IMAGING OF SURGICAL TOOLS AS A NEW PARADIGM FOR SURGEON COMPUTER -INTERFACE IN MINIMALLY INVASIVE SURGERY

by

Mandana Salajegheh
M.Sc, Azad University, 2001
B.Sc, Sharif University of Technology, 1996

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

MASTER OF APPLIED SCIENCE

In the
School of Engineering Science

SIMON FRASER UNIVERSITY

Fall 2008

APPROVAL

Name: **Mandana Salajegheh**

Degree: **Master of Applied Science**

Title of Thesis: **Imaging of surgical tools as a new paradigm for Surgeon Computer-Interface in Minimally Invasive Surgery**

Examining Committee:

Chair: **Dr. Marinko Sarunic**
Assistant Professor of Engineering Science

_____

Dr. Shahram Payandeh
Senior Supervisor
Professor of Engineering Science

_____

Dr. Parvaneh Saeedi
Supervisor
Assistant Professor of Engineering Science

_____

Dr. Belgacem Ben Youssef
Supervisor
Assistant Professor of School of Interactive Art and Technology

_____

Dr. Karim Qayumi
External Examiner
Professor of Department of Surgery, University of British Columbia

Date Defended/Approved: December 5th 2008

ABSTRACT

Today's Minimally Invasive Surgery (MIS) is performed by inserting an endoscope along with long stem surgical tools into the incision holes in the patient's body to capture real-time images. Although MIS has many advantages compared to an open surgery, it also introduces a number of constraints for surgeons. In this thesis, we propose a novel paradigm for surgeon computer interface which allows surgeons to use tracked surgical tools like a 3D input device to obtain more information regarding patients' anatomy during the MIS. We offer using tracked surgical tools to extract the boundaries of an organ from cluttered specular surfaces in MIS. Through real-time image analysis of the state and gesture of surgical tools, we also propose three augmented systems to access patient's pre-operative scans during the MIS.

ACKNOWLEDGEMENTS

Special thanks to my senior supervisor, Dr. Shahram Payandeh, for being supportive and encouraging throughout this thesis work. I truly appreciate all the time and efforts he has put into helping me with this project. This work would not be achievable without his great ideas, guidance and extensive knowledge in all aspects of my research.

Thanks to my supervisor Dr. Parvaneh Saeedi for providing valuable information, for her input and continuous and generous support throughout this thesis work.

Thanks to my supervisor Dr. Ben Youssef for his guidance and providing feedback on my research.

Thanks to Dr. Karim Qayumi for accepting to be my defence examiner and thanks to Dr. Marinko Sarunic for chairing my defence.

Thanks to Dr. Peter Doris of Surrey Memorial Hospital for providing feedback on my research and also for arranging a visit to a real surgical operation. Thanks to Siavash Rezaie for helping me with developing the source codes for this project.

Finally, I would like to thank my friends and family, especially my parents for their unconditional love and support throughout my entire life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Active contour	Computer generated curves that move within images to find object boundaries.
Blob	Connected regions of pixels within an image.
Endoscope	An instrument used to look inside the body for surgical or diagnostic purposes.
Gesture recognition	Recognizing human gestures with mathematical algorithms.
Human computer Interface	The interaction between people and computers.
Laparoscopic surgery	A modern surgical technique in which operations inside the abdominal area is performed through small incisions.
Minimally Invasive Surgery	A surgical procedure which is less invasive than an open surgery and it usually involves using an endoscope.
Robot	Mechanical systems under automatic control.
Robotic surgery	Using robots in surgical procedures.
Surgical tool/instruments	Tools used to perform surgical tasks such as graspers, scissors, forceps, retractors, clips and staplers.
Segmentation	Partitioning an image into meaningful regions.
Surgeon Computer Interface	The interaction between surgeons and computers.
Telesurgery	Surgical procedures performed over a distance.

LIST OF ACRONYMS

3D	3 Dimensional
BCI	Brain Computer Interface
CCD	Charged Coupled Device
EMG	Electromyogram
HCI	Human Computer Interface
MIS	Minimally Invasive Surgery
PC	Personal Computer
SCI	Surgeon Computer Interface

Chapter 1

Introduction

This chapter starts with a brief introduction of Minimally Invasive Surgery (MIS) and its limitations. Section 1.2 describes the problem statement and the motivation for this research. We discuss some of the existing limitations of MIS and also review some of the current technologies in the fields of MIS, robotics and telesurgery as well as interactive systems for MIS and Human Computer Interface (HCI) technology. In section 1.3, we list the contribution of this thesis. Section 1.4 explains the layout of the rest of this document.

1.1 Minimally invasive surgery and its limitations

About 20 years ago, surgeons began practicing a new approach to performing surgery, an approach that came to be known as Minimally Invasive Surgery, or MIS. Traditionally, surgeries had been performed in the open manner, in which large incisions were required for the surgeon to clearly observe and manipulate the surgical site. These incisions inevitably created significant patient trauma, substantial pain and suffering, extended recovery time, prolonged pain management and elevated costs. Today, MIS is performed by inserting an endoscope or laparoscope camera along with two or more long stems surgical tools into the incision holes in the patient's body to capture the real-time images inside the abdomen. In addition to be more cost effective, compared with conventional open surgery, MIS offers a lot of benefits to patients. These advantages include reduction of trauma to the patients' body, less blood loss, less risk of infection, shorter hospitalization time and

less scarring which speed up the patient's recovery. Figure 1.1 shows an operating room for a typical laparoscopic procedure which is a type of MIS.

Figure 1.1: Minimally invasive laparoscopic surgery for gallbladder removal

Although MIS has many advantages for patients, it also introduces a number of constraints for the surgeon. The reverse hand movement and limited force sensing of the remote surgical site, in conjunction with indirect vision and the straining body posture of surgeons, decrease their dexterity dramatically compared to open surgery [1]. These limitations often make the MIS a very exhausting and complicated procedure. As a result, this type of surgery turned out to be suitable for a narrow range of surgical procedures and only a small group of highly skilled surgeons routinely attempt complex procedures using a minimally invasive approach.

