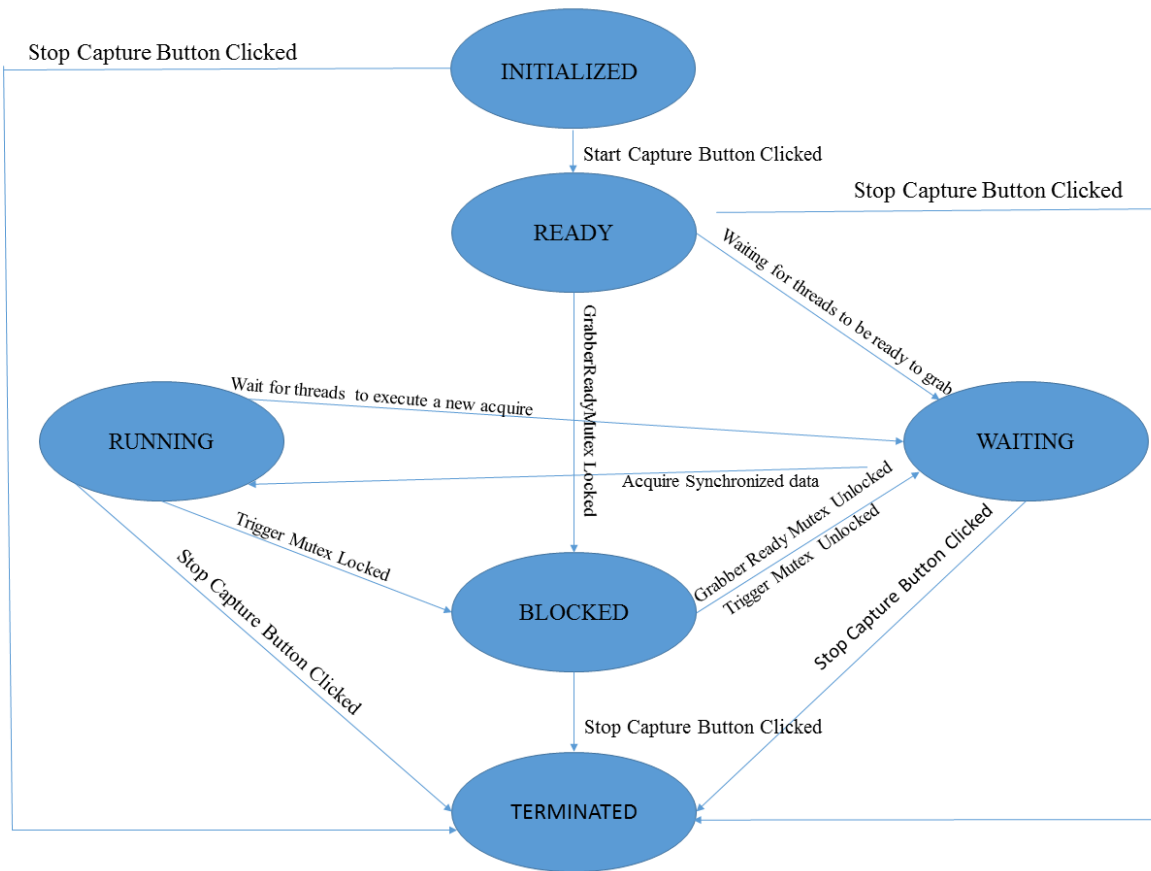
Figure 2.11: Video thread state diagram explaining the thread states and their interconnections.

2.3 Data Abstraction

Data Abstraction is the second component in the experiment's pipeline after data acquisition from various hardware devices. Abstraction is the process of reduction of the obtained data into a simplified form eliminating all non-essential information. In our experiment, abstraction corresponds to the extraction of the 3D location of the da Vinci® needle driver from the obtained stereo frames. The depth change in the tool's position while pressing the phantom is an essential feature for stiffness analysis. In section 2.3.1.3, we explain the da Vinci® tool-tip tracking algorithm for depth computation from the series of stereo frames acquired over the period of the experiment.

2.3.1 Articulated Surgical Tool Tracking Algorithm

MIS requires an additional training and skill set due to limited force feedback, reduced access and less guidance than conventional open surgical procedures [51]. Over the times, as the conventional operation procedures which were highly invasive in nature have shifted to MIS, it has become important to localize the tool in 3D space so as to accurately determine its relative positioning from other instruments and the organ. Surgical tool tracking improves awareness of the operational environment thus enhancing surgical planning, assessment, and interaction. The surgeon's interaction with the operational environment relies on visual feedback obtained from an endoscope camera. Therefore, MIS and RAMIS require training with a suitable metric for performance assessment. The development of such objective metrics depends on the surgical task as well as parameters such as completion time and tool's motion etc. The tool's motion analysis involves an objective evaluation of its trajectory, velocity, as well as its deviation from the pre-planned path derived from pre-operative imaging and surgical planning [52] [53]. The research by Cotin *et al.* [52] shows that tool-tip motion can result in a meaningful analysis of laparoscopic skills. Therefore, tool-tip localization can prove beneficial in the determination of these parameters for surgical skills assessment. The real-time feedback of the position of the tool can also assist in determining the distance of the instrument from tissues and organs. So, it is critical to localize the tool-tip for proper implementation of surgical planning, collision avoidance, and skills assessment. In our work, we have used tool-tip tracking to analyze the depth change of the tool as it presses on the phantom for tissue stiffness characterization.

2.3.1.1 Surgical Tool Tracking Methods

There are various works in the literature corresponding to tool-tip localization during MIS. The preferred characteristics of tool-tip tracking techniques are marker-less, real-time, and wireless communication. The review of methods used for tracking in MIS by Chmarra *et al.* [54] broadly classifies these systems into mechanical, acoustic, electromagnetic, ultrasonic and

vision-based techniques.

Mechanical tracking relies on kinematics and gimbal mechanism for the localization of a tool in 3D space about the base. Many systems in the past deployed mechanical tracking due to their highly accurate localization of the end effector. Earlier tracking devices used in operating rooms were mostly mechanical digitizers which had encoders at every joint such as the articulated socket joint assembly endoscopic tool mentioned in [55] [56]. These digitizers were highly accurate and reliable but were ruled out as they could not track multiple objects. Also, they were relatively slow, bulky and interfered with the surgical field.

Electromagnetic tracking consists of an electromagnetic field generator, a sensor, and an interpreter to decode the recorded signal from the sensors. It contains three coils arranged in a right-handed coordinate system that generates an electromagnetic field at a frequency of $8 - 14$ kHz [57]. This induced voltage, which is proportional to magnetic flux, tracks the position of the sensor. Northern Digital Inc. and Ascension technologies have introduced miniaturized electromagnetic tracking sensors that enable multiple sensor-based localization in an electromagnetic field. This is advantageous as it does not suffer from the limitation of line of sight that is present in vision-based tracking but its accuracy is not comparable to that of optical tracking. This technique may result in an inconsistent computation in the presence of metals due to distortion of the magnetic field.

Ultrasonic-based tracking applies segmentation to an image obtained from the US to differentiate background from surgical tools. As the tool moves, segmentation is achieved either using prior knowledge of the tool's shape or some image processing algorithm identifying a particular shape, color or texture. There are various image processing methods used for segmentation one being the use of the Hough Transform [58]. The ultrasonic-based tracking method is relatively more advantageous as it does not need any calibration, line of sight maintenance or correction for field distortion.

Vision-based tracking algorithms are broadly categorized as marker-based and marker-less tracking systems that are designed to identify a particular color, texture, pattern, wavelength or

shape. Active tracking corresponds to a presence of an attached physical marker on the tool that emits infrared light such as an LED for easier detection. Active markers are robust in nature with a high range of angular visibility for long distances. Passive markers, on the other hand, are retro-reflectors that reflect the incoming infrared light into a narrow range of angles that is easily identifiable. Northern Digital Inc. and Advanced Real Time Tracking commercially produce these markers that either emit or reflect light and are used for detection and localization of an object in 3D. Other commercially available optical markers from Claronav Technology have a specifically designed black and white pattern passive markers that reflect visible light and are detected by Micron Tracker's stereo camera and software. This technique uses image processing and computer vision-based technique for 3D object localization and has been used in research work for tool-tip tracking in image-guided procedures. Figure 2.12 shows one such application with markers on the tool and head skull phantom for registration and efficient navigation [59]. Other works in the literature based on marker detection are bar-code based design markers [60], alternating strips based markers for depth estimation [61] and colored markers designed such that it is different from the background color [62]. Attaching a physical marker with an identifiable texture, pattern or color is computationally less expensive to localize and more robust due to the presence of a salient feature, but it is not a feasible solution for surgical procedures due to the risk of losing the marker inside the body during MIS.Though marker-based vision tracking is computationally less expensive and robust due to the presence of a salient feature, it is not practical to use it in MIS.

Marker-less tracking harnesses features such as texture, color, gradient change, and the prior knowledge of the shape of the tool. Previous work in tool-tip detection involves the use of template matching [63] [64] to recognize the pre-determined tool template in all the re-occurring frames. Another work on da Vinci® tool-tip localization by Reiter *et al.* uses multiple features such as position, hue, saturation, gradient, and orientation combination in a covariance region descriptor matrix [65]. These feature descriptors are often combined with learning algorithm for robust tool detection. One such work that involves offline learning is

based on the Haar classifier which is trained using a template data set that contains the tool's picture in various orientations and scales [66]. Pezzementi *et al.* designed an offline based learning procedure where manually labeled pictures are used for training the classifier, which is later used in segmenting the tool shaft, metallic tip and background [67]. Therefore, it is more practical and safer to use marker-less tracking that uses computer vision, image processing and learning for localization.

Some systems combine electromagnetic tracking with optical tracking to resolve the problem of line of sight and other inaccuracies for better localization [68]. These hybrid system designs are much more accurate than the standalone use of other mentioned techniques, but they are not cost effective in general.
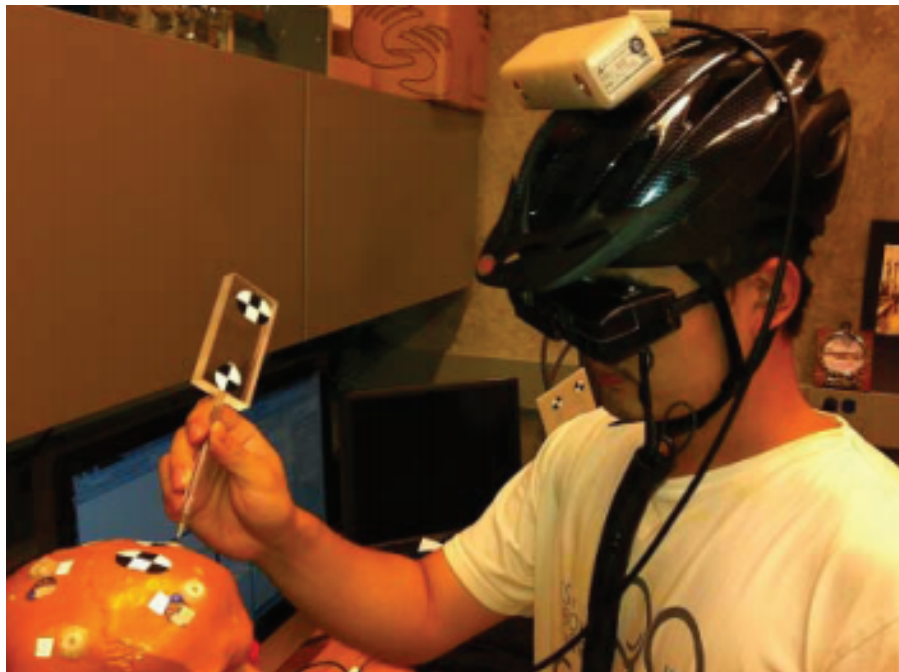
Figure 2.12: Head-mounted navigational system developed using micron trackers for use in neurosurgery [59].

To sum up, in this section, we discussed various surgical tool tracking methods developed over the decade, and their advantages and limitations. Our framework is designed for use in RAMIS palpation, so its design should be wireless and markerless. Also, we have SEM that can be utilized for 3D localization of tool-tip. Therefore, we have used a vision-based image

processing algorithm for tool-tip localization in 3D that we explain in detail in section 2.3.1.3.

2.3.1.2 Markerless Vision-Based Tracking Algorithms

Real-time computation, illumination variation, occlusions, scale and rotation changes in the feature detection are some major bottlenecks addressed during the development of tracking algorithms. Essentially all these tracking algorithms are an amalgamation of the following three processes: Inter-Frame Motion Estimation (IFME), Feature Localization (FL), and learning. **IFME** means to compute the shift in feature position as it makes a transition from one frame to another. Many algorithms detecting IFME use temporal coherency for shift detection. It is an estimation problem where translation, scale change, rotation, and deformation of the feature is predicted in the next frame based on knowledge of these parameters in the previous frame. Either these trackers search and localize the feature by searching in the vicinity of the previous feature position or they register the change in state (scale, rotation, deformation) of the defined parametric model. Several trackers are suggested in the literature such as Lucas-Kanade [69], mean shift computation [70], etc. **Detection** can be defined as localizing a feature based on visual information in the search space. Feature-based detection depends on the salient feature extraction from an image sequence for the computation of motion parameters. Detection can be defined as the localization of object of interest using features such as histograms [71], the gradient [72] or a combination of multiple features in the spatial domain [65]. These features are matched in the recurrent frame using some similarity metric to localize the point in the search space in the next frame. In case of large inter-frame distance where optical coherency is not present, we use feature-based descriptors such as SIFT [73], SURF [74], BRISK [75], etc. The problem associated with the use of these feature based descriptors is that they rely on the presence of a stable feature and they are time-consuming in nature due to the dense mapping, and outlier rejection computation. These descriptors are mostly used for initialization as they give better localization in the initial frame. Some algorithms incorporate learning in the tracking algorithm for a robust detection. These learning algorithms can be classified as

offline and online learning. The offline learning as the name says is performed offline. Before tracking, an offline classifier is trained on a data-set of templates to learn the salient feature for accurate detection during the tracking. In [66] offline based learning is used for da Vinci® tool-tip localization and in [76] random forests are used to learn the changes in feature.

2.3.1.3   3D da Vinci® tool-tip localization algorithm

2.3.1.3.1   Purpose   In the coming paragraphs, we present an algorithm capable of detecting 3D location of the da Vinci® instrument. Template based matching is a simple and an efficient tracking algorithm that detects a known template's presence in the image. This tracking method is conceptualized based on the degree of similarity between a template and the base image. The goal of this algorithm is to find the occurrence of the template in the search image by computing a pixel-wise similarity match. The similarity metric is designed to evaluate the image patch that has the best match corresponding to the template image. The design of the similarity metric is dependent on the parameters such as insensitivity to noise and luminance change, computationally effectiveness. Normalized Mean Square Error (NMSE), Mean Absolute Difference (MAD), Normalized Cross Correlation (NCC), and Hausdorff Distance (HD) are some common similarity metrics used for matching. The problem associated with the application of template matching for localization is that the template continuously changes in scale and orientation. The template undergoes scale changes as the tool moves closer and farther from the camera. Also, it is not possible to maintain the same orientation throughout the experiment. Therefore, the template at any instance is a scaled and rotated version of the initially selected template. If a fixed initial template is used throughout the analysis, then it may result in a faster detection but will suffer from inaccuracies. If the template is updated at each iteration, then it is more robust to appearance change, but it will result in template drifting, where the template drifts from the area of interest over the progression of the tracking algorithm. Every time we iteratively update the template corresponding to the current template obtained from the search space, the inaccuracies keep on accumulating, and the tracking

is no longer robust over the period of tracking. Therefore, we need to accurately register the changes in scale, rotation, and translation as the tool moves during the experiment. Translation, rotation, and scaling all have their components in the frequency domain. So, we continuously register the original template with the updated template to account for the inaccuracies generated due to motion. The updated template is a warped image obtained from original image accounting for rotation and scale changes. We have used gradient based template matching for tool-tip detection with a constant update of the template for scale and rotation changes. Before we proceed, we explain the symbols used while explaining the algorithm. The subscripts l and r denote left and right images acquired from SEM.

$I(x, y, t)$ denotes the grayscale brightness value at a location (x, y) in an image that was acquired at time t.

$T(x, y, t_0)$ denotes the grayscale brightness value at a location (x, y) in template image obtained initially at time t_0.

$T(x, y, t)$ denotes the grayscale brightness value at a location (x, y) in a template image obtained at time t.

$\nabla_x f(x, y, t)$ denotes the spatial gradient of the image $I(x, y, t)$ in the x direction at location (x, y) and time t.

$\nabla_y f(x, y, t)$ denotes the spatial gradient of the image $I(x, y, t)$ in the y direction at location (x, y) and time t.

2.3.1.3.2   Initialization   The user manually selects the template's position in the initial left frame in the GUI shown in Figure 2.8 with a mouse click. The template obtained from the left image is searched in the right image until convergence using the Matching Method described in the section 2.3.1.3.4. These captured templates are stored as a reference template pair in the memory that is updated throughout the experiment for changes in orientation and scale. The iterative template update for scale and rotation changes has eliminated the overhead of generating a database of the templates at various viewpoints and states.

2.3.1.3.3 Template Update The reference template is not a robust descriptor and therefore, we continuously update the template image for change in scale and orientation. In literature template based matching for articulated tool-tip localization in surgical domain has been used in [63] [77] [78]. These algorithms use a database of templates at various probable orientations and scales. A brute force algorithm is applied to compute the similarity match for each template and the one with the best similarity score is chosen as the correct match. This approach is computationally very intensive and less robust in nature. Other approaches in the literature involve kinematics fusion with the tracking algorithm to compute yaw and pitch of the tool at each iteration. This fusion enables the construction of a rotated and scaled template at every iteration for similarity computation [78]. We have extracted scale and orientation changes by studying the shift in the frequency domain of the log-polar transformed reference and the current template image. It can be inferred from Fourier shift theorem, two images that are translated copies of each other in the spatial domain shows a similar phase difference in their frequency domain [79]. The templates obtained from the row-aligned, rectified stereo images are theoretically a translated copy of each other. We phase correlated the left and the right template images to obtain the apparent disparity. In an ideal scenario with zero calibration RMSE, the Relation 2.5 would hold between a pair of rectified stereo images at any time instant t. Equations 2.6, 2.7, 2.8 show the Fourier transforms of the template image. The pair is registered for translation using the NCC metric to compute the disparity .

$$T_l(x, y, t) = T_r(x+\Delta x, y, t) \tag{2.5}$$

Let the Fourier transforms of the left template image and the right template images be $\tilde{T}_l(\omega_x, \omega_y)$ and $\tilde{T}_r(\omega_x, \omega_y)$ respectively; let the relation between the Fourier transforms of these images be as follows

$$\mathscr{F}\{T_r(x, y)\} = \tilde{T}_l(\omega_x, \omega_y)e^{j\omega_x\Delta x} \tag{2.6}$$

The NCC of the Fourier transforms of both images can be written as 2.7. The translation parameter Δx can be obtained by evaluating the maximum value of this similarity metric. This value will correspond to the best match among the two translated copies.

$$\widetilde{NCC}(\omega_x, \omega_y) \triangleq \frac{\tilde{T}_l(\omega_x, \omega_y)\tilde{T}_r^{\,*}(\omega_x, \omega_y)}{|\tilde{T}_l(\omega_x, \omega_y)||\tilde{T}_r^{\,*}(\omega_x, \omega_y)|} \tag{2.7}$$

$$\Delta x = arg \max_{x,y} \mathscr{F}^{-1}\{\widetilde{NCC}(\omega_x, \omega_y)\} \tag{2.8}$$

The Fourier-Mellin transformation is used to register images with misalignment due to translation, scaling and rotation [80–82]. Figure 2.13 shows the cartesian coordinate transformation to the log-polar domain. The polar domain transformation is defined in such a way that its axes denote the radial distance and the angle of a point from the center of the transformation.The rotations in Cartesian domain are manifested in the polar domain as translation. To be able to deduce scale and rotation changes, we have used the log- polar transformation where scale and rotation change become separable. Figure 2.13 shows the pixel transformation from Cartesian domain to polar domain and the template image transformation from Cartesian to the log-polar domain. The current template image is essentially a rotated and scaled replica of the original template image can be mathematically denoted as Equations 2.9 and 2.10 for the stereo image pair. Therefore, the current template and reference template are a translated copies of each other in polar domain. The DC component of the Fourier transform of this image pair, is related by rotation and scaling components [82]. The current template and the reference template are transformed into the log-polar domain using bi-linear interpolation and then phase correlated in the Fourier domain to estimate rotation and scale change. The translation in x and y coordinates in the log-polar domain corresponds to scale and rotation change respectively. We obtained parameters (s,θ) by phase correlating reference template with the current template in the log-polar domain. The updated template as shown in Figure 2.14 is obtained by warping the reference template with the obtained transformation matrix to compensate for scale and rotation changes. This update significantly reduces the template drift due to the accumulation

of incorrect estimation of inter-frame motion.

$$T_l(x,y,t) = T_{lo}(s_l * x * \cos(\theta_l) + s_l * y * \sin(\theta_l) + \Delta x, -s_l * x * \sin(\theta_l) + s_l * y * \cos(\theta_l) + \Delta y, t_o) \quad (2.9)$$

$$T_r(x,y,t) = T_{ro}(s_r * x * \cos(\theta_r) + s_r * y * \sin(\theta_r) + \Delta x, -s_r * x * \sin(\theta_r) + s_r * y * \cos(\theta_r) + \Delta y, t_o) \quad (2.10)$$

A) Cartesian Domain   B) Polar Domain

C) Cartesian Domain   D) Log-polar Domain
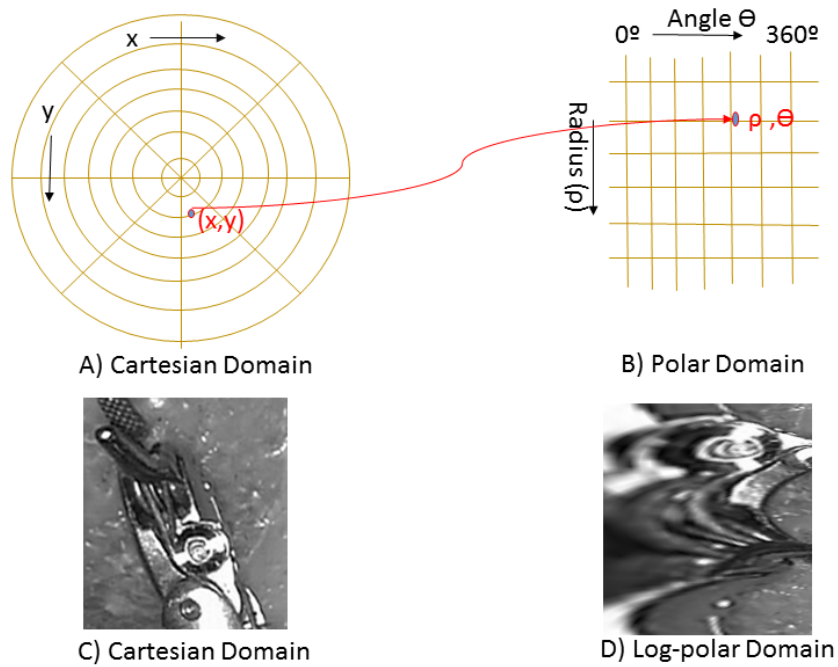
Figure 2.13: Figures demonstrating conversion from Cartesian coordinates (A) to polar domain (B) and conversion of the current template in Cartesian domain (C) to log-polar domain (D).
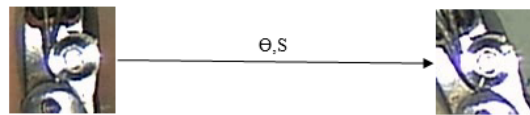
Figure 2.14: Updated template obtained by warping the reference template for changes in orientation and scale.

2.3.1.3.4 Matching Method After the template is updated for rotation and scaling changes, this section presents the matching method used to locate the updated template's presence in the

search space. The matching method compares characteristic features of the template image against the search image using NCC to find the best match with maximum similarity metric value. We have not used binarization of images as it require a threshold value for the construction of a binary image. This threshold is not a constant, and is dependent on scene illumination and motion. Therefore, a binary edge map is not a very robust feature. We have chosen gradient information as the characteristic feature descriptor in the spatial domain as it depicts the directional change in image intensity. Gradient descriptors are insensitive to illumination change and work well for texture-less object detection. We have used an orientation magnitude image convoluted with a 2D Gaussian for template-based matching. The template and search image orientation magnitude and Gaussian convoluted images are computed using the Equation 2.11 and the Equation 2.12 respectively. The best match from these processed images is calculated using Equation 2.13. Figure 2.15 shows the process of finding the updated template image presence in the search space.

$$\left\| \nabla f \right\| = \sqrt{(\nabla_x f(x, y, t))^2 + (\nabla_y f(x, y, t))^2} \tag{2.11}$$

$$Ori(x, y, t) = G * \left\| \nabla f \right\| \tag{2.12}$$

$$\varepsilon_m(\widetilde{Ori_B(x}, y, t), \widetilde{Ori_T(x}, y, t)) \triangleq \arg \max_{x,y} \frac{\widetilde{Ori_B(x}, y, t)\widetilde{Ori_T(x}, y, t)}{|\widetilde{Ori_b(x}, y, t)||\widetilde{Ori_T(x}, y, t)|} \tag{2.13}$$
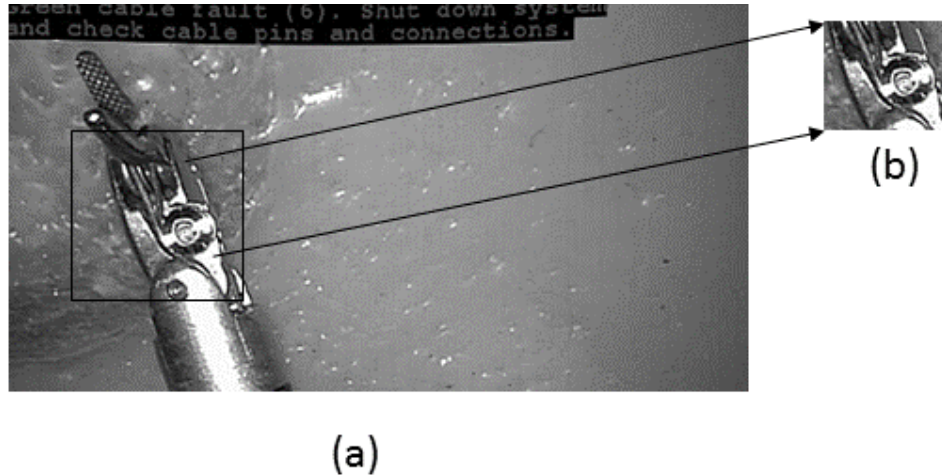
Figure 2.15: Iterative search for a template image in the search space image using NCC.

2.3.1.3.5 SubPixel Localization   Image is quantized in terms of discrete entity known as pixel and the matching method returns the pixel location that indicates the best match. The maximum of the correlated surface has a resolution of a pixel, but that location might not be the exact position where the best match occurs. Many applications, such as ours require an accuracy of a sub-pixel to determine the depth change in sub-millimeter for a precise evaluation. Sub-pixel determination involves computing the matching results better than a pixel accuracy. Generally, in the literature, two preferred approaches are used. One where the correlation surface is mapped to a higher resolution using interpolation algorithms and the second where a continuous analytic function is optimally fitted to the correlation surface function around the peak to obtain a sub-pixel accuracy. There are various functions such as bi-cubic interpolation, Gaussian fitting, orthogonal one-dimensional paraboloid pair fitting are used in the literature for the continuous fitting of the discrete correlation function. The NCC surface obtained from a pixel-based similarity score resembles a bell-shaped curve around the peak. Therefore for our purpose, we have used a least-squares regression method to obtain the quadratic function coefficients defined in [83]. Gleason *et al.* [83] defined a deterministic solution in their work by fitting a one-dimensional parabola through the peak and neighboring points. The NCC surface is expressed as a parabola with six unknown coefficients as shown in Equation 2.14; the nine

known points on the surface are used in modeling the parabola.

$$NCC(u, v) = au^2 + bv^2 + cuv + du + ev + f \tag{2.14}$$

The above equation with six unknown coefficients and nine known correlation points is an overdetermined system which is computed using least-square linear regression. After the coefficients are obtained, the maximum value of the parabola is calculated in terms of the coefficients. These results given in [83] are shown in Equations 2.15, 2.16. These are closed-form solutions that avoid iterative re-computation for fast and efficient subpixel localization.

$$x_{ls} = \frac{2db - ce}{c^2 - 4ab} \tag{2.15}$$

$$y_{ls} = \frac{2ae - dc}{c^2 - 4ab} \tag{2.16}$$

2.3.1.3.6 Extension to 3D Localization Depth computation is a very well researched problem of detecting stereo correspondence in the stereo images in computer vision. We have already discussed in detail the algorithm for tool-tip localization. The same algorithm is used for localizing the tool in the left and the right frame. The pair of rectified images is row aligned, so the shift along the column of corresponding points on the left and the right frames gives rise to the perception of depth. The process of backtracking the 3D point after localizing corresponding pixels in the left and the right images is known as triangulation. The geometric setup and positioning of the rectified image pair can be understood from the Figure 2.6. Equation 2.17 defines the relation between depth, disparity, focus and baseline:

$$z_{tool} = \frac{fb}{disparity} \tag{2.17}$$

The focus value for each camera and the baseline is estimated from stereo calibration,

and it remains constant throughout the experiment. The apparent disparity obtained by phase correlating the stereo template images and that observed after the localization of the tool-tip in stereo frames are fused together using Kalman filtering with equal weights in the measurement equation to obtain a better estimate of the predicted disparity. The resolution of the computed depth is dependent on the accuracy of the calculation of disparity as defined in Equation 2.18 [84].

$$\Delta z = \frac{z^2 \Delta d}{fb} \qquad (2.18)$$

The disparity computation accuracy is dependent on the localization algorithm and the stereo calibration accuracy. The error in depth measurement increases quadratically with the increasing depth. SEM also has a short baseline as it is designed to pass through a trocar port during RAMIS. A short baseline is not a good fit for long range sensing and also depth accuracy computation degrades with increasing depth. It is because of these reasons we kept the SEM close to the phantom while performing the experiment. In the results and discussion, we present the color-coded map onto a small region of the phantom because the area of focus while performing the experiment is a small subsection of the entire section.

Depth computation accuracy validation   We designed an experimental setup as shown in Figure 2.16 using a linear stage (Zaber Technologies Inc., Canada) motion for depth computation accuracy validation. The linear stage is commonly used for cycling applications where it is actuated to move in a particular waveform with defined amplitude, frequency and a total distance of travel. We used the triangular wave generation script from Zaber Technologies to actuate the linear stage in a triangular waveform with an amplitude of 16*mm*. Figures 2.17, 2.18 show that the depth computed from the camera follows the same variation as the linear stage's displacement with a resolution of 0.5*mm*. Figure 2.18 shows a fitted linear function of depth measurement and linear stage displacement with a slope of approximately unity indicating the accuracy of the depth validation experiment.

Figure 2.16: Experimental setup for depth validation using linear stage motion.
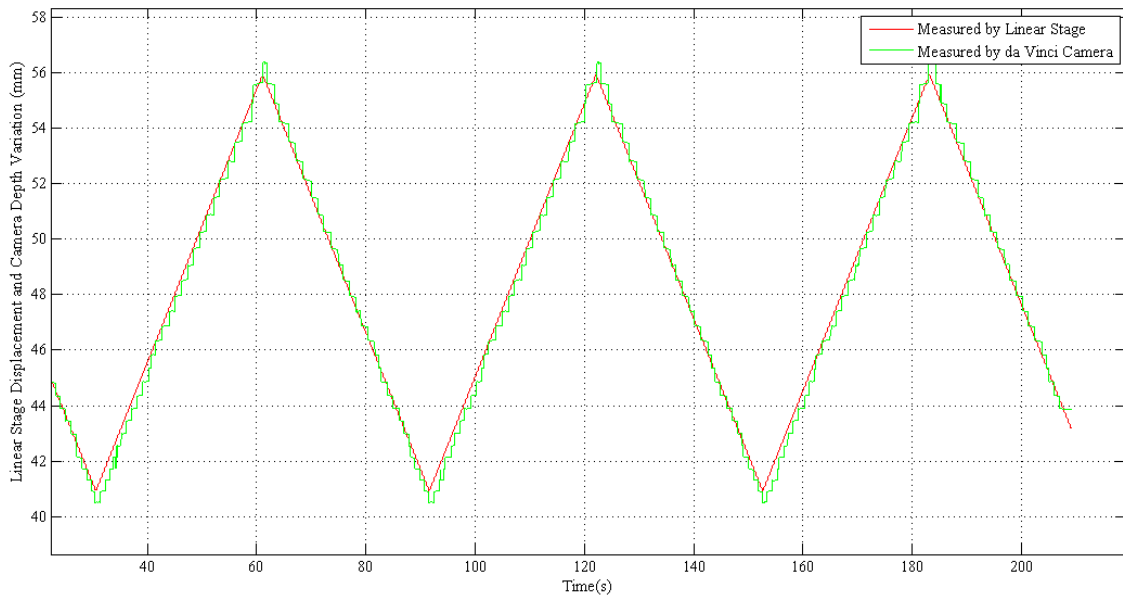
Figure 2.17: Experimental validation showing camera measurements following linear stage sinusoidal displacement with an amplitude of 16mm.
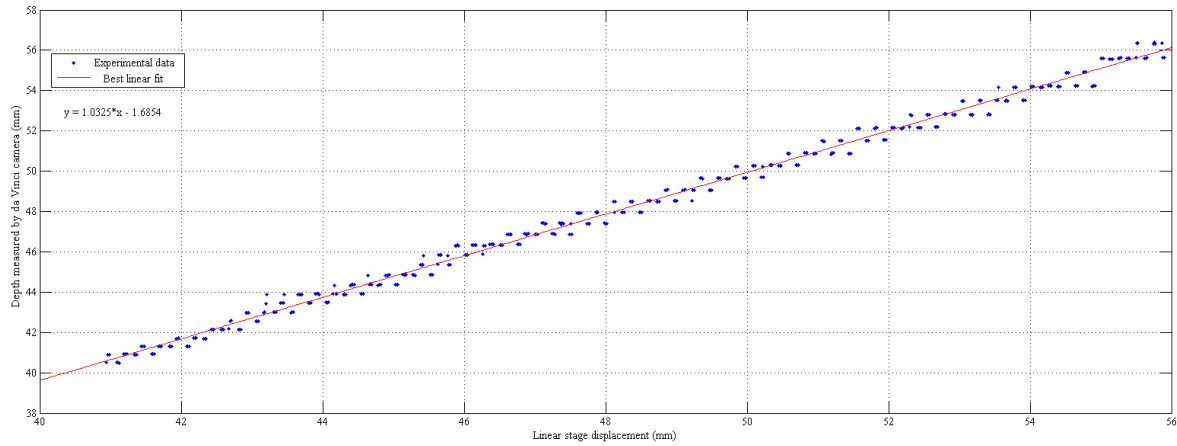
Figure 2.18: Validation of da Vinci® stereo camera depth tracking showing a linear best fit for depth measurement from camera with respect to linear stage displacement.

2.3.1.4 Tracking Algorithm Description
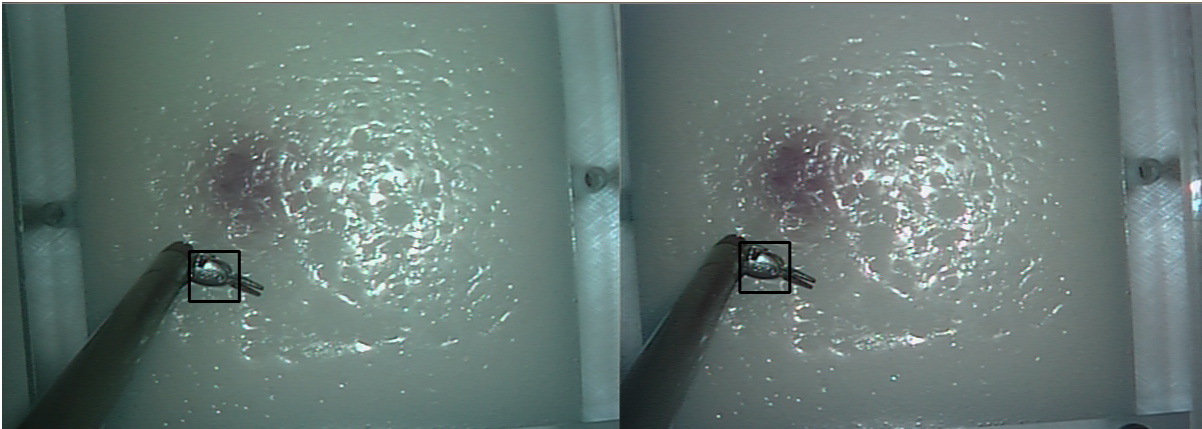
Figure 2.19: Left and right frame tool-tip localization

Algorithm 1 Tool-tip localization algorithm

Input: I_l , I_r, $calibration_{param}$
Output: x_{tool}, y_{tool}, z_{tool}
  1: **Initialization**
            $\hat{I}_l, \hat{I}_r \leftarrow stereoRemap(I_l, I_r, calibration_{param})$
            $x_{click}, y_{click} \leftarrow userIntialization()$
            $T_{lo}, T_{ro}, B_{lo}, B_{ro} \leftarrow computeSearchParameters(\hat{I}_l, \hat{I}_r, x_{click}, y_{click})$
  2: **while** $i \leqq num_{frames}$ **do**
            $\hat{I}_l, \hat{I}_r \leftarrow stereoRemap(I_l, I_r, calibration_{param})$
            $T_l, T_r, \Omega_l, \Omega_r \leftarrow computeSearchParameters(\hat{I}_l, \hat{I}_r, x_l, y_l, x_r, y_r)$
            $\theta_l, s_l \leftarrow logPolarTransform(T_{lo}, T_l)$
            $\theta_r, s_r \leftarrow logPolarTransform(T_{ro}, T_r)$
            $disparity_{apparent} \leftarrow phaseCorrelation(T_l, T_r)$
            $T_l, T_r \leftarrow warp(T_{lo}, \theta_l, s_l, T_{ro}, \theta_r, s_r)$
            $\hat{x}_l, \hat{y}_l, \hat{x}_r, \hat{y}_r \leftarrow matchMethod(T_l, \Omega_l, T_r, \Omega_r)$
            $\hat{x_{ls}}, \hat{y_{ls}}, \hat{x_{rs}}, \hat{y_{rs}} \leftarrow subPixelLocalization(\hat{x}_l, \hat{y}_l, \hat{x}_r, \hat{y}_r)$
            $disparity_{observed} = \hat{x}_l - \hat{x}_r$
            $disparity_{predicted}, x_l, x_r \leftarrow kalmanFilter(disparity_{apparent}, disparity_{observed}, \hat{x}_l, \hat{x}_r)$
            $x_{tool}, y_{tool}, z_{tool} \leftarrow 3DLocalization(disparity_{predicted}, Q)$
            $i = i + 1$
  3: **end while**

2.4   Experimental Setup
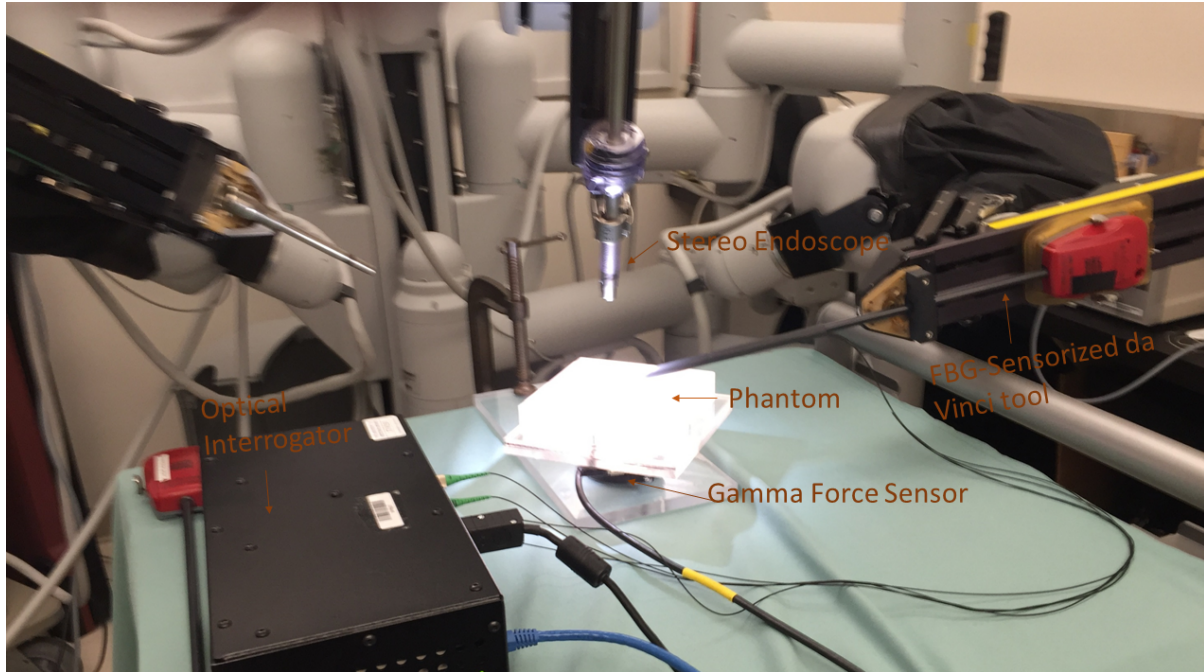
Figure 2.20: Experimental setup on the patient side showing the PSM, the FBG-sensorized tool, phantom, SEM and the optical interrogator.