1.2 Problem statement and motivation

Visual related problems have created major challenges in design of MIS systems. In this type of surgery, surgeons operate the surgical procedure using a standard 2D monitor instead of looking directly at their hands and the surgical site. The 2D image of the monitor has flattened the natural depth in the field of view. Lack of 3D visualization and stereoscopic view, limited field of view and resolution along with limited contrast and color fidelity are common problems with typical MIS procedure.

Due to the existence of such visual limitations, it would be desirable for the surgeons to obtain more information regarding patient's anatomy during the MIS procedures in order to visualize internal anatomical structures and navigate their way through out of sight and hidden parts of the body. It is very common to refer to patient's pre-operative scans during the surgical procedure. For example surgeons might want to see the patient's X-ray, MRI scan, or other medical records or even want to register and superimpose pre-operative medical images on top of intra-operative images of the organ of interest. Such access to medical information would be more desirable if the surgeon could invoke data mining and extraction by simple commands which require minimum effort and without being distracted during the surgery.

In addition to MIS, the proposed interface could have potential applications in the fields of robotics and telesurgery. Figure 1.2 is artificially created to show the applications of one of our proposed systems in a real operating room.

Unlike some of other techniques, this interface allows surgeons to obtain patient's medical records without using external hardware which could restrict surgeon's

movements. Since the instruments used in this interface are typical surgical tools which surgeons are already familiar with, there is no need for intense training sessions.

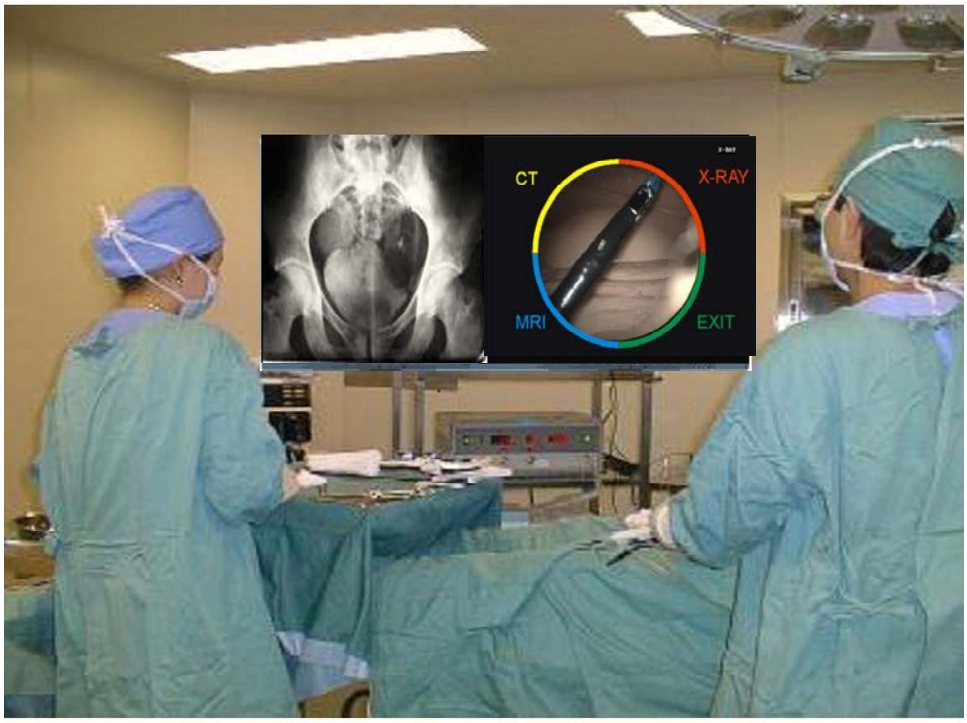


Figure 1.2: The ring menu enables surgeons to access preoperative scans during the MIS by touching the particular region of the menu with surgical tools

1.2.1 Robotics and telesurgery

The process of minimally invasive surgery has been augmented by specialized tools for decades. Researchers around the world have tried to use robotic and image processing systems in order to interactively assist the medical team, both in planning and execution of surgical intervention. The objective of these new techniques is to enhance the quality of surgical procedures while minimizing their side effects.

In recent years, many specialized devices and computational/visualization environments have been developed to assist surgeons to overcome some of the current limitations with the MIS. In the late 1990s, another evolutionary stage in the development of surgical techniques was achieved with the application of robotics in surgical technology. Telesurgery or remote surgery is one of the possible applications of robotics. It is the ability for a doctor to perform surgery on a patient even though they are not physically in the same location. Remote surgery combines elements of robotics and communication technology such as high-speed data connections and elements of management information systems.

Intuitive Surgical Inc. [2] has become one of the leaders in development and commercialization of robotic technology for MIS. The *da Vinci* and the *da Vinci S HD* surgical system integrate 3D high definition endoscopy and state of the art robotic technology to virtually extend the surgeon's eyes and hands into the surgical field. As shown in figure 1.3, the *da Vinci* Surgical System consists of an ergonomically designed surgeon's console, a patient-side cart with four interactive robotic arms, the high performance *InSite*® Vision System and proprietary *EndoWrist*® Instruments with 7 degrees of freedom. Surgeon's hand movements are scaled, filtered and seamlessly translated into precise movements of the *EndoWrist* Instruments.

IRCAD [3] , in partnership with the CNRS (French National Scientific Research Centre) and the Louis Pasteur University in Strasbourg, is currently working on the development of a new generation robot.