The stiffness of a tumor is generally higher than that of healthy tissue and can be easily differentiated during an OSP by manual palpation. The hypothesis of this experiment is that the two different materials will show different compression when pressed with the same amount of force. This behavior is the tissue's characteristic property and has been tested to determine elastic and plastic properties during uni-axial compression tests in [85]. In this framework design, we intend to develop a mechanism similar to a finger's pressing motion during manual palpation. A master-slave teleoperation system, i.e. the da Vinci®, is used to generate repetitive pressing motions by moving the MTM to manipulate the patient-side robotic arm. Figure 2.21 shows the experimental setup on the PSC with a phantom and the FBG force-sensorized da Vinci® tool. The force-sensorized tool is used for sensing lateral forces while a point is compressed to study the change in depth. The user has a continuous feedback of the amount of lateral force exerted while pressing on the phantom. The vision-based tool tracking explained

in Section 2.3.1.3 is used for depth sensing as the robotic arm mimics the user's motions. Figure 2.21 shows different orientations of the sensorized da Vinci® tool as it presses on the phantom. As shown in the figure, orientation A results in a higher angle between the phantom plane and the palpation plane. However, orientation B where the tool is arranged in a horizontal manner, results in a good alignment between the planes with minimal rotation. Orientation B maintains an impact along the normal axis and also results in efficient lateral force sensing. We have used orientation B for our experimental system design. Figure 2.22 depicts force sensing in orientation B for eight sampled points using the FBG force-sensorized tool and the commercial Gamma ATI force sensor placed underneath the phantom for validating the force estimated using the tool in orientation B.

(A)

(B)
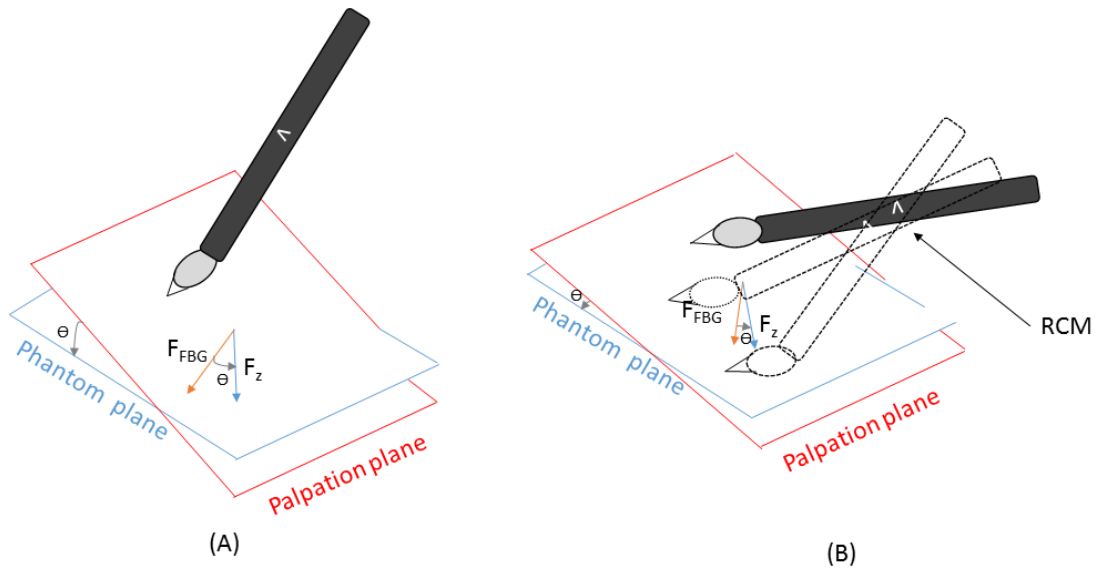
Figure 2.21: Probable orientations for palpating the surface during RAMIS.
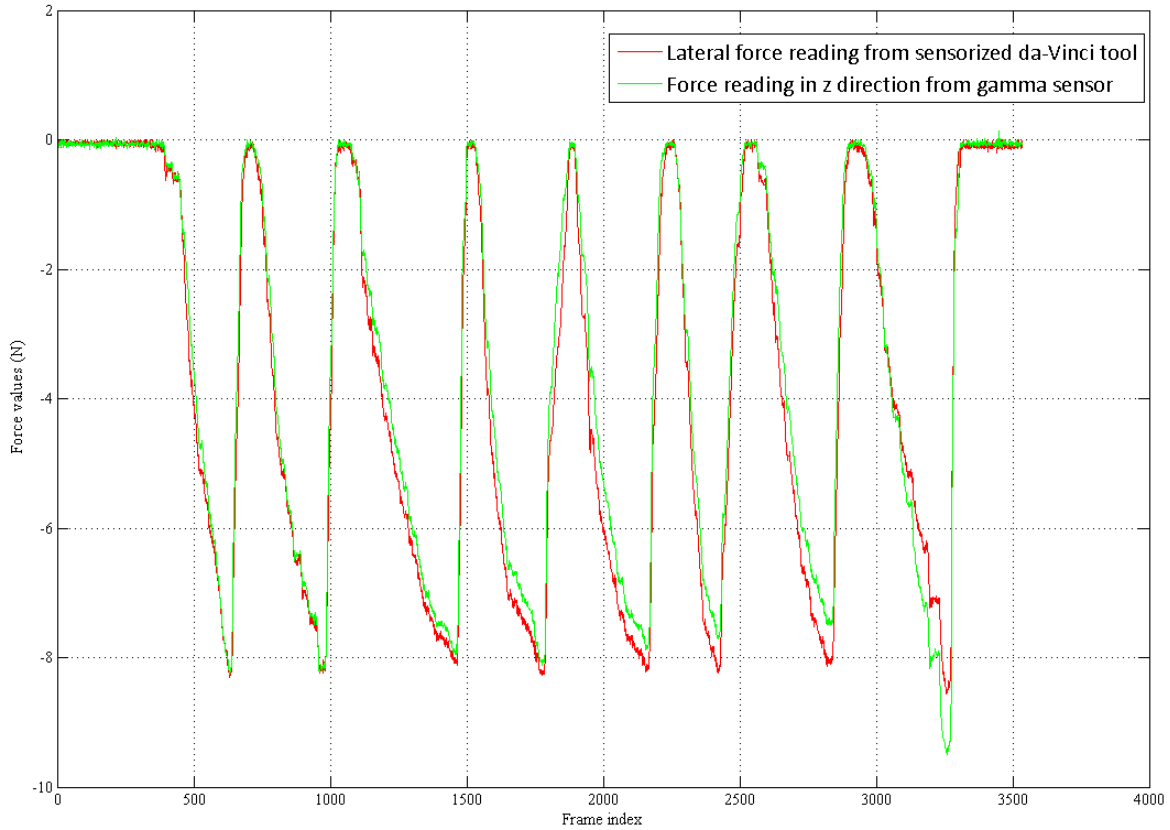
Figure 2.22: Force validation plot showing the forces measured by the FBG-sensorized tool and Gamma ATI force sensor in orientation B.

2.5 Data Interpretation

2.5.1 Sampled Point Stiffness Computation

Precise contact dynamics modeling of tool-phantom interactions can provide an estimation of conservative and non-conservative force parameters. Various localized continuous deformation models accurately estimate the contact and damping forces between a rigid sphere like object and a deforming plane as a combination of springs. KV is one such (linear) model that establishes a relation between the force and compression as a parallel connection of linear springs [86]. HC is another contact dynamics (non-linear) model that even takes into account the energy loss during contact [87]. Yamamoto *et al.* [34] compared seven contact models for best

tool-tissue interaction estimation. They computed that HC is the best model for delineating hard inclusion from soft surrounding material based on a parameter K estimated from the HC model. Lateral forces computed from the FBG-sensorized tool are equivalent to a reaction force applied to the phantom's surface. The HC contact model is defined in Equation 2.19 where K is the stiffness, λ is the damping factor, n is dependent on the local geometry around the contact surface.

$$F(t) = \underbrace{K x^n(t)}_{\text{Non-linear elastic force component}} + \underbrace{\lambda x^n(t) \dot{x}(t)}_{\text{Non-linear viscous force component}} \qquad (2.19)$$

Yamamoto *et al.* [34] used K (stiffness) as an estimator for differentiation visualization. Instead of K we characterize the tissue stiffness based on the maximum indentation depth attained on pressing the sampled point with the same force. We used the nonlinear grey-box model script (Matlab) for unknown coefficients estimation from the lateral force, depth and depth rate. The non-linear grey box model estimates the unknown coefficients from known data points based on the differential state equation of the HC model. In the Table 2.1 we present the parameters estimated from the HC model and maximum indentation depth obtained by pressing on the sampled points. We have highlighted the points in Region A (stiffer) in red and points in Region B (softer) in blue for differentiation.

Table 2.1: HC parameters and maximum depth obtained for eight sampled points.

Point Index	n	λ	K	d_{max}
1	0.823	0.6674	-3.7492	4.0506mm
2	0.9533	1.825	-3.0671	3.7203 mm
3	0.9620	0.5996	-2.4827	4.6906 mm
4	2.5071	0.0002	-0.0494	8.3410 mm
5	1.1942	0.0680	-0.8429	7.5510 mm
6	1.5673	-0.0018	-0.2319	10.614mm
7	1.4549	0.0092	-0.3649	9.8145 mm
8	1.6655	0.0315	-0.3729	7.7777 mm

We have presented the HC parameters and maximum depth in the above table to show that depth can characterize the difference in stiffnesses. The estimation of HC parameters for tissue property characterization is computationally expensive and may increase the error if the nonlinear grey box model is not able to estimate the coefficients accurately from the data obtained. Our work has simplified the process from multiple parameter estimations to a single parameter analysis. Now, when unsupervised clustering algorithms are applied for data evaluation instead of a multidimensional analysis on the parameters, the computations are done only on maximum depth variation. When a similar amount of force is applied to different materials, the compressive reaction i.e. the indentation is indicative of its stiffness.

2.5.2 Clustering Technique

Clustering approaches are broadly classified into parametric and non-parametric based clustering. The choice of the clustering algorithm is dependent on the type of data, application, and computational cost. Clustering refers to assigning the data into groups based on similarity of data points depending on common characteristic parameters. The K-means is a parametric hard partitioning clustering approach that divides the data into groups where the number of groups are known a *priori*. The basic K-means functioning is defined in Algorithm 2. The initial centroids are randomly placed to divide the data point into k clusters with k centroids. After every iteration, the data points are re-assigned to a class based on their distance from the nearest centroid. After reassigning the point by evaluating Euclidean distances, the new centroid position for each of the k clusters is computed. This whole recursive task is performed until the convergence condition is satisfied, i.e, the position of the k centroids stops varying. The run time of the K-means algorithm is dependent on these four parameters: the number of clusters (n_c) , the number of iterations (n_i), the number of data points (n_d) and the dimension of the data points (d).

Algorithm 2 Basic K-means Algorithm

Input: Data points, Number of clusters (k)
Output: k clusters
    Initialization Select K initial cluster centroid $C_1, C_2...C_k$
 2: **while** (*convergenceCondition*$! = True$) **do**
    Form clusters by labeling every point to the closest centroid.
       Recompute the centroid for each class by averaging the data points in respective class.
 4: **end while**

The K-means algorithm produces tighter clusters with small distances within the group and a large separation of groups from each other. The K-means algorithm is computationally more effective than other hierarchical clustering algorithms and also easier to implement. But it also suffers from a major drawback that the user has to decide the number of groups *a priori*. It is very sensitive to noisy perturbations and also being a greedy algorithm, it can converge to a local optimum. Therefore this clustering approach is dependent on the initialization of the centroid; hence different initializations can result in different partitions due to convergence at local minima rather than a global optimum.

Algorithm 3 Clustering algorithm for stiffer area segmentation.

Input: Maximum depth change for sampled points, Number of clusters ($k = 2$)
Output: Maximal area spanned by stiffer region color coded ($I_{colorCoded}$)
    Initialization Select K initial cluster centroid C_1, C_2
    $(C_A, C_B) \leftarrow KMeans(dmax_i, 2)$
    $dMean_A = \sum\limits_{i \in C_A} dMax_i/n_A$
    $dMean_B = \sum\limits_{j \in C_B} dMax_j/n_B$
  if $dMean_A < dMean_b$ **then**
    $C_A = True$;
  else
    $C_B = True$;
 3: **end if**
    $x_{min}, x_{max} = (\min\limits_{i \in C_A} x_i, \max\limits_{i \in C_A} x_i)$
    $y_{min}, y_{max} = (\min\limits_{i \in C_A} y_i, \max\limits_{i \in C_A} y_i)$
    $I_{colorCoded} = MaximalAreaS pan(x_{min}, y_{min}, x_{max}, y_{max})$

Algorithm 3 shows the computation of the area spanning the cluster points that show less compression. After the K-means algorithm has segregated the sampled points into two clusters,

we compute the group with the minimum mean compression to label it as the stiffer region. Then for that group, we compute the maximal area spanned by the group on the image and segment that cluster to highlight the stiffer region.

2.6 Results and Discussion

The proposed procedure is a 3-stage process, and we have already discussed synchronized data acquisition, data abstraction and data interpretation in Sections 2.2, 2.3, 2.5 respectively. In this section, we conclude this chapter's work by presenting the results obtained from the interaction of the FBG-sensorized tool and the phantom in a manner similar to the finger's pressing motion during manual palpation. We give a detailed description of the results, experimental verification and possible future enhancements for the framework presented in this chapter.

As, mentioned previously, the accuracy of depth measurement decreases quadratically, so the SEM is placed close to the phantom for accurate measurements. Figure 2.23 shows the top view of the phantom section captured from the da Vinci®'s SEM. The labeled phantom's section shown in Figure 2.23 depicts the boundary separating the stiffer surface (Region A) from the soft surface (Region B). The phantom model shown in Figure 2.23 is prepared from Ecoflex0030 with a silicone thinner for making Region B. The different colors of Region A and Region B are also indicative of different materials used while preparing the phantom. The experimental setup and the motion of the tool-tip are shown in Figure 2.20, 2.21 respectively. The user palpates four points in Region A and five points in Region B while getting real-time force and video feedback on the stereoscopic head rest. In this experiment, the user has tried to maintain a maximum force of $8N$ for each palpated point. The force feedback enables the user to stop when a force of $8N$ is applied on each point. There is some variability in controlling the maximum force applied; the error usually ranges within $\pm 1N$. While palpating point 8, the user has applied higher force than the chosen empirical value and while palpating point 9, the user has applied lesser force than $8N$. The results are dependent on the control and action

of the user while performing the experiment. The chosen maximum force value of $8N$ is an empirical selection, and we have conducted the experiments at $10N$ as well. Figure 2.24 shows the change in force and depth as points are palpated with a maximum force of $8N$. Since the user tries to push only up to $8N$, the maximum depth change is indicative of the nature of the region. The strength of a material (modulus of elasticity) is defined as the ratio of stress to strain as given by Equation 2.20. The stiffness of an object is the extent of the resistance to deformation when a force applied. In this experiment since the user controls the maximum force applied to the point, the change in deformation is indicative of the material property as shown in Figure 2.24 where there is a difference in depth change for Region A and Region B. Region A being stiffer shows less change in depth than Region B.

$$E = \frac{\sigma}{\epsilon} \qquad (2.20)$$
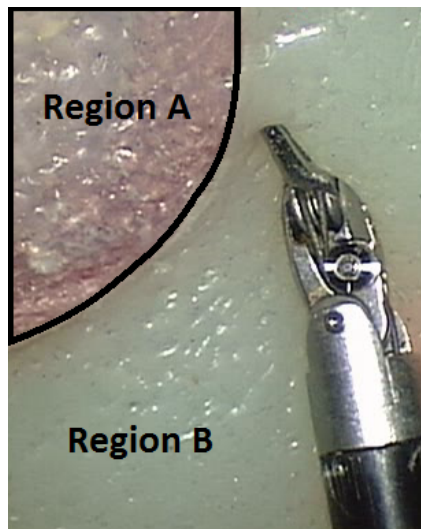
Figure 2.23: Section of the phantom outlining Region A and Region B captured from the da Vinci's® SEM.
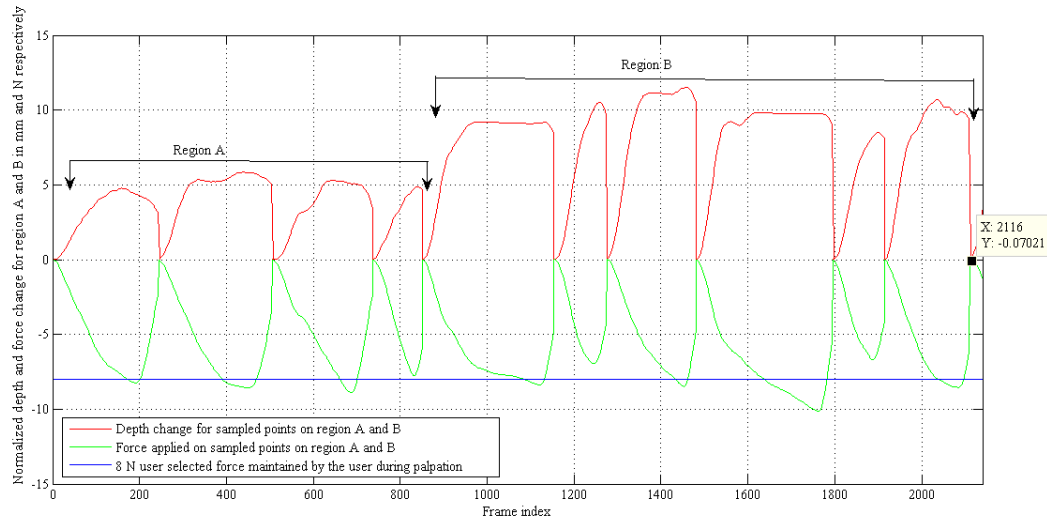
Figure 2.24: Variation in depth and force while palpating the points minimally invasively using the MTM to manipulate the PSM.

Figure 2.25 shows a plot of depth change versus the force applied to all 9 points palpated in Region A and Region B. There is a clear separating plane that distinguishes Region A from Region B after an application of a force of $4N$. This limit is dependent on the material property, and it can be deduced from this plot that property of the material in Region A can be differentiated from that of the material in Region B when the applied force is above $4N$, where the deformation change is able to characterize the points. Figure 2.26 also shows a change in depth versus force graph when the tool is removed from the surface point after the application of the maximum force. This plot is sparse in nature as the tool was removed quickly after the maximum force was applied. The rate at which the points were pressed was relatively slower then when the force was released. These points are also well separated into groups for Regions A and B. These plots indicate that the K-means clustering algorithm would be able to partition these feature vectors accurately.
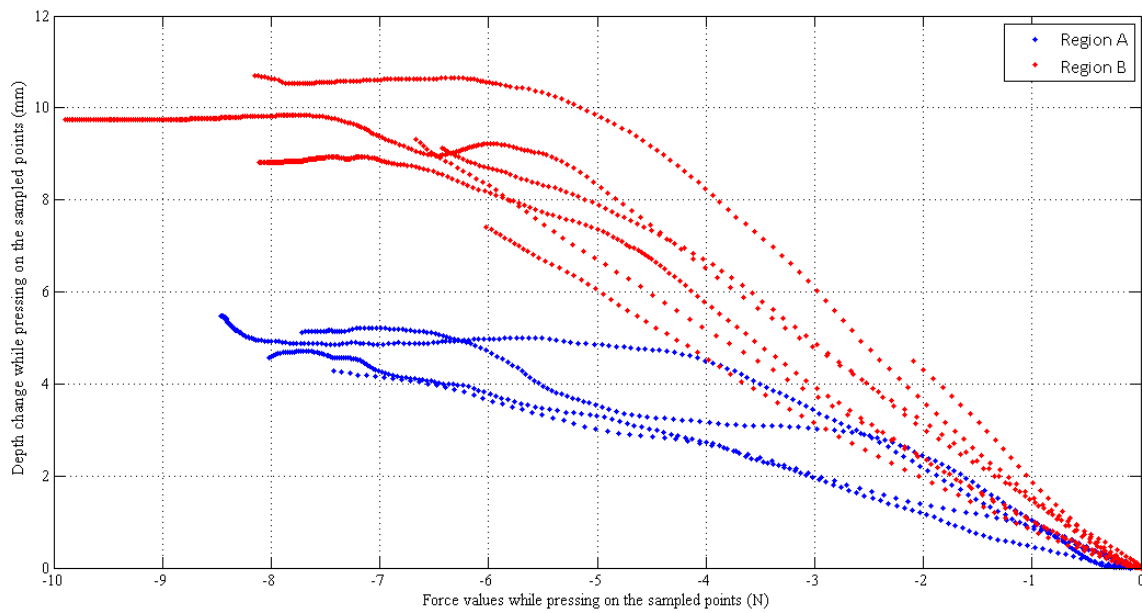
Figure 2.25: Depth change versus applied force plot while pressing the tool-tip on the phantom during RAMIS
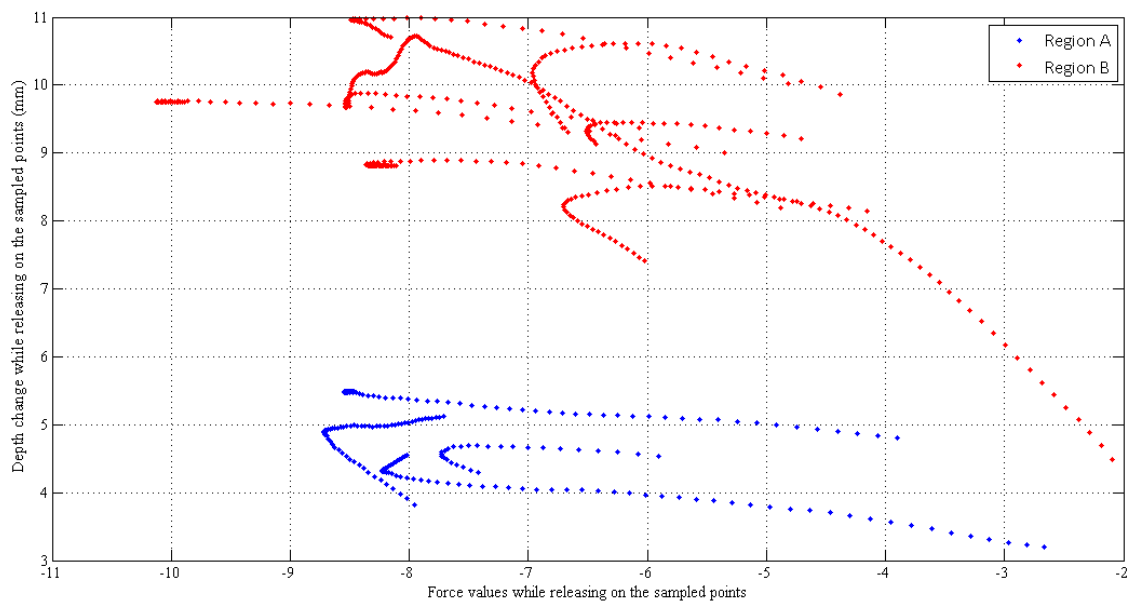
Figure 2.26: Depth change versus applied force while releasing the tool-tip from phantom during RAMIS

We conducted the same experiment on the phantom shown in Figure 2.27. This phantom was dyed with a brown color while it was prepared so that the user has no visual cue from color

difference about the phantom property while performing the experiment. The same experiment was performed and fourteen points were palpated teleoperatively. The K-means clustering resulted in the color-coded image shown in Figure 2.28. The figure shows the original boundary on the section of the phantom depicted in black color and the boundary estimated from the framework in red.

(a)

(b)
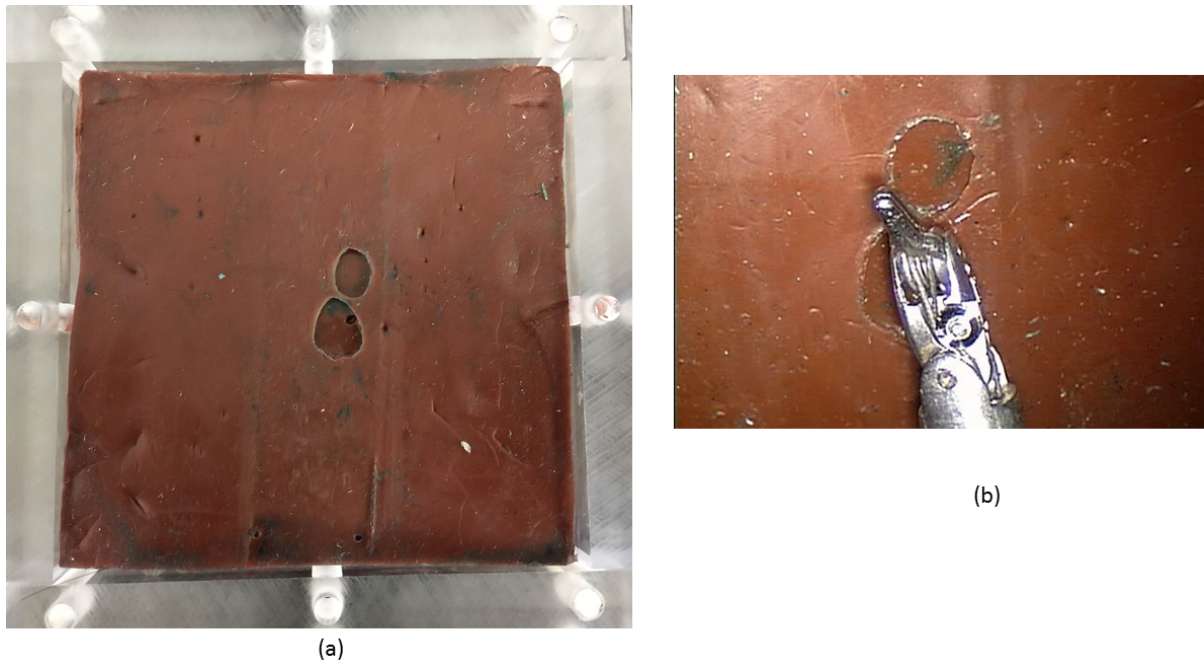
Figure 2.27: (a) Phantom designed for a blind test of the framework. (b) Image of the section of the phantom palpated, captured from da Vinci®'s SEM.

Figure 2.28: The original tumor boundary and the tumor boundary computed using the framework for a palpated section of the phantom.

We presented an algorithm for color coding the areas of higher stiffness in Section 2.5.2. The performance of the algorithm was evaluated on a phantom with a single continuum stiffer area. This algorithm could show variable performance if there are tumors present in small discrete lumps or in arbitrary shapes and also could give a false detection in the absence of a tumor. All four cases are enlisted in the Figure 2.29 and show the performance of the algorithm in these test scenarios. In the case of a spherical or arbitrarily shaped tumor, the algorithm can be used to provide a maximal spanning boundary that will include some portion of healthy tissue as the algorithm is not designed to closely fit to the tumor boundary. In the presence of multiple small tumors, the algorithm could define a boundary that spans all of the tumors including the non malignant regions between the tumors. The K-means clustering algorithm works with prior knowledge about the number of clusters. We chose this number to be two for identifying regions with high and low stiffnesses. In the absence of a tumor in the palpated region, K-means clustering will still attempt to differentiate the region into two groups. In such cases, the use of preoperative imaging data can reduce or eliminate the occurrence of false positives. Such data can help to determine the number of tumors as well as to better identify regions where they may be located. These regions can be palpated further using the framework proposed in this chapter to determine the maximal area spanned by each tumor. Information about the number of tumors from pre-operative imaging can be used to decide on the number of clusters.
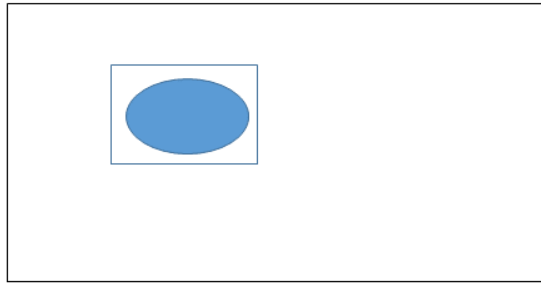
Case1 Oval shaped tumor, rectangle detection
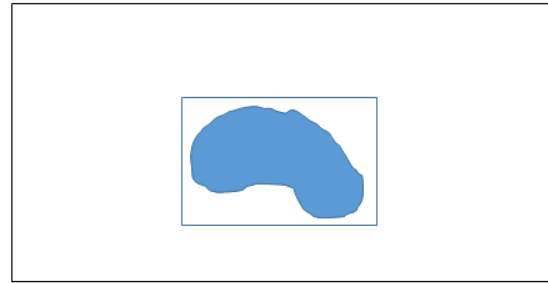
Case2: Arbitrarily shaped tumor , rectangle detection
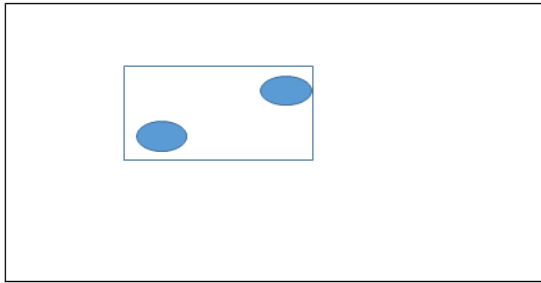
Case3 Multiple tumor not spatially continuous , erroneous detection
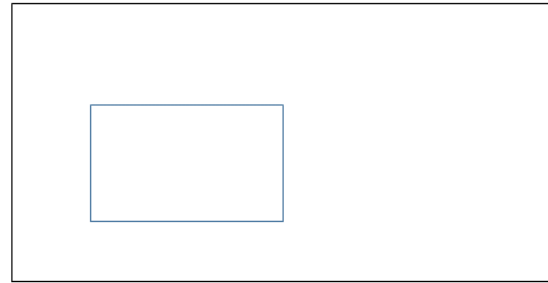
Case4 No tumor, false detection

Figure 2.29: Theoretical performance of the framework on varied test case scenario.

2.7 Conclusions

This chapter presented a framework for delineation of the hard object from soft material using a color-coded map. A da Vinci® needle driver sensorized with thin FBG sensors for force sensing was used, with an SEM for vision-based tracking. This design makes the framework a suitable candidate for integration in an actual surgical scenario without any major modifications to the existing procedure. This framework does not require any additional tool or extra trocar port. The framework though can still be improved to present a fitted boundary that deforms exactly to define the palpated points. A segmentation approach with an active contour model such as snakes [88] needs to be integrated with the existent framework for arbitrarily shaped tumor boundary recognition. We also propose to semi-automate this whole procedure to control the rate at which the tool presses onto the point for much more accurate results elim-

inating any inaccuracies that may result from a manual approach. Future work will consist of incorporating this assistive framework with other intraoperative imaging modalities such as US and tactile sensing to render a virtual image overlayed onto the actual surgical scene for additional guidance.

Chapter 3

Tissue Stiffness Characterization Based on Lung Surface Tracking

3.1 Motivation

In Chapter 2, we presented an assistive framework that is capable of characterizing tissue stiffness minimally invasively and telemanipulatively when using the da Vinci® surgical robotic system. We tested the performance of the framework on a stationary silicon phantom with materials of different stiffnesses. The surgical scene during an *in-vivo* thoracic surgery is much more challenging due to the presence of motion in the lung surface. The target lung is deflated during a VATS for stabilizing the surface [89]. Even after the lung has been deflated, it still has some motion due to the motion of the contralateral lung and the heart as the patient lies in a lateral recumbent position. The framework should be designed to take into account this motion or at least its most dominant part (frequency), i.e., the frequency of respiratory motion. Also, robot-assisted palpation should be performed either at the rising part of the breathing cycle or at the falling part of the cycle to produce consistent results. In other words, a more realistic validation experiment should involve a phantom which can simulate the motion of a deflated

lung. Therefore, surface tracking is an important upgrade to the current framework that will enable us to compute the relative indentation incorporating the motion of the tissue surface as well as the interacting tool.

Deflation of the lung results in a significant shift in the location of a tumor that is initially obtained from preoperative CT scans. Also, an intra-operative imaging modality such as the US produces distorted images of the lung because some amount of air is still present in the alveoli of the deflated target lung which interferes with the reflected US signals. Naini *et al.* [90] developed an algorithm that can determine the lower and the upper-intensity thresholds for segmenting lung's soft tissue and air-volume based on the principles of tissue incompressibility and lung's air-mass conservation. They analyzed the histograms obtained from the CT scans at different inflation and deflation levels during respiration for initializing the threshold value. Then using an optimization algorithm, they computed the optimum threshold values for segmentation. Naini *et al.* [20] further worked on computing the global and the local non-rigid registration transformations between CT scans at different deflation levels. The resulting transformation parameters were optimized to compute a new set of parameters that can extrapolate the CT scan of the deflated lung. These results assist in constructing a CT scan for a deflated lung which can determine the location of the shifted tumor thereby enhancing surgical planning. Even after the tumor location is predicted intra-operatively by extrapolation, the lung tissue surface still undergoes deformation due to continuous cardiac and respiratory motions which can introduce further uncertainty during a surgical procedure. Therefore, it will be helpful to provide visual cues to the surgeon about the real-time movement of a tumor during VATS.

Apart from providing visual guidance, surface motion information can also act as an input to the robotic control system so as to cancel any active relative motion between the tool and the lung surface to achieve more accurate control. One such example is the macro-micro robotic tool designed to compensate for tissue surface motion in [91]. Lung Brachytherapy is relatively a new minimally invasive interventional procedure where a radioactive seed is placed in a tumor

using a hollow needle at a target location to destroy cancerous cells effectively. This procedure is dependent on an accurate needle placement and steering. Farokh *et al.* [91] designed an impedance controller that synchronizes the motion of a needle with the motion of a lung for effective seed placement. In such a motion stabilizer and synchronizer design, estimation of lung surface motion can act as an important feedback for the controller. In this chapter, we present a tracking algorithm, designed to obtain the data about the motion of the lung surface intraoperatively. The tracking algorithm is implemented on *in-vivo* VATS videos for analyzing the motion of distinctive landmarks on the surface of a lung. We also extend this study by evaluating the motion of a pair of distinctive landmarks to compute the lateral strain variation between them over the period of the experiment. This analysis helps us to validate an important property that the lateral strain variation can characterize tissue stiffness.

3.2   Literature Review

The literature regarding tracking can be broadly classified into feature-based tracking, contour-based tracking, region-based tracking and learning-based tracking. Feature-based tracking generally employs feature descriptors such as SURF [74], SIFT [92] and BRISK [75] for matching a distinctive feature in consecutive images. All of these correspondence matches are scale and rotation invariant, hence they are robust for accurate tracking if the feature is repeatable and stable in consecutive frames. These feature matching algorithms are robust but time-consuming because the matching method is followed by an outlier rejection mechanism which is a computationally intensive task.
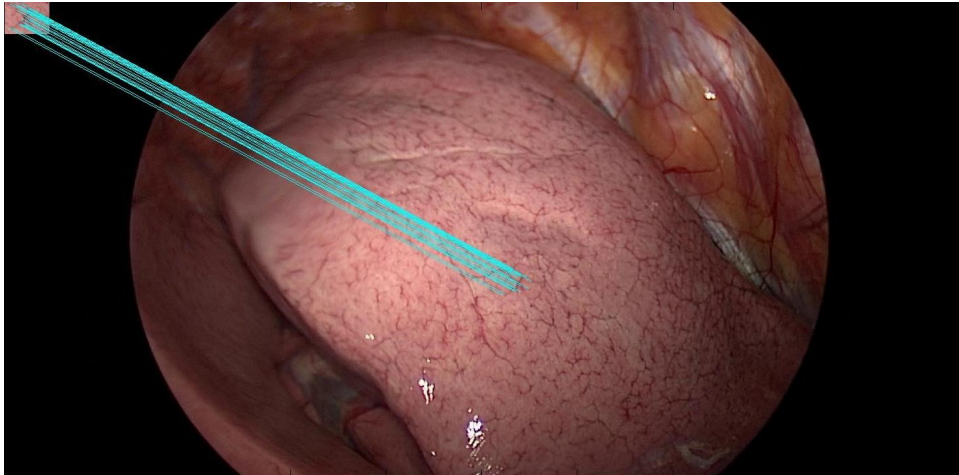
Figure 3.1: Matching for a single feature using SIFT.

The contour-based matching algorithms are essentially modeled to match the contour's segment in consecutive images. These algorithms employ prior knowledge of the structure. Seminal work in these matching algorithms indicates a dependence on the extraction of an edge map using either motion or texture cue [93]. The matching algorithm most commonly used for contour matching is chamfer matching that computes the amount of similarity between two contours to recognize the best match [94]. Several variations of basic contour matching algorithms are available such as using the generalized Hough transform instead of chamfer matching for analytic curve detection [95] or using a multi-scale contour detection that detects contour fragments by defining a basis function that optimizes for the correct match [96].

Region-based matching algorithms determine the overlap corresponding to the best match that minimizes the difference. Similarity metrics used for the computation of the best match in the literature are Sum-of-Squared Difference (SSD), Normalized Cross Correlation (NCC) and Mutual Information (MI). Non-rigid region-based registration uses a mesh that fits the deformed surface initially, the mesh is iteratively registered to the surface based on the transformation of the control points [97]. Template-based matching is a region-based matching algorithm where a template with a specific feature is searched in the base image using the similarity metrics such as SSD, NCC and MI. We have used template-based matching for our purpose that estimates the best similarity match corresponding to the defined template. Since,

it can be time-consuming to do the match for the whole image, we choose a small search space of width (w) and height (h) and predict its motion over time using a Kalman Filter (KF) for an accurate estimation.