Figure 1.3: The da Vinci robot system for MIS

The objective is to automate the surgical gesture so that it becomes more secure: thanks to the combination of augmented reality and robotics, it is possible to carry out an action on the patient that has previously been programmed on the digital copy of that patient. The ODYSSEUS project [3], shown in figure 1.4, is another venture coordinated by IRCAD in collaboration with four other European partners. They have developed a software for 3D patient reconstruction and simulation of surgical intervention in digestive cancerology. This project combines the knowledge and expertise in the fields of medicine, computer sciences, robotics and surgical industry. ODYSSEUS exploits various technologies such as telecommunications, virtual reality and surgical robotics. From a CT-scan or MRI, softwares will provide a 3D copy of the patient on which medical experts will prepare their surgery.

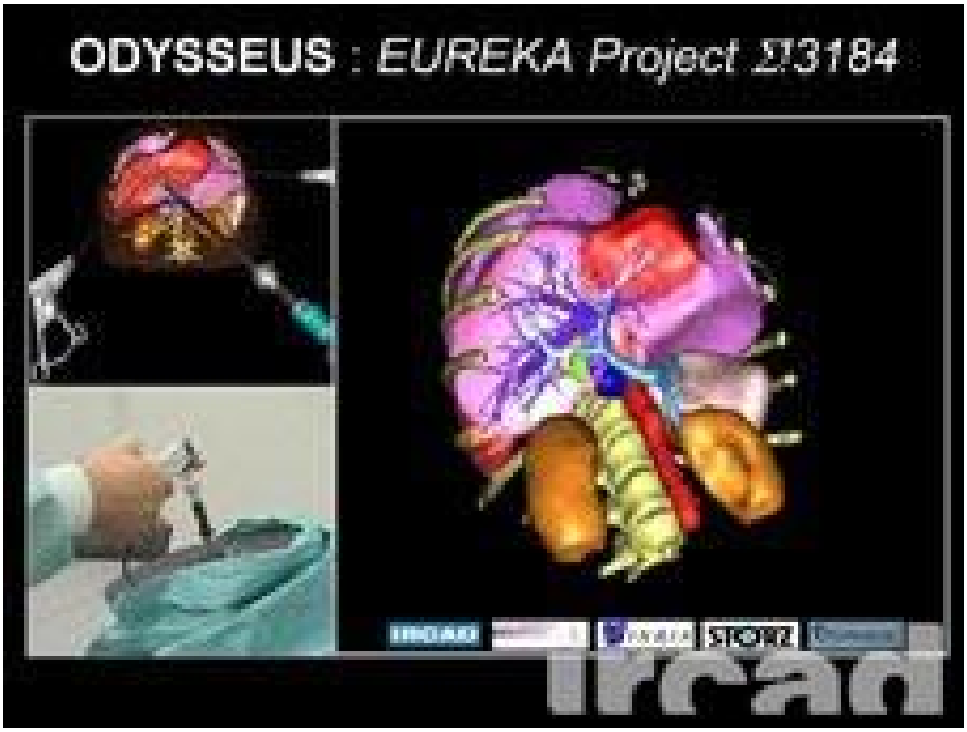


Figure 1.4: The ODYSSEUS project

1.2.2 Review of interactive HCI systems

Some of the difficulties with the current imaging in the MIS include the lack of stereoscopic imaging for depth perception, geometric distortion, limited resolution and contrast. Also, the limited dexterity of surgical instruments and the lack of tactile and kinesthetic feedbacks are examples of some of the additional challenges which surgeons face. Moreover, endoscopic techniques are more difficult to learn comparing to conventional surgeries.

In general every human activity requires a suitable environment to be performed properly. However this is not the case with most current MIS techniques. A study by Faraz & Payandeh [1], suggested a possible method of enhancing the dexterity by an external mechanism which provides support for

surgical tools outside the body. They investigated different designs of external supports to create the environment by considering surgical needs and requirements. They have developed a multi-arm robotic system which provides support as well as capability to lock several movable tools in various positions. They indicated that the design should also satisfy the following requirements: avoid obstruction of the workspace, avoid interference with other surgical tools and comply with the kinematics constraints of the incision points.

Generally most of the above problems with MIS are related to the challenges associated with human interface where their potential solutions lie in the design of proper interface techniques [4][5]. The study by Peters [6] indicated that another big challenge in the design of image guided surgery systems is that the design should be simple to operate by the users. It should also minimize additional technical support and eliminate keyboard and inconvenient switching between mechanisms in operating rooms. In fact the performance of a proper and intuitive Surgeon Computer Interface (SCI) between the surgeon and the surgical site would be an essential requirement for the quality of the surgery.

Commercially available tracking systems such as OPTOTRAK and POLARIS optical tracking systems made by Northern Digital [7], which are able to provide position and orientation information with millimetre accuracy have been generally used in MIS researches [8]. These tracking systems use two or three cameras to track the position of passive or active markers which are infrared light emitting diodes (LEDs) attached to the objects to be tracked. Some of the more recent optical systems use passive tracking technology where reflecting markers attached to the probe are recognized by a multi-

receptor optical imaging system. The new smart wireless markers allow for fast and easy experimental setup with more easy wire management, 80 percent reduction in the number of wires and more freedom of movement.

Virtual endoscopic navigation, combined with a standard display of a set of orthogonal planes has been utilized for enhancing clinical diagnosis and for planning surgery and treatment [9]. Mapping anatomical landmarks derived from pre-operative MRI images to live video has been done with a tracked endoscope [10] . Image overlay systems for medical data visualization have been proposed for intra-operative tool localization and guidance as well as surgical education and training [11].

Hand movement tracking, facial feature tracking, eye gaze tracking, electromyogram signal, or brain signal based HCI systems have been developed to help paralyzed patients or people with significant physical disabilities gain access to computers.

Wearable HCI such as a data glove also provides a natural and convenient means for a human being to communicate with a computer. Chou *et al.* [12] developed a data glove with tactile feedback and an exoskeleton arm type haptic interface. The data glove has eleven degrees of freedom, with vibrators that are mounted on the tip of five glove fingers. The arm type haptic interface has five degrees of freedom. It can be worn on an operator's arm, and the operator can obtain real time force feedback. The combination of data glove and an arm type haptic device provide both external forces of contact and internal forces of grasping, which are useful for completing complex works such as teleoperation and interacting with virtual environments.