The template matching algorithm has evolved over time to be a better tracking estimator by implementation of corrections for occlusion detection and continuous template update for scale and rotation changes. Matthew *et al.* [98] designed a drift correction update algorithm that constantly updates the template in each iteration by registering the old template to the current template. Schreiber *et al.* [99] suggested a more robust variation by using adaptive weights for each pixel that maximize the effect of dominant data along with the use of a drift correction algorithm. Chambers *et al.* [100] designed a higher learning based template matching algorithm that learns the template parameters. In our work, we have used template matching with a continuous template update for scale and rotation changes as explained in Section 2.3.1.3.3. The KF estimates the shift in search space with higher accuracy that significantly improves the computation time by reducing the size of the search space for searching the template.

3.3   Lung Surface Tracking Algorithm

Analyzing the shape changes, the deformations and the temporal motion from a sequence of image data acquired using an endoscope is a challenging problem. Seminal works in tracking a deformable surface involves a mathematical template mesh model that is deformed over time with the motion of the control points [101]. In [102] [103], the authors suggested attaching artificial fiducial markers over the surface of a beating heart to assess the displacement of those markers. Though these markers, if fixed properly to the tissue surface, can result in a robust detection, the problem lies in designing a fixing mechanism. Also, it is not deemed safe to introduce miniaturized markers in the body due to the possibility of dislocation and other complications such as bio-compatibility of the material used in the fabrication of those markers. Ortmaier *et al.* [104] suggested tracking of a distinctive biological landmark on the

heart surface. Lung surface also has prominent markers and especially, a smoker's lung will contain scavenger cells filled with the impurities absorbed from smoking. Black lines in case of a smoker as shown in Figure 3.13, and red arteries-based biological markers in case of healthy tissue as shown in Figure 3.2, can serve as a template for tracking the lung surface. The basic flow of the tracking algorithm is depicted in the Figure 3.3.
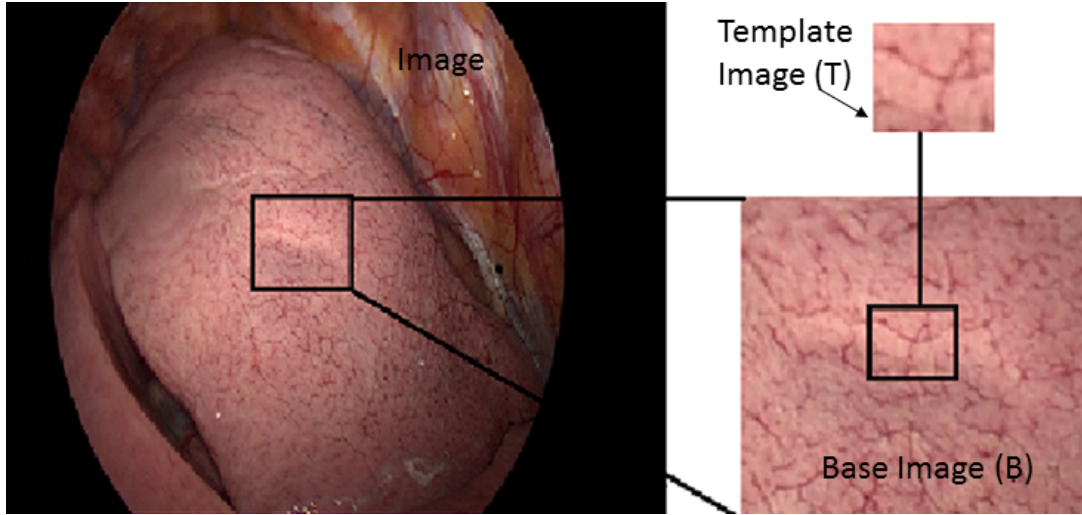
Figure 3.2: A user selected template on a lung's surface.
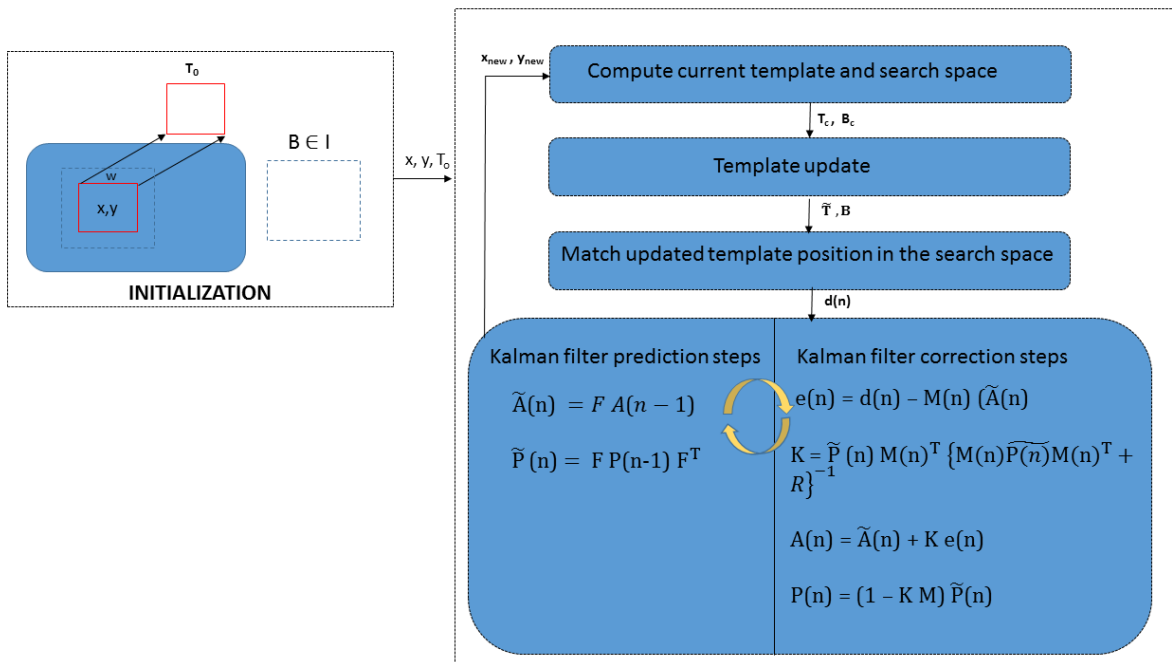
Figure 3.3: Basic block-diagram of the tracking algorithm.

We have used a template-based matching algorithm for tracking multiple biological landmarks on a tissue surface. The user manually selects a biological marker with a mouse click (x, y) as depicted in Figure 3.2 and the initial template (T_0) image of size 50×50 is computed and saved in the memory. The template is updated recursively using the Template Update method defined in Chapter 2 in Section 2.3.1.3.3. This template update procedure resolves the issue of drifting over the search space that contains the marker with time. To localize the presence of a template in the computed search space (B), the updated template is matched within the computed search space using the Matching Method explained in Chapter 2 in Section 2.3.1.3.4. The Matching Method returns the location of the center of the template found in the search space corresponding to the best match from the NCC similarity metric analysis. This approach can be repeated for multiple templates, and these selected templates can be localized simultaneously over the period of the experiment. The motion of a deflated lung is acquired from the motion of the heart and the contralateral lung which is quasi-periodic in nature. Therefore, the motion of the deflated lung surface is inherently periodic, and the Fourier theorem states that any periodic signal can be expressed as a sum of sine and cosine functions with different amplitudes and harmonic frequencies. Richa *et al.* [105] have utilized this theorem to model the 3D coordinates of the motion of the heart surface as a dual non-stationary Fourier series. The unknown parameters of the series were recursively estimated using the Extended Kalman Filter (EKF). The motion of the deflated lung is an amalgamation of the periodic motion of the heart and the contralateral lung. By incorporating this knowledge about the bandwidth of the motion signal of the deflated lung, we have modeled the motion as a truncated Fourier series with unknown amplitudes. We discuss various filters that were used in literature to model a reference signal with prior knowledge about the frequency and bandwidth to estimate the unknown weights. We have chosen a KF with a motion model that incorporates the knowledge of the bandwidth of lung's surface motion and it adapts to the determine the unknown weights recursively by using the motion model as well as the value obtained from the tracking algorithm.

3.3.1 Fourier linear combiner

Fourier Linear Combiner (FLC) is an adaptive filter that can estimate the Fourier series coefficients of a quasi-periodic signal of a known frequency using Least Mean Squares (LMS) [106]. If the base-frequency of a signal is known *a priori*, FLC can adapt to the amplitude and phase change of the signal. The reference signal equation is as follows:

$$Y(n) = \sum_{r=1}^{M} a_r \sin(rw_{known}n) + b_r \cos(rw_{known}n) \tag{3.1}$$

where M is the number of harmonics in the model.

3.3.2 Band Limited Multiple Linear Fourier Combiner

The FLC filter assumes a fixed known base frequency for modeling the unknown signal, but our problem involves a motion signal within a frequency band. The motion signal of the deflated lung has a lower cut-off frequency corresponding to the frequency of respiration and an upper cut-off frequency corresponding to cardiac motion. Band Limited Multiple Fourier Linear Combiner (BMFLC) is a filter that estimates the modulated signal within a band of frequencies. BMFLC can be considered as the combination of multiple FLCs for the estimation of a band-limited signal. The reference signal is modeled as a weighted sum of sine and cosine functions at specific frequencies within the band, given by Equation 3.2 which is defined in [107]. A user can decide the number of divisions in the bandwidth for the signal modeling.

$$Y(n) = \sum_{r=0}^{L} a_r \sin(2\pi(f_0 + \frac{r}{P})n) + b_r \cos(2\pi(f_0 + \frac{r}{P})n) \tag{3.2}$$

where P is the number of divisions in the frequency band and $L = (f - f_0)P$.

3.3.3 Kalman Filtering

The motion of the template on the surface of the deflated lung is modeled similarly to the reference signal for the BMFLC defined in Equation 3.2. The motion model of the displacement between the consecutive frames, also known as the inter-frame motion, is modeled as the weighed sum of sine and cosine functions at sampled frequencies in the bandwidth. The deflated lung surface motion's bandwidth has the lowest frequency component corresponding to the respiration of $0.16Hz$ and the highest frequency component corresponding to the cardiac motion of $1.83Hz$. The higher the number of sampled frequencies within the bandwidth the better will be the motion model work to estimate the inter-frame motion. The motion model for the surface motion is defined as the weighed sum of the sine and the cosine at five dominant frequencies $(0.16Hz, 0.58Hz, 1Hz, 1.42Hz, 1.83Hz)$ as shown in Equation 3.3. The KF adapts to the changes in the phase and amplitude of the inter-frame motion and predicts the shift for the next frame. The KF is an estimator that recursively predicts the future state of a dynamic system model that may be excited by random disturbances and noisy measurements. It is a minimal variance estimator for a linear time varying system containing statistical noise and other inaccuracies. The motion model of the template's movement in 2D space (x, y) is defined using a linear time varying equation described as follows:

$$
\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \sum_{r=0}^{4} a_{1r} \sin(2\pi(f_0 + \frac{r}{P})n) + b_{1r} \cos(2\pi(f_0 + \frac{r}{P})n) \\ \sum_{r=0}^{4} a_{2r} \sin(2\pi(f_0 + \frac{r}{P})n) + b_{2r} \cos(2\pi(f_0 + \frac{r}{P})n) \end{bmatrix} \tag{3.3}
$$

Let the 2D motion vector at time instant 'n' be denoted as:

$$
d(n) = \begin{bmatrix} x \\ y \end{bmatrix} \tag{3.4}
$$

The measurement equation can be given by Equation 3.5, where M is the transformation matrix that maps the state vector into the measurement domain and is given by Equation 3.6.

$$d(n) = M(n)A(n) + \lambda(n) \tag{3.5}$$

$$M(n) = \begin{bmatrix} sin(w_1 n) & cos(w_1 n) & \cdots & sin(w_5 n) & cos(w_5 n) & 0 & \cdots & 0 \\ 0 & \cdots & 0 & sin(w_1 n) & cos(w_1 n) & \cdots & cos(w_4 n) & sin(w_5 n) & cos(w_5 n) \end{bmatrix} \tag{3.6}$$

KF works in a state-space representational framework where a set of first-order differential equations relate the input, output and state variables. The amplitude also known as the weight of the truncated Fourier series is the state vector. The state update equation is defined in 3.7. Since, we have no prior information of the weights, it is described as a random walk [108]. KF uses noisy measurements and inputs corrupted with white noise to estimate the unknown variable of the system. The measurement error $\lambda(n)$ and the state error $\mu(n)$ are assumed to be uncorrelated, zero mean Gaussian white noise as described in Equation 3.9 and 3.10 respectively.

$$\tilde{A(n)} = A(n-1) + \mu(n) \tag{3.7}$$

where $A(n)$ is a 20×1 column vector containing all the unknown weights.

$$A(n) = \begin{bmatrix} a_{11}(n) & b_{11}(n) & \cdots & b_{15}(n) & a_{21}(n) & b_{21}(n) & \cdots & a_{25}(n) & b_{25}(n) \end{bmatrix}^T \tag{3.8}$$

$$\mu \approx N(0, Q) \tag{3.9}$$

$$\lambda \approx N(0, R) \tag{3.10}$$

$$R = 0.01 \times I_2 \ \ and \ \ Q = 0.1 \times I_{20} \tag{3.11}$$

where R is the measurement error covariance whose value is estimated by studying the variance of the measurement error signal $\lambda(n)$ around its mean value and Q is the state error covariance whose value is empirically set higher than R to give higher weight to the values obtained from the camera measurements.

3.3.3.1   Prediction Step

This step computes the new state estimate at every discrete time by applying a linear operator to the current state based on the state equations defined in 3.12, 3.13. $A\tilde{(n)}$ and $P\tilde{(n)}$ represents *apriori* estimates of amplitude and covariance at time step n.

$$A\tilde{(n)} = FA(n-1) \tag{3.12}$$

where F is the state transition matrix and defined as I_{20}.

$$P\tilde{(n)} = FP(n-1)F^T \tag{3.13}$$

3.3.3.2   Correction Step

This step involves the combination of current observation with the state estimate obtained from the prediction step to refine the value of the state vector. The calculation of innovation $e(n)$ in Equation 3.14 refers to the deviation of the measured variable from the estimated variable. Kalman gain (K) computation is defined in Equation 3.15 and this value is used to correct the predicted state vector and predicted covariance matrix in Equation 3.16 and 3.17 respectively. KF can effectively determine the next state by not only taking into account the process model, but also giving different weighing to measurements and process values by adjusting the Kalman gain (K).