Over the past two decades, a variety of studies have evaluated the possibility that brain signals recorded from the scalp or from within the brain could provide new augmentative technology that does not require muscle control. Brain-Computer-Interface (BCI) research and development is a complex interdisciplinary endeavour that depends on careful evaluation and comparison of various different brain signals, signal processing methods, and output devices. Schalk *et al.* [13] have developed a general-purpose BCI research and development a platform called BCI2000. BCI systems measure specific features of brain activity and translate them into device control signals. Progress and realization of practical BCI applications depend on systematic evaluations and comparisons of different brain signals, recording methods, processing algorithms, output forms, and operating protocols.

Ko *et al.* [14] proposed a robust, fast and low-cost scheme for locating the eyes, lip-corners, and nostrils for the Eye-Head Controlled Human Computer Interface on a facial image with non-constrained background. The proposed algorithm computes the similarity between all pairs of objects after the image is thresholded. The two objects with the greatest similarity are then selected as eyes.

The study by Nishikawa *et al.* [15] showed that surgeon can control the movement of endoscope by performing an appropriate facial gesture which could be recognized by the system. It has also been shown that it is more convenient for the surgeon when there is no need for body contact devices such as head-mounted sensing devices, hand or foot switches or voice commands.

Eye gaze studies have an increasingly important role in examining usability of human computer interfaces [16]. Lyons *et al.* [17] proposed a hybrid hands-off HCI that

uses infrared video eye gaze tracking and electromyogram (EMG) signals. This system combines the advantages of the sub-systems, providing quick cursor displacement in long excursions and steady, accurate movement in small position adjustments. The hybrid system also provides a reliable clicking mechanism. Test results also show that the hybrid system is, on average, faster than the EMG-only system by a factor of two or more.

In a study by Atkins *et al*. [18], eye gaze tracking data were used in designing radiology workstations. They showed that eye gaze data is useful to analyze computer interaction techniques, visual search strategies, and certain kinds of errors in their artificial look like radiology task. This may prove helpful in designing radiology work stations. The great advantage of an eye gaze tracking system is that it could eliminate the head mounted optics.

In another study by Atkins *et al.* [19], they compared three different interactive navigation techniques for radiology images. These techniques include wheel mouse just the wheel, wheel mouse using click and drag, and Shuttle Xpress jog wheel. The result of the conducted user study among students and radiologists showed that unlike most students, the radiologists preferred using the mouse wheel over the jog wheel. This preference may result from their extensive experience with the mouse.

1.2.3 Thesis objectives

In this research, we explore image analysis of surgical tools as a new paradigm for the MIS and we also investigate the use of surgical tools as an interactive input device for surgeons. Our main objective is to design a natural and intuitive interface that allows surgeons to obtain more information regarding patient's anatomy during the MIS. Within

the concept of the new proposed interface, we propose novel augmented systems which permit surgeons to employ surgical tool as an input device and access to patient's pre-operative scans during the MIS.

We consider the usability of the new interface by minimizing additional technical support and eliminating keyboard and inconvenient switching between different tasks. These considerations could eventually reduce the number of assistants and use of voice commands in the operating rooms. We assume that using surgical tools like input device could be a natural and convenient way of human computer interface for the MIS. Taking into account the limitations of the experimental setup and the differences between the experimental setup and the actual surgery, we conduct preliminary user studies to test our hypothesis and show the practicality of our proposed systems.

1.3 Contribution of the thesis

- **Track surgical tools using blob analysis of coloured images**

Unlike previous researches which have used blob analysis of grey scale images to track the surgical tools, in this research we process coloured images of the endoscope. Having different colours for different surgical tools eliminates the possibility of mistaking various blobs from each other while tracking them. For example, instead of calculating complex shape attributes such as feret elongation, roughness and compactness to identify graspers from scissors we can only refer to their colours to identify them. This reduces the possibility of misidentification of surgical tools from each other especially when two tools, i.e. open graspers and open scissors look very similar in some orientation angles and it is hard to distinguish between them based on their shape attributes.

- **Propose a practical method for segmentation of live endoscope images**

Automatic segmentation of specular surfaces such as laparoscope images is a very challenging task and it usually leads to false contouring. Employing tracked surgical tools like a 3D mouse to select the seed points along the boundaries of the organ is another contribution of this thesis. We show that this method is practical and eliminates the false contouring effect. We introduce the application of a small wireless switch which is attached to the surgical tools handle and could be easily pressed with user's finger as an interactive device. This switch could replace the use of computer mouse and result in more accuracy and consistency in the proposed segmentation task. The wireless switch could allow surgeons to obtain more information regarding patient's anatomy which is traditionally provided by an assistant.

- **Introducing the cube overlay system to manipulate extracted contours and pre-operative scans in MIS**

We introduce the cube overlay system as a potential solution which eventually allows surgeons to modify the extracted contour. Assuming that the extracted contour is surrounded by the virtual cube, we can grab, resize, translate and rotate the virtual cube and fit the contour more closely to the actual boundary of the organ by using tracked surgical tools. The cube overlay could also be utilized as a calibration block to measure the size of an internal organ. We propose extending the application of this system to superimpose 3D pre operative scans on top of live endoscope images.

- **Introducing three interactive systems to retrieve pre operative medical scans during the MIS procedure**

It is often desirable for surgeons to obtain some information regarding patient's medical records such as pre-operative scans during the surgery. We propose three intuitive and interactive systems which allow surgeons to retrieve information regarding patient's medical records during the MIS. The first system lets the surgeon switch to augmented mode of the system by forming a particular gesture with surgical tools and obtain this information. The second and third systems enable access to patient's medical records via displaying a narrow ring around the field of view of the endoscope camera. However, in the second system, scans become activated automatically as the user enters the tip of the surgical tool in the corresponding section of the menu whereas in the third system scan activation also requires pressing the wireless switch on the surgical tool handle. Not only these three proposed systems allow surgeons to have access to pre-operative scans without taking their hands off from the surgical tools, but also could reduce the time, voice commands and the number of assistants in an operation room.

1.4 Thesis layout

The remainder of this document is structured as follows. In chapter 2, we explain the method that we have used to track surgical tools in the field of view of the endoscope camera. The tracking system will allow us to obtain 3D coordinates of the surgical tools which will later be used in designing interactive systems for the MIS. In chapter 3, we will propose a new segmentation method for live endoscope images. The extracted contours could be used for further registration of pre-operative scans with live endoscope

images. We conduct some user studies and recommend two new systems to improve the proposed segmentation method. We suggest replacing the computer mouse with a small wireless switch attached to the surgical tool handle. Then we introduce the cube overlay system to manipulate the extracted contours. In chapter 4, through analysis of gesture and position of the surgical tools, we propose three new systems to access patient's medical records during the MIS. We conduct user studies and compare the three proposed systems and discuss the potential advantages and disadvantages of each of them. In chapter 5, we discuss the proposed interface and compare it to some other HCI techniques. We also make suggestions for future work in this area and offer some ideas to improve and extend the applications of the proposed interface.

Chapter 2

Automatic tracking of surgical tools using blob analysis

Automatic tracking of surgical tools is a key step in the design of our new paradigm for surgeon computer interface for minimally invasive surgery. We start this chapter with explaining the experimental setup as well as the hardware and software used for this research. In order to track surgical tools, we have to estimate their 3D coordinates in camera coordinate system. Camera calibration is a major step in tracking surgical tools which is described in section 2.2. In order to calculate the 3D coordinates of the surgical tools we have used image processing method of blob analysis of coloured images which is explained in section 2.3 of this chapter. Section 2.4 describes the developed tracking program and finally the result and accuracy of the tracking program in estimating 3D coordinates is shown in section 2.5.

2.1 Components of the experimental setup

The experimental setup used for this research consists of a plastic stomach model made by *Limb & Things* Company and some typical long stem surgical tools such as gaspers and scissors as shown in figure 2.1. The rubber organs and rubber tissue inside the plastic stomach mimic a real laparoscopic surgery. Images are captured by the *Karl Storz Supercam 9050B* CCD coloured camera. The CCD sensors inside the camera are electronic systems which transform the real image (photons) into electronic images that may be read on a screen. The *Surpercam* is a small, light weight, easy to operate and high resolution camera for video monitoring and recording of laparoscopic procedures.

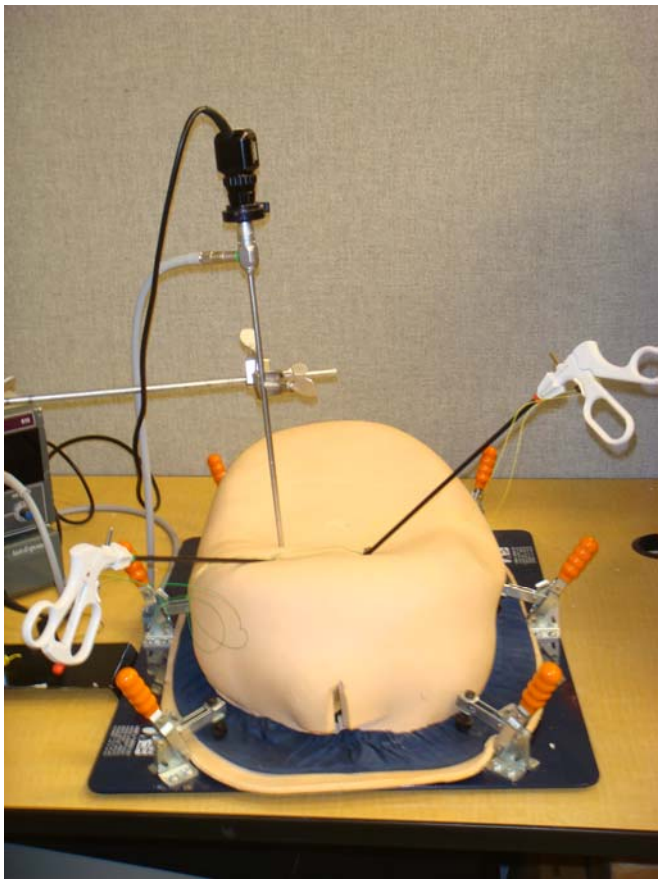


Figure 2.1: Plastic stomach with long stem surgical tools

Figure 2.2 shows the *Karl Storz 615* xenon light source, which provides illumination through a fibre optic cable. The xenon lamp colour temperature approximates bright sunlight and is considered unmatched for visual and photographic colour rendition. The brightness can be adjusted by the switch on the control panel. Since the light source is co-located with the endoscope camera, the illumination is not even throughout the entire field of view of the endoscope. In addition to this, changes in illumination level might affect the image processing results such as blob detection process. Therefore in order to be consistent, the illumination was set to 54 during the entire experiment.