$$e(n) = d(n) - MA\tilde{(n)} \tag{3.14}$$

$$K = P\tilde{(n)}M(n)^T\{M(n)P\tilde{(n)}M(n)^T + R\}^{-1} \tag{3.15}$$

$$A(n) = A\tilde{(n)} + Ke(n) \tag{3.16}$$

$$P(n) = (1 - KM(n))P\tilde{(n)} \tag{3.17}$$

3.4 Results

The proposed tracking algorithm in Section 3.3 is validated on a video obtained from an *in-vivo* VATS, conducted by a thoracoscopic surgeon at LHSC, Dr. Richard Malthaner. Figure 3.4 shows a snapshot of the video with four selected templates that are tracked iteratively. The lung is deflated but still has motion due to the contralateral lung and the heart. Figure 3.5 and 3.6 shows the displacement signal of these four templates in x and y directions respectively. Figure 3.7 shows a plot of the power spectral density for template T_1's motion in the x-direction, and it has peaks at frequencies corresponding to the bandwidth of respiratory motion and cardiac motion. The peaks for dominant frequencies in the signal are at $0.17Hz$, $0.35Hz$ and $1.64Hz$ where the first two highlighted peaks have frequencies indicating the motion corresponding to the lung and third highlighted peak indicates the motion due to the heart. The inter-frame displacement prediction from the KF improves the tracking capability by reducing the search space image window size from 500×500 to 150×150. Figures 3.8 and 3.10 show the actual position of the template motion in the next frame computed using the camera and the motion predicted using the KF considering the motion model and the computed values for the current frame. The search space is now initialized with the center position predicted by the KF in the next frame. This reduces the computational complexity and the chance of false detection which can occur when the search space is not updated for the movement over time. Figure 3.9 and 3.11 report the Normalized Root Mean Square Error (NRMSE) for the difference in the shift computed using the KF and measured by the camera for x and y directions respectively.
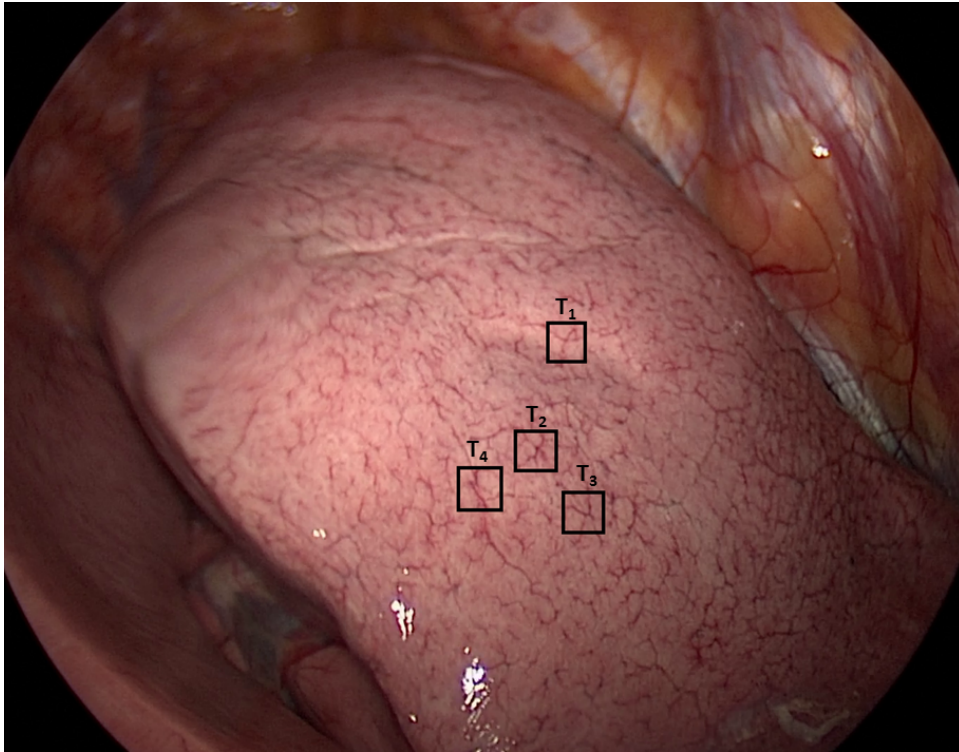
Figure 3.4: A snapshot from the video of a VATS procedure on a deflated lung with the selected templates depicted in the box with their respective notations.
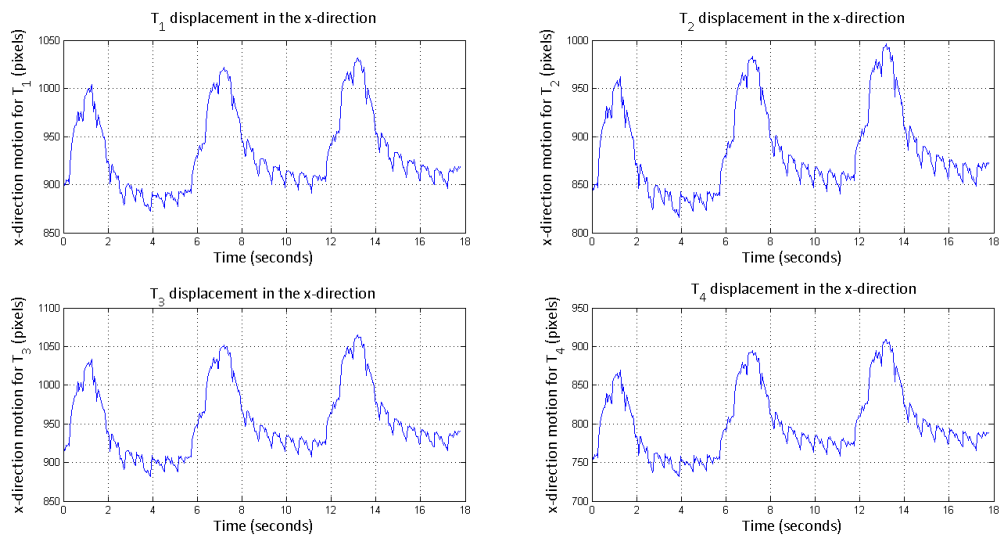
Figure 3.5: The displacement of the template in the x-direction for T_1, T_2, T_3 and T_4 respectively.
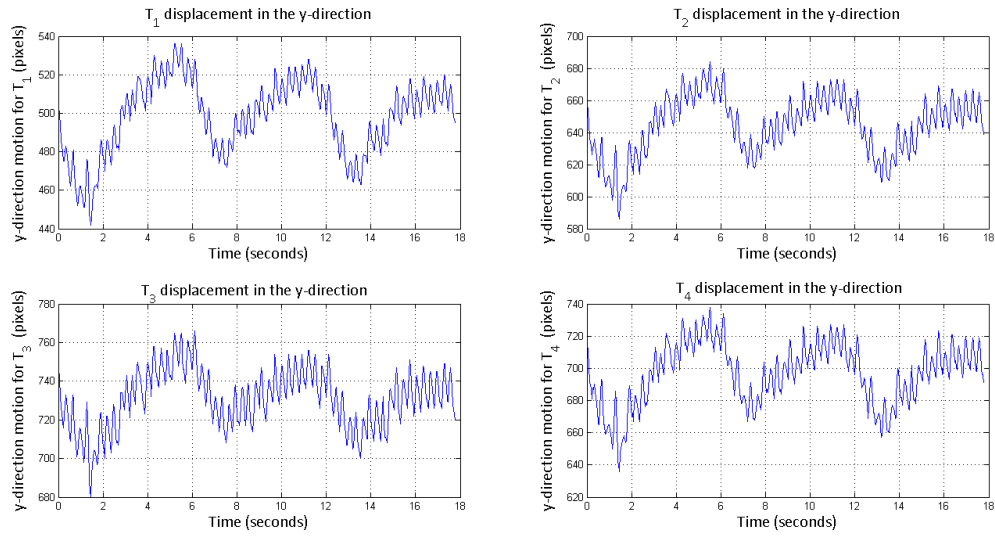
Figure 3.6: The displacement of the template in the y-direction for T_1, T_2, T_3 and T_4 respectively.
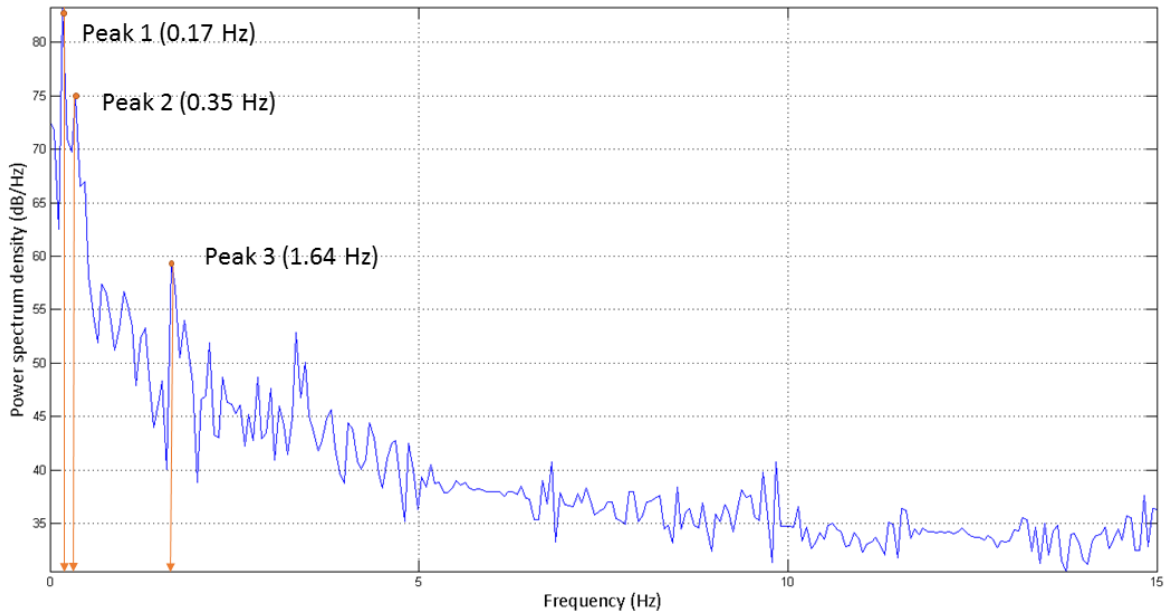
Figure 3.7: Power spectral density plot for T_1 in the x-direction showing peaks corresponding to the bandwidth of respiratory and cardiac frequencies.
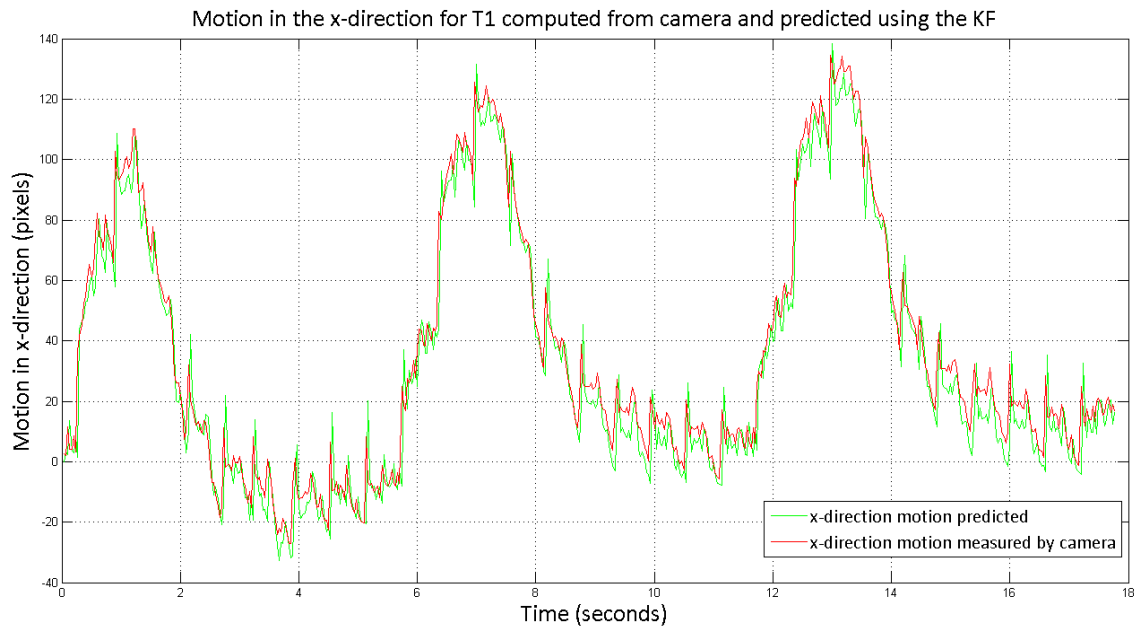
Figure 3.8: The motion of T_1 in the x-direction measured using the camera and predicted by the KF.
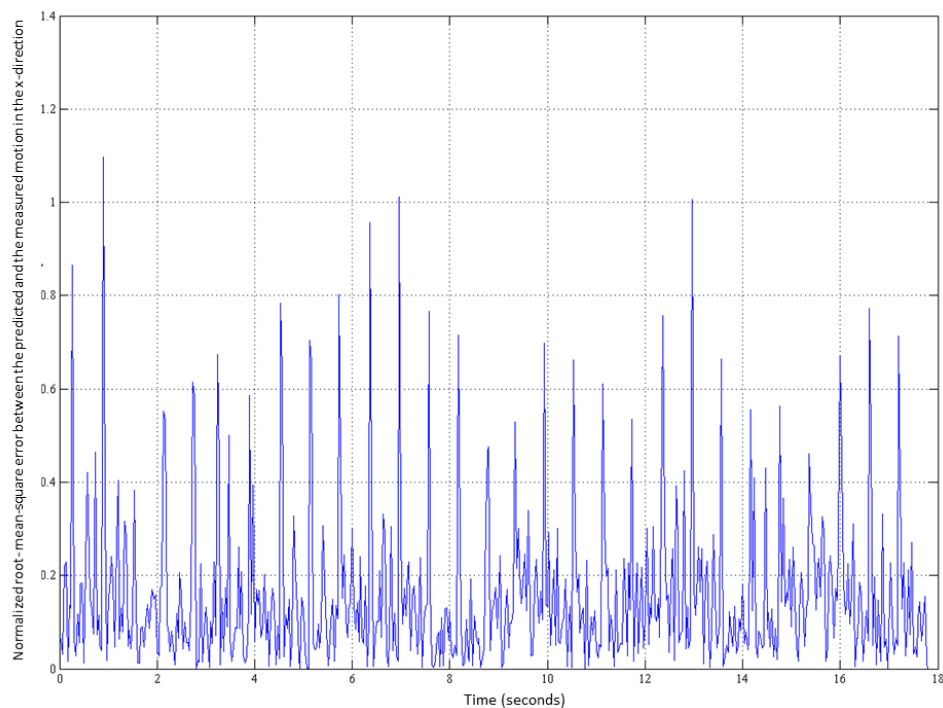
Figure 3.9: NRMSE for motion prediction in the x-direction using the KF.
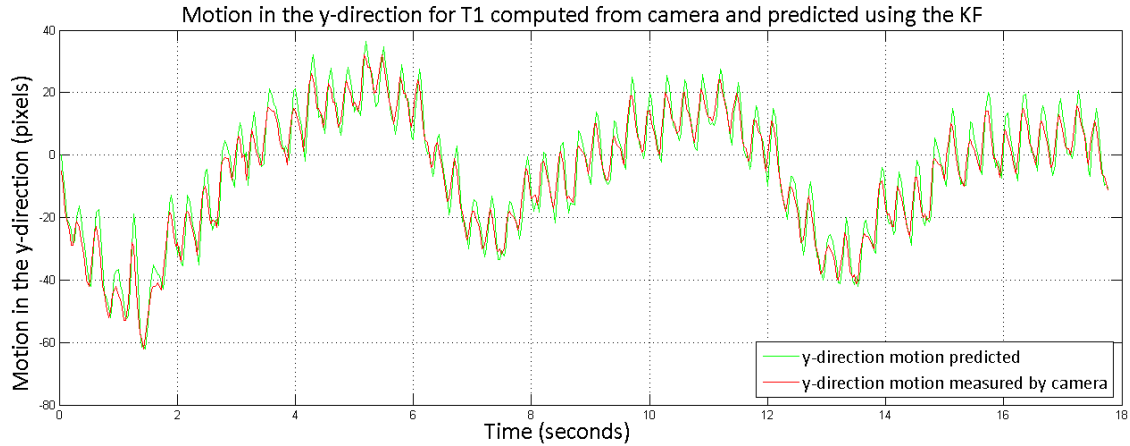
Figure 3.10: The motion of T_1 in the y-direction measured using the camera and predicted by the KF.
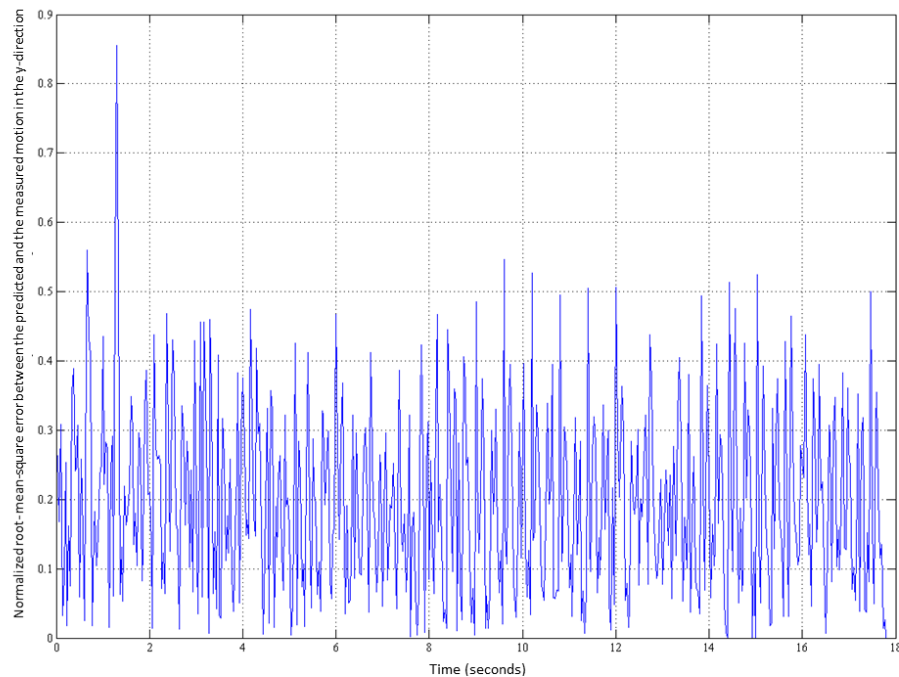
Figure 3.11: NRMSE for motion prediction in the y-direction using the KF.

Strain is a dimensionless quantity that is defined as the ratio of the change in dimension to the original dimension when the body is subjected to stress. The strain is generated when there is a change in the spatial domain state such as length, volume or angle. When normal stress is applied on a material, it results in a change in length in the direction of the force which is termed linear or longitudinal strain whereas if the change appears in the direction perpendicular

to the force it is termed lateral strain. In Chapter 2, we characterized stiffness based on the measurement of the longitudinal strain when a pressing normal force is applied to a set of sampled points on the phantom surface. The motion of a lung is associated with a continuous change in its volume that generates a lateral strain on the surface. Therefore, the measurement of the lateral strain can assist in characterizing the tissue surface. The tumorous tissue is stiffer than the surrounding non-malignant tissue, and it has been experimentally validated for breast tissue in [109]. Tumorous tissue being stiffer than the surrounding healthy tissue should ideally show more resistance to deformation and stretch less. To verify this, a pair of points on the tumorous region and two pairs of points on the non-tumorous region at an increasing distance from the tumorous region are sampled. Figure 3.12 shows the deflection curve in black that stretches and contracts due to the motion of the lung surface. We chose a small element ds on the deflection curve between a pair of points (p_1, p_2) and observe the curvature variation at different time instances. The chosen deflection curve between template pair is a small fraction of the deflection curve spanning the lung surface therefore the arc-length can be approximated as the Euclidean distance between the points (p_1, p_2). The tracking algorithm described in Section 3.3 returns the best match corresponding to the center of the template found within the search space. The computation of centers for a pair of templates over time can assist in evaluating the variation of the arc-length between the two chosen templates. The strain is assumed to be zero at the initial configuration, i.e., $ds(n_0)$. The Lagrangian strain (ϵ_L) measure for the computation of lateral strain between two points at a given time instant (n) is given by Equation 3.18. Equation 3.19 shows that the lateral strain is proportional to the inter-template distance with respect to the initial distance between the points. This inter-template distance variation with respect to the initial distance between the pair ($\beta(n)$) can be used as a measure to differentiate regions with different stiffnesses.

$$\epsilon_L(n) = \frac{ds(n)}{ds(n_0)} - 1 \tag{3.18}$$

$$\epsilon_L(n) = \beta(n) - 1, \, \epsilon_L(n) \propto \beta(n) \tag{3.19}$$

where

$$ds(n) = \sqrt{(p_1 x(n) - p_2 x(n))^2 + (p_1 y(n) - p_2 y(n))^2}$$

$$ds(n_0) = \sqrt{(p_1 x(n_0) - p_2 x(n_0))^2 + (p_1 y(n_0) - p_2 y(n_0))^2}$$
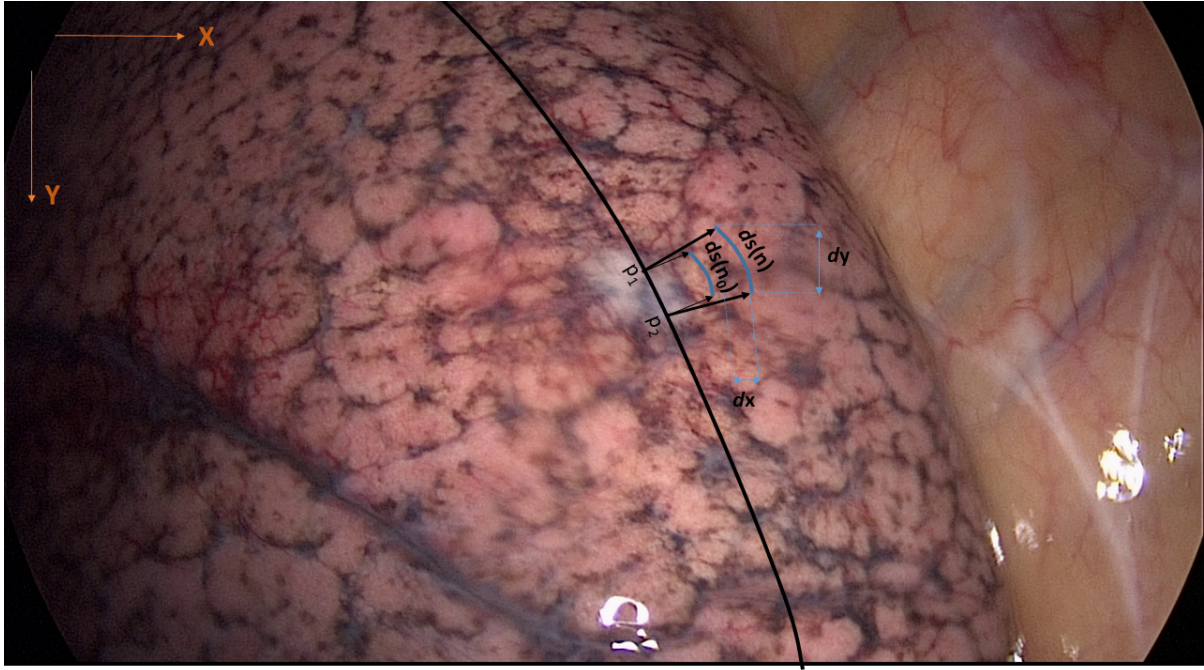
Figure 3.12: A snapshot from a video of a VATS, indicating the deflection curve and a small element ds between a pair of points (p_1, p_2).

A region with tumor will have high stiffness and therefore will stretch less than the surrounding non-tumorous region with low stiffness. Figure 3.13 shows a snapshot from a video of an *in-vivo* VATS performed on a human lung. In the figure, the tumorous region can be seen as the tissue with the distinctive gray color and a template pair (T_5, T_6) is placed on the tumor surface. $T_{1...4}$ lie on the nonmalignant surface where template pair (T_1, T_2) is placed farther from the tumor than the pair (T_3, T_4). The variation of the lateral strain for the selected template pairs over time is depicted in Figure 3.14. This plot concludes that the lateral strain variation between pairs of templates can characterize tissue stiffness. To do a comprehensive analysis, we sampled template pairs along a line that does not pass through the tumor to verify whether the lateral strain variation is due to the inherent motion of the lung surface or due to the presence of a stiffer region. Figure 3.15 shows the line along which the sampled template pairs

are selected, and the results of lateral strain variation among the template pairs are depicted in Figure 3.16. Figure 3.16 shows that lateral strain variation signal does not show a significant difference in the absence of a tumor whereas Figure 3.14 shows that the variation of the signal corresponding to tumor is damped compared to the signal for the healthy tissue. Therefore, we can conclude that the analysis of motion of the tissue surface can characterize tissue stiffness by analyzing the change in lateral strain.
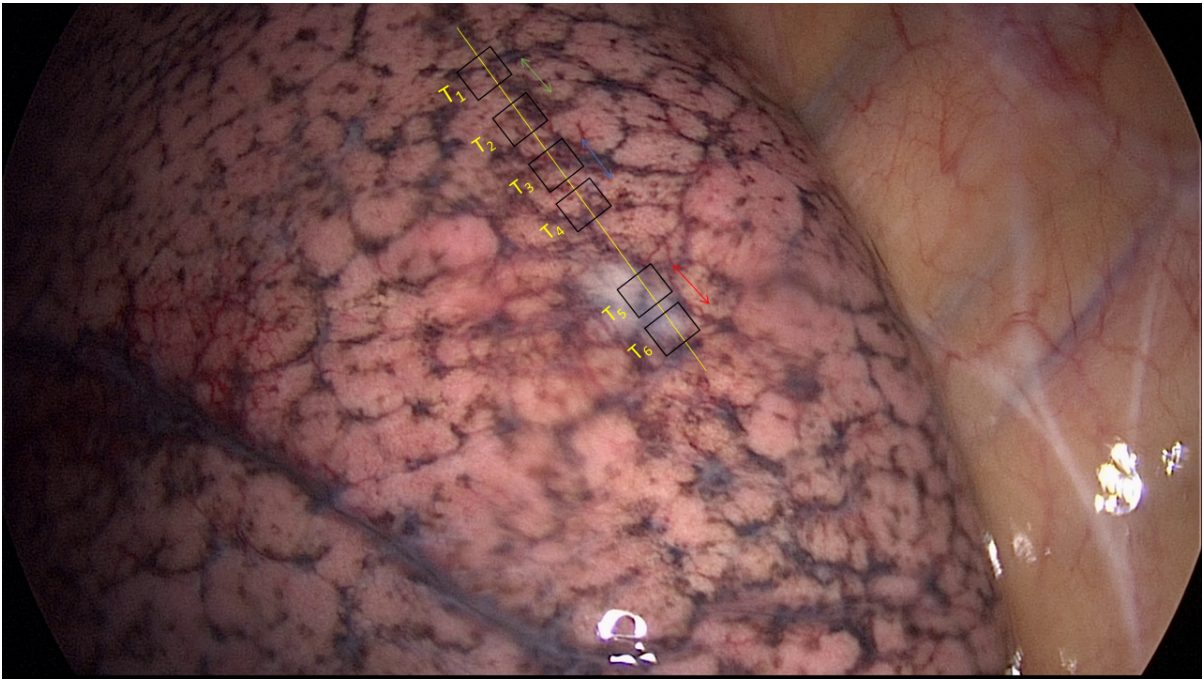
Figure 3.13: A snapshot from a video of a VATS, indicating the templates positions and their notations on the lung surface. The template pairs (T_1, T_2), (T_3, T_4) lie on the non-tumorous region and (T_5, T_6) lies on the the surface of the tumor.
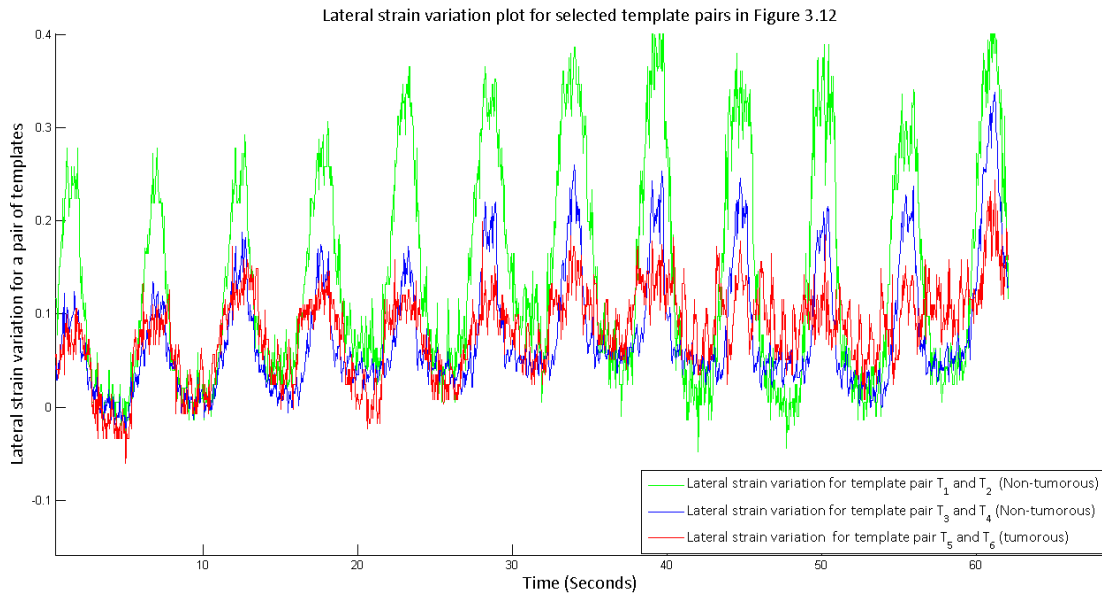
Figure 3.14: Lateral strain variation plot for the selected template pairs in Figure 3.13.
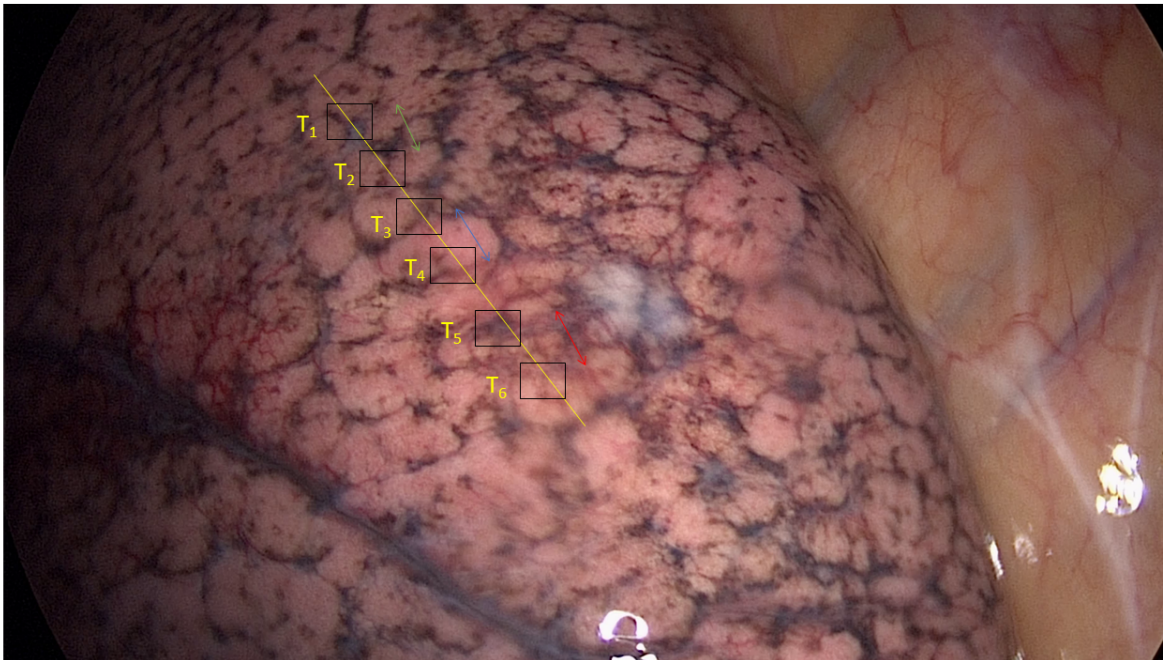
Figure 3.15: A snapshot from a video of a VATS, indicating the templates positions and their notations on the lung surface. The template pairs (T_1, T_2), (T_3, T_4) and (T_5, T_6) lies on the non-tumorous region.

Lateral strain variation plot for selected templat pairs in Figure 3.14
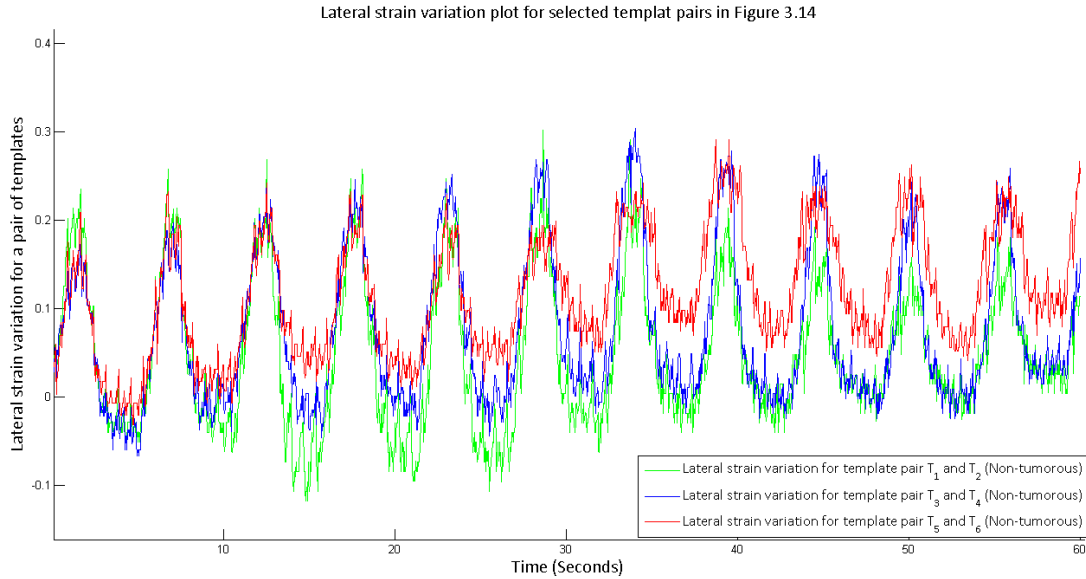
Figure 3.16: Lateral strain variation plot for the selected template pairs in Figure 3.15.

3.5 Conclusions

In this chapter, we presented a method for tracking multiple biological markers on the lung surface and implemented the tracking algorithm on videos obtained from an *in-vivo* VATS. The motion detected by the algorithm can serve as feedback for the robotic controller used to cancel the active relative movement between the tool and the tissue surface. We also verified a hypothesis that the tumorous surface will have a less lateral strain as compared to the strain of the non-tumorous surface. We concluded that the extraction and analysis of the 2D motion of the lung surface can be used to characterize tissue stiffness. The tumor localization mechanism based on this principle can be developed further by analyzing along multiple scan lines and sampled template pairs to localize the tissue surface that has a higher probability of being a tumorous region. The lateral strain determination will assist in narrowing down to the region that contains the tumor and later the suspicious area can be probed with the force-sensorized da Vinci® tool for more accurate localization of the tumor. Therefore, the surface tracking algorithm will serve as an important upgrade to the framework developed in Chapter 2 for the

determination of the most probable region that may contain a tumor. Also, the motion of the lung surface will interfere with the robot-assisted palpation motion suggested in Chapter 2. The integration of tool-tip tracking with lung surface tracking will make it possible to determine the relative depth change while pressing the surface. Despite being close to the manual palpation process, one of the major limitation of the probing device is the slow speed of palpation. The work presented in this chapter has the potential to accelerate the tumor localization process by helping to identify the probable area to be palpated instead of probing the entire surface.

Chapter 4

Conclusions and Future Work

This chapter concludes the research presented in this thesis on the topic of the development of an assistive framework for tumor localization in RAMIS. A brief summary of the key aspects, concluding remarks and possible extensions as future work are presented.

4.1   Summary

The sense of touch is a critical feedback for surgeons during open surgical procedures to localize tumors and assess tissue properties for surgical resections. To overcome the loss of this crucial information, researchers have tried to design frameworks that mimic hand-tissue interaction similar to manual palpation. The aim of this research was to develop a framework that can increase the probability of tumor localization in RAMIS where feedback from manual palpation is not possible. In Chapter 1, we reviewed different feedback modalities used to recover soft tissue stiffness information in the absence of manual palpation. We concluded from the literature review that a good framework should require minimal redesign to the existing setup and should be sufficiently miniaturized for use in RAMIS. Also, its design principle should mimic manual palpation, and provide fast and accurate results.

Based on the extensive review regarding the preferred characteristics of the probe design, a

vision-based tissue stiffness characterization framework for RAMIS was developed in Chapter 2. The framework was experimentally validated on a silicon phantom with regions of different stiffnesses, using a telemanipulation setup involving the da Vinci® surgical system and a force-sensorized da Vinci® surgical tool. An FBG force-sensorized da Vinci® needle driver was used for sensing the force applied to the phantom's surface as the tool was made to perform pressing motions on the tissue teleoperatively. The use of da Vinci® tool provides better control and higher DOFs, and this tool is also used for suturing. Therefore, no additional port or setup re-design is required to use it during teleoperated palpation. We also developed a vision-based tracking algorithm for 3D tool-tip tracking using the inbuilt da Vinci® stereo endoscope, for the measurement of indentation depths. The framework analysis depends on two parameters for characterizing the phantom stiffness: indentation depth and force applied. We concluded in Chapter 2 that when the amount of force applied to a set of sampled points is controlled, the indentation depth alone can characterize the phantom stiffness. The palpated points were differentiated into two groups based on stiffness using K-means clustering. The lower stiffness points were segmented to present the localized boundary of the tumor within the palpated region. This chapter presented a model for robot-assisted tumor localization using a standard da Vinci® tool by integrating force sensing, vision-based sensing, and machine learning.

The framework presented in Chapter 2, required a mechanism to compensate for the intraoperative motion observed on the lung surface during an actual surgical scene to ensure reliable and robust stiffness measurements. In Chapter 3, we presented an algorithm for tracking distinctive biological markers on the surface of the lung. The tracking efficiency was further improved by using the Kalman filter to predict the inter-frame motion. The proposed tracking algorithm was tested on the videos obtained from an *in-vivo* VATS. The lung surface undergoes a continuous volume change due to its motion which results in a strain on its surface. We computed the values of strain variations on the surface corresponding to different stiffnesses from an *in-vivo* VATS video with a tumor and validated an important result that the lateral strain variation can be used to characterize tissue stiffness. This characterizing parameter can be employed

to determine the region that is likely to contain a tumor. The probing mechanism can then be used to evaluate the relative stiffness of this region. The integration of surface tracking in the framework proposed in Chapter 2, will result in a robust framework that can be validated in an *in-vivo* experiment. The work presented in the thesis forms the basis for characterizing lung tissue stiffness by analyzing the longitudinal strain change during robot-assisted palpation and the lateral strain change due to the motion of the lung tissue surface.

4.2 Future Work

Though the framework proposed in this thesis has attempted to address the problem of tumor localization in a comprehensive manner, there is significant scope for further improvements and testing. Some suggestion for future work are outlined in this section:

1. The experiment designed in Chapter 2 can be performed in a controlled manner by automating certain segments of the task of robot-assisted palpation. We can make the process semi-autonomous and move the da Vinci® tool in a pre-defined number of discrete steps so that the amount of force applied can be controlled better. The automation will reduce the variability in the way that the user presses the tissue surface. Also, the relative orientation of the tissue and the palpation plane can be maintained while pressing the surface. A semi-autonomous approach will allow the surgeon to focus on other aspects of the surgical procedure.

2. The framework is designed to provide continuous kinesthetic feedback about tool-tissue interaction (pseudo-haptics). Tavakoli *et al.* [110] suggested that haptic feedback decreases the time required for palpating a tissue surface. Therefore, the visual force feedback while maneuvering an indenter tool can be replaced with haptic feedback at the master console. A better framework design would consist of a pressure-sensitive touch-pad that would create an immersive perception of interacting with either a tumor or non-tumorous soft tissue.

3. We used an unsupervised clustering algorithm, i.e., K-means clustering for grouping points into two groups based on their stiffnesses. This division can be considered as an initial partition which can be used to train a classifier that can classify any new palpated point into either of two groups. Also, instead of computing the maximal area spanned by the points with lower stiffness, we can segment a set of points to present a fitted boundary that deforms around the contour of each palpated point.

4. The lung surface undergoes a continuous volume change due to its motion which develops a strain on its surface. We showed that lateral strain variations can be used to characterize tissue stiffness. This algorithm can be further developed by evaluating strain variations for template pairs along multiple sampled parallel scan lines for detecting the pair that behave differently. For the detection of a confined region, the algorithm can be run iteratively for sets of parallel scan lines in different directions. This simultaneous analysis of template pairs along multiple scan lines will be a very time intensive process. It can be speeded up by using a GPU that can process the data in parallel along the scan lines. This framework can accelerate the palpation process by confining the search of a tumor to a narrow region. Also, this approach needs no marker or physical device for its implementation and can be performed using an endoscopic camera and software analysis for determining regions with low stiffness.

5. The framework developed in Chapter 2 and the approach suggested above can be combined to first confirm the probable region having one or more tumors and then palpated in the specified zone using the FBG-sensorized tool. The sequential analysis will significantly reduce the computational time because the probing device now has to palpate only the suspicious region instead of the entire region. This framework will need to be evaluated in phantoms where motions corresponding to respiration can be simulated.

Bibliography

[1] B. Stewart, C. P. Wild, *et al.*, "World cancer report 2014," *World*, 2016.

[2] J. Y. Luh and C. R. Thomas Jr, "Radiation therapy for non-small cell lung cancer," *Lung Cancer Choices*, p. 45, 2012.

[3] N. J. Soper and D. B. Kaufman, *Northwestern handbook of surgical procedures*. Landes Bioscience, 2011.

[4] J. B. Brodsky and E. Cohen, "Video-assisted thoracoscopic surgery," *Current Opinion in Anesthesiology*, vol. 13, no. 1, pp. 41–45, 2000.

[5] V. Hornblower, E. Yu, A. Fenster, J. Battista, and R. Malthaner, "3d thoracoscopic ultrasound volume measurement validation in an ex vivo and in vivo porcine model of lung tumours," *Physics in medicine and biology*, vol. 52, no. 1, p. 91, 2006.

[6] J.-M. Baste, P. Rinieri, M. Sarsam, and C. Peillon, "Place de la chirurgie robotique dans les pathologies tumorales thoraciques," *Oncologie*, pp. 1–8, 2016.

[7] "Simple prostatectomy," http://emedicine.medscape.com.proxy1.lib.uwo.ca/article/445996-overview#a3, accessed: 2010-11-21.

[8] C. L. Bennett, S. J. Stryker, M. R. Ferreira, J. Adams, and R. W. Beart, "The learning curve for laparoscopic colorectal surgery: preliminary results from a prospective analysis of 1194 laparoscopic-assisted colectomies," *Archives of Surgery*, vol. 132, no. 1, pp. 41–44, 1997.

[9] M. Wilson, M. Coleman, and J. McGrath, "Developing basic hand-eye coordination skills for laparoscopic surgery using gaze training," *BJU international*, vol. 105, no. 10, pp. 1356–1358, 2010.

[10] P. Breedveld, H. Stassen, D. Meijer, and L. Stassen, "Theoretical background and conceptual solution for depth perception and eye-hand coordination problems in laparoscopic surgery," *Minimally invasive therapy & allied technologies*, vol. 8, no. 4, pp. 227–234, 1999.

[11] J. H. Palep *et al.*, "Robotic assisted minimally invasive surgery," *Journal of Minimal Access Surgery*, vol. 5, no. 1, p. 1, 2009.

[12] P. Zehnder and I. S. Gill, "Cost-effectiveness of open versus laparoscopic versus robotic-assisted laparoscopic cystectomy and urinary diversion," *Current opinion in urology*, vol. 21, no. 5, pp. 415–419, 2011.

[13] T. Wilson and R. Torrey, "Open versus robotic-assisted radical prostatectomy: which is better?" *Current opinion in urology*, vol. 21, no. 3, pp. 200–205, 2011.

[14] W.-b. Ji, Z.-m. Zhao, J.-h. Dong, H.-g. Wang, F. Lu, and H.-W. Lu, "One-stage robotic-assisted laparoscopic cholecystectomy and common bile duct exploration with primary closure in 5 patients," *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, vol. 21, no. 2, pp. 123–126, 2011.