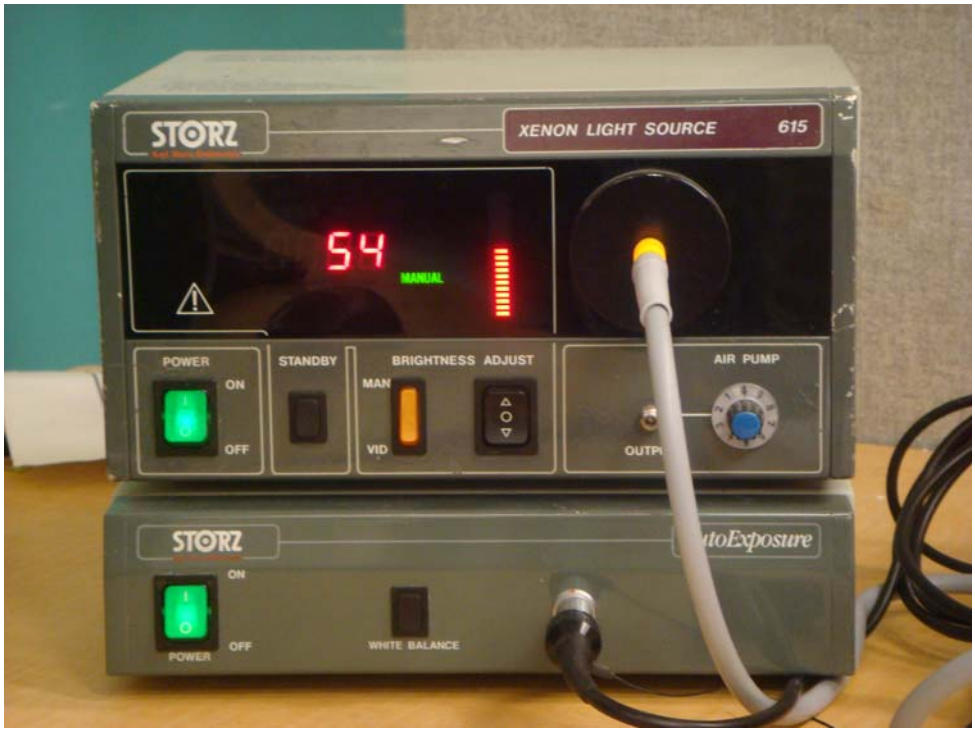


Figure 2.2: The xenon light source

The captured images are transferred to the camera unit and the camera unit then feeds the images to the *Matrox Coronosplus* frame grabber in the image processing PC. *Matrox CronosPlus* is a standard monochrome and composite color analog frame grabber for extremely cost-sensitive imaging applications. This board is available in a PCI form factor and can acquire different types of standard video formats using its video decoder. The video decoder can accept composite (CVBS) and component S-video (Y/C) in NTSC/PAL formats, and monochrome video in RS-170/CCIR. The *Matrox CronosPlus* board allows the transfer of live video to Host memory or off board display memory. It also features a 32-bit/33 MHz PCI bus master to reduce CPU usage. The board can also generate interrupts for the start and end of a field, frame, and sequence capture.

Programs for this research are written in C++ using *Matrox Imaging Library (MIL) version 8.0* which is a hardware-independent, modular 32-bit imaging library developed for image capture, image processing and analysis. MIL provides an extensive list of functions used to capture, process, analyze, transfer, display, and archive images. The processing and analysis operations include: spatial filtering, morphological, measurement, blob analysis, optical character recognition (OCR), pattern recognition (normalized grey scale correlation and Geometric Model Finder), matrix/bar code reading, and calibration operations.

2.2 Camera calibration method

Defining pixel to world mapping is known as calibration. Once the camera has been calibrated, we can transform image pixel coordinates to their real world equivalents. Camera calibration is a critical first step in many applications such as tracking, camera on robot configuration and robot vehicle guidance. In general, calibration parameters are classified into two categories: intrinsic and extrinsic parameters. Intrinsic parameters include the effective focal length of the camera, the real image center and the scale factors whereas extrinsic parameters incorporate the rigid body transformation from the world coordinate system to camera coordinate system.

2.2.1 Calculating intrinsic camera parameters

The method that we have used here to find intrinsic camera parameters is similar to the method proposed by Zhang & Payandeh [20]. The first step in camera calibration is to find the real image center. A fibre optic cable transmits light to the side of the

endoscope. The light source is evenly distributed through the rod lens. If we point the endoscope perpendicular to a white background and record the image, we can easily find the real center of the image, which is the same as the center of the white circle. The second step is to find the focal length (*f*) of the camera which is the distance between the image plane (*o*) and optical center (o_c). In order to calculate the focal length of the camera, we assume to have a perfect pinhole camera model. Figure 2.3 shows the basic geometry of the camera.

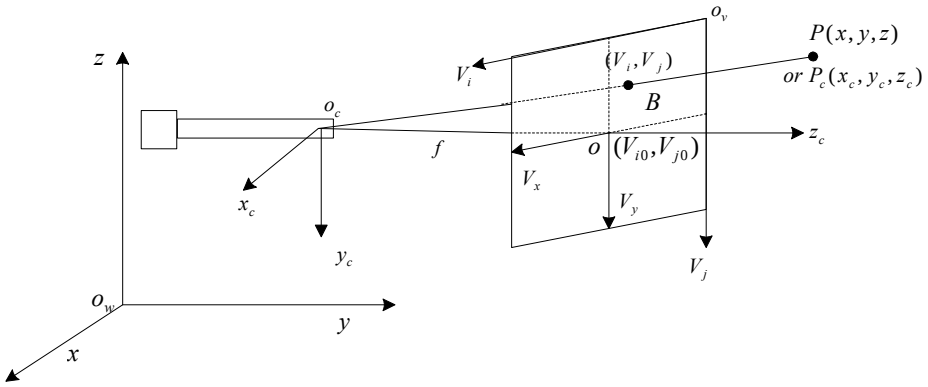


Figure 2.3: Camera geometry with perspective projection

The calibration grid consists of a series of black dots with known center to center distances. We measure the distances between the centers of two adjacent circles on the screen at various depths from 20-150 mm and take their mean value. Based on the geometry of the triangles, the transformation from 3D camera coordinates (x_c, y_c , z_c) to the 2D image coordinate (V_{io}, V_{jo}) with pinhole camera model, we have:

$$\frac{f}{z_c} = \frac{(V_i - V_{io})S_x}{x_c} = \frac{(V_{j-}V_{jo})S_y}{y_c} \quad (1)$$

Where $f = oo_c$ is the focal length of the camera or the distance between the center of the camera and the center of the image buffer, ($V_{io,}$ V_{jo}) is the image center in pixel units, ($V_{i,}$, V_j) is the location of the 2D projection of the point (x_c, y_c , z_c) on the image plane and both S_x and S_y are the scale factors that map the points in camera coordinate system to the image coordinate system in horizontal and vertical dimensions respectively. In the previous study [20] it has been shown that the S_x and S_y values are the same, therefore, when V_i, V_j, X_c and Z_c are known parameters, the effective focal length can be calculated in pixels units from equation (1). Also, when f, V_i, V_j and x_c, y_c are known parameters, we can get the depth information Z_c of an object in the field of view.

2.2.2 Using a grid for pixel to world mapping

An image associated with a calibration object is known as calibrated image. A calibration object could be a grid which consists of a series of black dots on a white background. When we associate a calibration object to an image, the image receives a copy of the calibration object's current relative coordinate system and current relative camera position as well as a reference to the calibration object for all other settings. As shown in figure 2.4, the circle in the top left corner of the grid image is associated to the origin, (0, 0), of the real world coordinate system, the first column of circles is aligned with its Y-axis, and the first row of circles is aligned with its X-axis. This is because, in

general, we know where that first circle is in the real-world, so we need results with respect to that position.

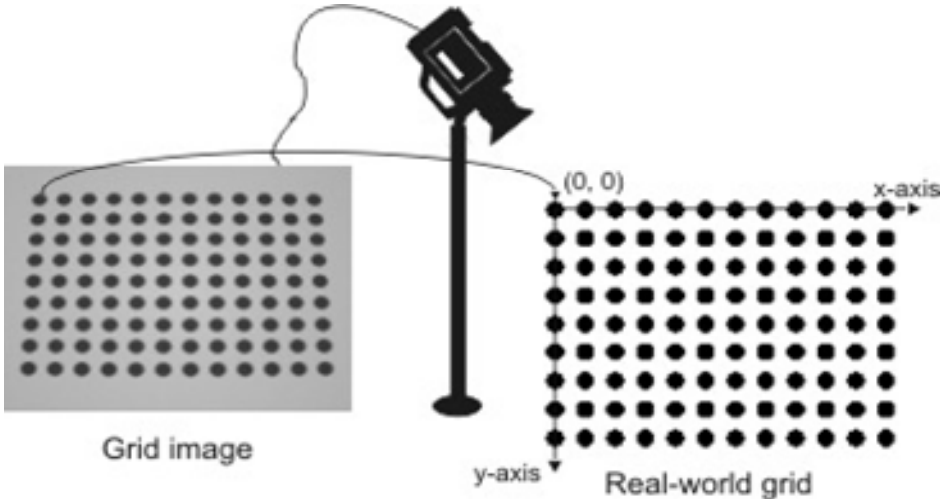


Figure 2.4: Using a grid to map image coordinates to world coordinates

If necessary, we can associate the top-left circle of the grid image to a different position within the real-world coordinate system. This could be done by considering an appropriate offset. In our research, we have shifted the origin from the top left corner to the center of the field of view. When an image is associated with a calibration object, if we change the relative coordinate system or relative camera position after association, the change will not affect the image.

2.3 Blob analysis procedure

In real surgery, the instruments might enter the field of view from different angles, therefore, using pattern matching techniques such as the normalized grey scale

correlation method would not be useful for tracking surgical tools since the projected shape of the tracked object might change at different angles. In order to track surgical tools, we have to use a method to estimate their real time 3D coordinates regardless of their orientation and shape in the image.

Blob analysis seems to be an appropriate image processing method for tracking surgical instruments. It allows us to identify connected regions of pixels within an image, and then calculate selected features of those regions. These regions are commonly known as blobs. Blobs are areas of touching pixels that are in the same logical pixel state. This pixel state is called the foreground state, whereas the alternate state is called the background state. Typically, the background has the value of zero and the foreground has a value other than zero. Since computation is time-consuming, blob analysis is often performed as an elimination process whereby only blobs of interest are considered in further analysis. Below is a series of major steps that we have used to perform a blob analysis:

1. Grab or load an image that was captured under the best possible conditions by focusing the lens and adjusting the illumination to minimize the amount of preprocessing required.
2. If necessary, reduce the amount of noise in the image. Based on the existing type of noise in the image, we can reduce it by applying different types of filters. For example we can apply a low-pass spatial filter to the image in order to reduce Gaussian random noise or use a median filter to reduce salt and pepper noise.

3. Segment the image so that blobs are separated from the background and from each other. Typically, this involves binarizing the image so that the background is in one state (zero or non-zero) and the blob pixels are in the other state.
4. Performing an opening operation (for non-zero blobs) or a closing operation (for zero blobs) will remove most of the noise particles without significantly affecting real blobs.
5. Allocate a buffer for blob analysis results.
6. Allocate a feature list (i.e., center, length, breath, feret elongation, etc.) in order to specify the features that should be calculated.
7. Calculate the required features and analyze the results.

2.3.1 A new approach to blob analysis of the surgical tools using coloured images

Blob analysis is usually performed for grey scale images and involves binarizing images. A binarizing operation reduces an image into two grayscale values. In general, these values are 0 and the maximum value in the image (for example, 255 if the image is 8-bit). Binary images are useful when trying to identify geometric patterns and objects in the image. A binarizing operation is performed by comparing each pixel value in the image against one or two specified threshold values. Pixels that meet the specified condition are set to the maximum value in the image while other pixels are set to 0.