[15] S. Masroor, C. Plambeck, and M. Dahnert, "Complex repair of a barlow's valve using the da vinci robotic surgical system," *Journal of Heart Valve Disease*, vol. 19, no. 5, p. 593, 2010.

[16] A. M. Okamura, "Haptic feedback in robot-assisted minimally invasive surgery," *Current opinion in urology*, vol. 19, no. 1, p. 102, 2009.

[17] D. McKay and G. Blake, "Technical notes and tips: Optimum incision length for port insertion in laparoscopic surgery," *Annals of the Royal College of Surgeons of England*, vol. 88, no. 1, p. 78, 2006.

[18] P. Risholm, S. Pieper, E. Samset, and W. M. Wells III, "Summarizing and visualizing uncertainty in non-rigid registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2010, pp. 554–561.

[19] X. Gao, N. V. Navkar, D. J. Shah, N. V. Tsekos, and Z. Deng, "Intraoperative registration of preoperative 4d cardiac anatomy with real-time mr images," in *IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), 2012*. IEEE, 2012, pp. 583–588.

[20] A. S. Naini, R. V. Patel, and A. Samani, "Ct image construction of the lung in a totally deflated mode," in *2009 IEEE International Symposium on Biomedical Imaging*. IEEE, 2009, pp. 578–581.

[21] A. Wittek, G. Joldes, M. Couton, S. K. Warfield, and K. Miller, "Patient-specific non-linear finite element modelling for predicting soft organ deformation in real-time; application to non-rigid neuroimage registration," *Progress in Biophysics and Molecular Biology*, vol. 103, no. 2, pp. 292–303, 2010.

[22] A. S. Naini, R. V. Patel, and A. Samani, "Ct-enhanced ultrasound image of a totally deflated lung for image-guided minimally invasive tumor ablative procedures," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2627–2630, 2010.

[23] T. Aoki, T. Ohashi, T. Matsumoto, and M. Sato, "The pipette aspiration applied to the local stiffness measurement of soft tissues," *Annals of Biomedical Engineering*, vol. 25, no. 3, pp. 581–587, 1997.

[24] P. Schiavone, T. Boudou, E. Promayon, P. Perrier, and Y. Payan, "A light sterilizable pipette device for the in vivo estimation of human soft tissues constitutive laws," in

30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.   IEEE, 2008, pp. 4298–4301.

[25] A. Nava, E. Mazza, M. Furrer, P. Villiger, and W. Reinhart, "In vivo mechanical characterization of human liver," *Medical image analysis*, vol. 12, no. 2, pp. 203–216, 2008.

[26] S. Umale, C. Deck, N. Bourdet, P. Dhumane, L. Soler, J. Marescaux, and R. Willinger, "Experimental mechanical characterization of abdominal organs: liver, kidney & spleen," *Journal of the mechanical behavior of biomedical materials*, vol. 17, pp. 22–33, 2013.

[27] B.-M. Ahn, J. Kim, L. Ian, K.-H. Rha, and H.-J. Kim, "Mechanical property characterization of prostate cancer using a minimally motorized indenter in an ex vivo indentation experiment," *Urology*, vol. 76, no. 4, pp. 1007–1011, 2010.

[28] Z. Cui, Z. Han, H. Pan, Y. Shao, and D. Zhu, "Design of a 3-axial force/torque sensor for arthroscopy force sensing," in *2011 IEEE International Conference on Mechatronics and Automation.*   IEEE, 2011, pp. 243–248.

[29] K. S. Shahzada, A. Yurkewich, R. Xu, and R. V. Patel, "Sensorization of a surgical robotic instrument for force sensing," in *SPIE BiOS.*   International Society for Optics and Photonics, 2016, pp. 97 020U–97 020U.

[30] U.-X. Tan, B. Yang, R. Gullapalli, and J. P. Desai, "Triaxial mri-compatible fiber-optic force sensor," *IEEE Transactions on Robotics*, vol. 27, no. 1, pp. 65–74, 2011.

[31] R. Sargeant, H. Liu, and K. Althoefer, "An mri compatible optical multi-axis force/torque sensors robotic surgery," in *The Hamlyn Symposium on Medical Robotics*, 2012, pp. 79–80.

[32] G. L. McCreery, A. L. Trejos, M. D. Naish, R. V. Patel, and R. A. Malthaner, "Feasibility of locating tumours in lung via kinaesthetic feedback," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 4, no. 1, pp. 58–68, 2008.

[33] T. Yamamoto, N. Abolhassani, S. Jung, A. M. Okamura, and T. N. Judkins, "Augmented reality and haptic interfaces for robot-assisted surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 8, no. 1, pp. 45–56, 2012.

[34] T. Yamamoto, B. Vagvolgyi, K. Balaji, L. L. Whitcomb, and A. M. Okamura, "Tissue property estimation and graphical display for teleoperated robot-assisted surgery," in *IEEE International Conference on Robotics and Automation, 2009. ICRA'09.*   IEEE, 2009, pp. 4239–4245.

[35] M. Jia, J. W. Zu, and A. Hariri, "A new tissue resonator indenter device and reliability study," *Sensors*, vol. 11, no. 1, pp. 1212–1228, 2011.

[36] A. Faragasso, A. Stilli, J. Bimbo, H. Wurdemann, and K. Althoefer, "Multi-axis stiffness sensing device for medical palpation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2015.*   IEEE, 2015, pp. 2711–2716.

[37] H. Liu, P. Puangmali, D. Zbyszewski, O. Elhage, P. Dasgupta, J. S. Dai, L. Seneviratne, and K. Althoefer, "An indentation depthforce sensing wheeled probe for abnormality identification during minimally invasive surgery," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 224, no. 6, pp. 751–763, 2010.

[38] S. McKinley, A. Garg, S. Sen, R. Kapadia, A. Murali, K. Nichols, S. Lim, S. Patil, P. Abbeel, A. M. Okamura, *et al.*, "A disposable haptic palpation probe for locating subcutaneous blood vessels in robot-assisted minimally invasive surgery," *IEEE CASE*, 2015.

[39] M. I. Tiwana, S. J. Redmond, and N. H. Lovell, "A review of tactile sensing technologies with applications in biomedical engineering," *Sensors and Actuators A: physical*, vol. 179, pp. 17–31, 2012.

[40] S. Schostek, C.-N. Ho, D. Kalanovic, and M. O. Schurr, "Artificial tactile sensing in minimally invasive surgery–a new technical approach," *Minimally invasive therapy & allied technologies*, vol. 15, no. 5, pp. 296–304, 2006.

[41] M. B. Cohn, M. Lam, and R. S. Fearing, "Tactile feedback for teleoperation," in *Applications in Optical Science and Engineering*. International Society for Optics and Photonics, 1993, pp. 240–254.

[42] A. L. Trejos, J. Jayender, M. T. Perri, M. D. Naish, R. V. Patel, and R. A. Malthaner, "Experimental evaluation of robot-assisted tactile sensing for minimally invasive surgery," in *2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE, 2008, pp. 971–976.

[43] M. Tavakoli, A. Aziminejad, R. Patel, and M. Moallem, "Methods and mechanisms for contact feedback in a robot-assisted minimally invasive environment," *Surgical Endoscopy and Other Interventional Techniques*, vol. 20, no. 10, pp. 1570–1579, 2006.

[44] A. Talasaz and R. V. Patel, "Remote palpation to localize tumors in robot-assisted minimally invasive approach," in *IEEE International Conference on Robotics and Automation (ICRA), 2012*. IEEE, 2012, pp. 3719–3724.

[45] A. Srinivasan Naidu, "Design of novel sensors and instruments for minimally invasive lung tumour localization via palpation," 2015.

[46] Å. B. Vallbo, R. S. Johansson, *et al.*, "Properties of cutaneous mechanoreceptors in the human hand related to touch sensation," *Hum Neurobiol*, vol. 3, no. 1, pp. 3–14, 1984.

[47] "da vinci surgical system user manual," accessed: 2016-08-10. [Online]. Available: http://www.vghtc.gov.tw/GipOpenWeb/wSite/public/Attachment/f1314187097823.pdf

[48] A. Reiter, P. K. Allen, and T. Zhao, "Appearance learning for 3d tracking of robotic surgical tools," *The International Journal of Robotics Research*, p. 0278364913507796, 2013.

[49] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[50] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.

[51] M. E. Arregui, J. Robert Jr, N. Katkhouda, J. B. McKernan, H. Reich, *et al.*, *Principles of laparoscopic surgery: basic and advanced techniques*. Springer Science & Business Media, 2012.

[52] S. Cotin, N. Stylopoulos, M. Ottensmeyer, P. Neumann, D. Rattner, and S. Dawson, "Metrics for laparoscopic skills trainers: the weakest link!" in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2002, pp. 35–43.

[53] V. Datta, A. Chang, S. Mackay, and A. Darzi, "The relationship between motion analysis and surgical technical assessments," *The American journal of surgery*, vol. 184, no. 1, pp. 70–73, 2002.

[54] M. Chmarra, C. Grimbergen, and J. Dankelman, "Systems for tracking minimally invasive surgical instruments," *Minimally Invasive Therapy & Allied Technologies*, vol. 16, no. 6, pp. 328–340, 2007.

[55] W. P. Williamson, S. Burnett, J. E. Smith, M. Fogelberg, and E. D. Allen, "Articulable socket joint assembly for an endoscopic instrument for surgical fastner track therefor," 1995, US Patent 5,405,344.

[56] A. D. Ramans, "Surgical tool with mechanical advantage," Mar. 27 2001, US Patent 6,206,903.

[57] W. Birkfellner, J. Hummel, E. Wilson, and K. Cleary, "Tracking devices," in *Image-Guided Interventions*. Springer, 2008, pp. 23–44.

[58] M. Ding and A. Fenster, "A real-time biopsy needle segmentation technique using hough transform," *Medical physics*, vol. 30, no. 8, pp. 2222–2233, 2003.

[59] E. Azimi, J. Doswell, and P. Kazanzides, "Augmented reality goggles with an integrated tracking system for navigation in neurosurgery," in *2012 IEEE Virtual Reality Workshops (VRW)*. IEEE, 2012, pp. 123–124.

[60] T. Zhao, W. Zhao, and W. C. Nowlin, "Configuration marker design and detection for instrument tracking," Apr. 23 2009, uS Patent App. 12/428,691.

[61] X. Zhang and S. Payandeh, "Application of visual tracking for robot-assisted laparoscopic surgery," *Journal of Robotic systems*, vol. 19, no. 7, pp. 315–328, 2002.

[62] D. R. Uecker and Y. Wang, "Image analysis for automated tracking in robot-assisted endoscopic surgery," in *Proc. Int. Conf Pattern Recognition*, 1994, pp. 88–92.

[63] D. Burschka, J. J. Corso, M. Dewan, W. Lau, M. Li, H. Lin, P. Marayong, N. Ramey, G. D. Hager, B. Hoffman, *et al.*, "Navigating inner space: 3-d assistance for minimally invasive surgery," *Robotics and Autonomous Systems*, vol. 52, no. 1, pp. 5–26, 2005.

[64] A. Reiter, P. K. Allen, and T. Zhao, "Marker-less articulated surgical tool detection," in *Computer assisted radiology and surgery*, 2012.

[65] ——, "Learning features on robotic surgical tools," in *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 2012, pp. 38–43.

[66] R. Dockter and T. M. Kowalewski, "A low-cost computer vision based approach for tracking surgical robotic tools," *Journal of Medical Devices*, vol. 7, no. 3, p. 030917, 2013.

[67] Z. Pezzementi, S. Voros, and G. D. Hager, "Articulated object tracking by rendering consistent appearance parts," in *IEEE International Conference on Robotics and Automation, 2009. ICRA'09.* IEEE, 2009, pp. 3940–3947.

[68] M. F. Khan, S. Dogan, A. Maataoui, S. Wesarg, J. Gurung, H. Ackermann, M. Schiemann, G. Wimmer-Greinecker, and T. J. Vogl, "Navigation-based needle puncture of a cadaver using a hybrid tracking navigational system," *Investigative radiology*, vol. 41, no. 10, pp. 713–720, 2006.

[69] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/kanade meets horn/schunck: Combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.

[70] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 2. IEEE, 2000, pp. 142–149.

[71] S. T. Birchfield and S. Rangarajan, "Spatiograms versus histograms for region-based tracking," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, pp. 1158–1163.

[72] Y. Zhong, A. K. Jain, and M.-P. Dubuisson-Jolly, "Object tracking using deformable templates," *IEEE transactions on pattern analysis and machine intelligence*, vol. 22, no. 5, pp. 544–549, 2000.

[73] A. E. Abdel-Hakim and A. A. Farag, "Csift: A sift descriptor with color invariant characteristics," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 2. IEEE, 2006, pp. 1978–1983.

[74] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*. Springer, 2006, pp. 404–417.

[75] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *2011 International conference on computer vision*. IEEE, 2011, pp. 2548–2555.

[76] A. Bosch, A. Zisserman, and X. Munoz, "Image classification using random forests and ferns," in *2007 IEEE 11th International Conference on Computer Vision*. IEEE, 2007, pp. 1–8.

[77] G. D. Hager and P. N. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 10, pp. 1025–1039, 1998.

[78] A. Reiter, P. K. Allen, and T. Zhao, "Marker-less articulated surgical tool detection," in *Computer assisted radiology and surgery*, 2012.

[79] C. Kuglin, "The phase correlation image alignment method," in *Proc. Int. Conf. on Cybernetics and Society, 1975*, 1975, pp. 163–165.

[80] D. Casasent and D. Psaltis, "Position, rotation, and scale invariant optical correlation," *Applied optics*, vol. 15, no. 7, pp. 1795–1799, 1976.

[81] Q.-s. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 12, pp. 1156–1168, 1994.

[82] Y. Keller, A. Averbuch, and M. Israeli, "Pseudopolar-based estimation of large translations, rotations, and scalings in images," *IEEE Transactions on Image Processing*, vol. 14, no. 1, pp. 12–22, 2005.

[83] S. S. Gleason, M. A. Hunt, and W. B. Jatko, "Subpixel measurement of image features based on paraboloid surface fit," in *Fibers' 91, Boston, MA*. International Society for Optics and Photonics, 1991, pp. 135–144.

[84] M. Kytö, M. Nuutinen, and P. Oittinen, "Method for measuring stereo camera depth accuracy based on stereoscopic vision," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2011, pp. 78 640I–78 640I.

[85] C. Lally, A. Reid, and P. Prendergast, "Elastic behavior of porcine coronary artery tissue under uniaxial and equibiaxial tension," *Annals of biomedical engineering*, vol. 32, no. 10, pp. 1355–1364, 2004.

[86] P. Song, P. Kraus, V. Kumar, and P. Dupont, "Analysis of rigid-body dynamic models for simulation of systems with frictional contacts," *Journal of Applied Mechanics*, vol. 68, no. 1, pp. 118–128, 2001.

[87] D. W. Marhefka and D. E. Orin, "A compliant contact model with nonlinear damping for simulation of robotic systems," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 29, no. 6, pp. 566–572, 1999.

[88] H. H. Ip and D. Shen, "An affine-invariant active contour model (ai-snake) for model-based segmentation," *Image and Vision Computing*, vol. 16, no. 2, pp. 135–146, 1998.

[89] I. Nagahiro, A. Andou, M. Aoe, Y. Sano, H. Date, and N. Shimizu, "Pulmonary function, postoperative pain, and serum cytokine level after lobectomy: a comparison of vats and conventional procedure," *The Annals of thoracic surgery*, vol. 72, no. 2, pp. 362–365, 2001.

[90] A. S. Naini, T.-Y. Lee, R. V. Patel, and A. Samani, "Estimation of lung's air volume and its variations throughout respiratory ct image sequences," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 152–158, 2011.

[91] S. F. Atashzar, I. Khalaji, M. Shahbazi, A. Talasaz, R. V. Patel, and M. D. Naish, "Robot-assisted lung motion compensation during needle insertion," in *IEEE International Conference on Robotics and Automation (ICRA), 2013*.   IEEE, 2013, pp. 1682–1687.

[92] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[93] P. Dollar, Z. Tu, and S. Belongie, "Supervised learning of edges and object boundaries," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 2.   IEEE, 2006, pp. 1964–1971.

[94] J. Shotton, A. Blake, and R. Cipolla, "Contour-based learning for object detection," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 1.   IEEE, 2005, pp. 503–510.

[95] D. H. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[96] D. Wu and W. Zhang, "Multiscale contour matching with basis functions," *Optical Engineering*, vol. 47, no. 3, pp. 037 009–037 009, 2008.

[97] S. K. Balci, P. Golland, and W. Wells, "Non-rigid groupwise registration using b-spline deformation model," *Open source and open data for MICCAI*, pp. 105–121, 2007.

[98] I. Matthews, T. Ishikawa, S. Baker, *et al.*, "The template update problem," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 6, pp. 810–815, 2004.

[99] D. Schreiber, "Robust template tracking with drift correction," *Pattern recognition letters*, vol. 28, no. 12, pp. 1483–1491, 2007.

[100] N. Chambers and D. Jurafsky, "Template-based information extraction without the templates," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*.   Association for Computational Linguistics, 2011, pp. 976–986.

[101] C. Cagniart, E. Boyer, and S. Ilic, "Probabilistic deformable surface tracking from multiple videos," in *European conference on computer vision*.   Springer, 2010, pp. 326–339.

[102] Y. Nakamura, K. Kishi, and H. Kawakami, "Heartbeat synchronization for robotic cardiac surgery," in *IEEE International Conference on Robotics and Automation, 2001. Proceedings 2001 ICRA.*, vol. 2.   IEEE, 2001, pp. 2014–2019.

[103] R. Ginhoux, J. Gangloff, M. de Mathelin, L. Soler, M. M. A. Sanchez, and J. Marescaux, "Active filtering of physiological motion in robotized surgery using predictive control," *IEEE Transactions on Robotics*, vol. 21, no. 1, pp. 67–79, 2005.

[104] T. Ortmaier, M. Groger, D. H. Boehm, V. Falk, and G. Hirzinger, "Motion estimation in beating heart surgery," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 10, pp. 1729–1740, 2005.

[105] R. Richa, A. P. Bó, and P. Poignet, "Motion prediction for tracking the beating heart," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.   IEEE, 2008, pp. 3261–3264.

[106] C. Vaz, X. Kong, and N. Thakor, "An adaptive estimation of periodic signals using a fourier linear combiner," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 1–10, 1994.

[107] K. C. Veluvolu, U.-X. Tan, W. T. Latt, C. Shee, and W. T. Ang, "Bandlimited multiple fourier linear combiner for real-time tremor compensation," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.   IEEE, 2007, pp. 2847–2850.

[108] M. P. Tarvainen, J. K. Hiltunen, P. O. Ranta-aho, and P. A. Karjalainen, "Estimation of nonstationary eeg with kalman smoother approach: an application to event-related synchronization (ers)," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 3, pp. 516–524, 2004.

[109] P. Wellman, R. D. Howe, E. Dalton, and K. A. Kern, "Breast tissue stiffness in compression is correlated to histological diagnosis," *Harvard BioRobotics Laboratory Technical Report*, pp. 1–15, 1999.

[110] M. Tavakoli, A. Aziminejad, R. Patel, and M. Moallem, "Tool/tissue interaction feedback modalities in robot-assisted lump localization," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*. IEEE, 2006, pp. 3854–3857.

Curriculum Vitae

Name:

Post-Secondary Education and Degrees:
The University of Western Ontario
London, Ontario, Canada
2013 - 2017 M.E.Sc (Electrical and Computer Engineering - Robotics and Control)
Indian Institute of Technology Jodhpur
Jodhpur, Rajasthan, India
2009 - 2013 B.Tech (Computer Science and Engineering)

Related Work Experience:
Teaching Assistant
University of Western Ontario
2013 - 2015
Research Assistant
University of Western Ontario
2013 - 2017