Previous studies regarding surgical tool tracking and surgical tools gesture recognition which have also employed the blob analysis method, processed grey scale images of the endoscope to get their results [20][21] . In our study, we use blob analysis

of coloured images to track surgical tools. The basis of image processing is like the blob analysis procedure which is described above. In order to identify blobs in grey scale images, we use a certain threshold value to binarize the image and extract blobs below or above that threshold value. In addition to the threshold value, sometimes we need to extract some complex shape attributes to distinguish different blobs. When processing coloured images, we can clearly specify each blob with its unique HLS values (Hue, Luminance and Saturation) before binarizing the image. Blob analysis of coloured images makes blob identification easier and more accurate especially when we are tracking more than one blob. In grey scale images, sometimes it is difficult to identify blobs which have close grey level values. Moreover, uneven light reflection from the surface and different level of illumination could considerably alter the accuracy of blob identification in grey scale images. In the following sub-section, we describe a method to identify blobs in coloured images and how to overcome the illumination problem.

2.3.1.1. Grabbing and processing coloured images

In order to simulate the real surgical procedure, we need to capture real time coloured images of the endoscope. The live images from the endoscope are grabbed into a color image buffer (multi-band or RGB) via a color digitizer for further processing. When grabbing from a color digitizer, each component is transmitted simultaneously and the destination buffer has the same number of bands as the digitizer. The data is simultaneously stored in the appropriate component of the image buffer. When grabbing RGB, the red component is stored in the first color band, the green component is stored in the second color band, while the blue component is stored in the third color band. To

display a color image buffer, the first band is routed to the first output channel (usually red), the second band is routed to the second output channel (usually green), while the third band is routed to the third output channel (usually blue).

During the processing of coloured images, sometimes it is required to perform colour conversions such as converting an RGB image into a HLS (Hue, Luminance and Saturation) image and vice versa. As shown in figure 2.5 a-b, in the HLS color space, color is represented as a combination of hue, luminance, and saturation. The hue corresponds to the wavelength of the main colour and is represented as an angular position on a circular color disk. The luminance corresponds to the brightness of the color, while the saturation can be thought of as the measure of colour purity or concentration. The HLS color space is similar to the human way of describing colours. Each colour has its own hue value (such as red, orange, or green). Once the hue value is chosen, changes to the luminance or the saturation alter only the colour quality and not the basic color.

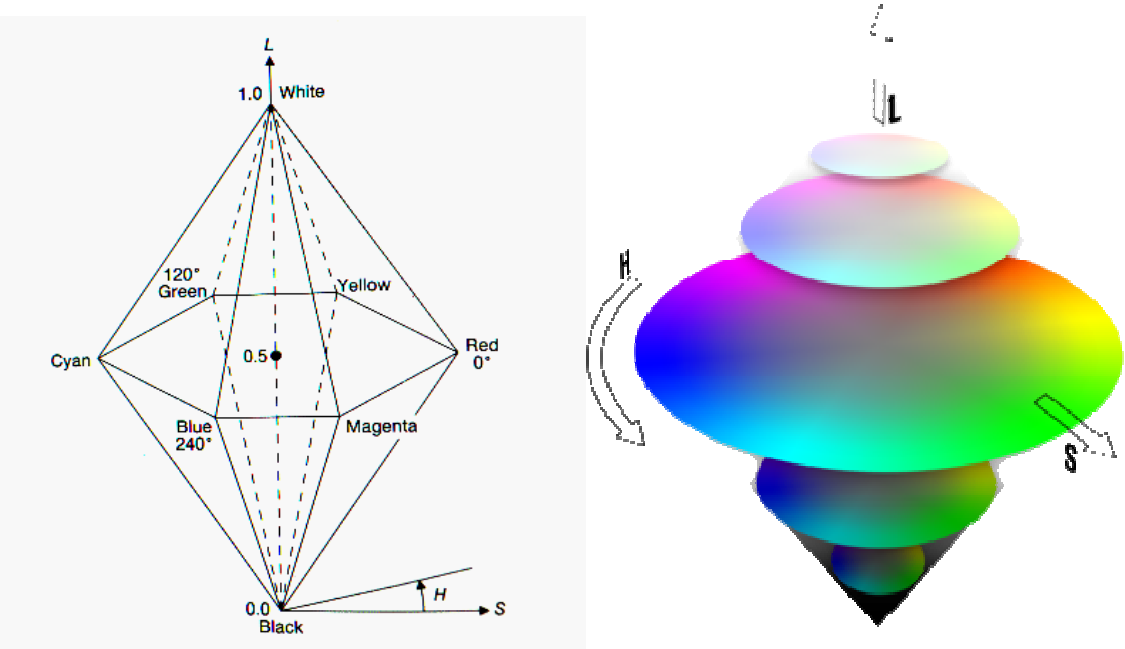


Figure 2.5 a-b : HLS color space

When converting RGB to HLS, we can calculate just the hue component of the HLS into a one band buffer. We can also extract the luminance (intensity) from an RGB image or copy the luminance component of an image into a three-band buffer to create a monochromatic (gray) RGB buffer.

2.3.1.2. Blob analysis for coloured images

Unlike previous studies in this area, in this research, tracking of surgical tools is accomplished through blob analysis of coloured images of such tools. The stainless steel tip of the surgical tool does not provide even light reflection because of its curved surfaces. The darker areas would break up the tool tip blob into multiple blobs in the binarized image. Smooth light reflection from a surgical tool tip is critical in segmenting the tool tip blobs. To have a better contrast between the surgical tools and the background and also to make the light reflected from the stainless steel surgical tools more evenly distributed, the tips of the surgical tools were coated with paint.

Since the live images from the endoscope are captured coloured during the actual surgery, we also capture color images of the endoscope while processing them as binary images. Therefore, the endoscope images which are displayed in the computer monitor are coloured. In order to discriminate and track a particular object with a specific color in the image, we need to find the HLS values for that specific color. Then we can binarize the coloured images based on these HLS values. Any objects with colors other than the defined HLS values will be set to zero (black) and only objects with those specific HLS values which also meet the defined size and shape criteria will be detected and tracked.

Figure 2.6 shows a small blue marker at the tip of the surgical tool in a cluttered image of the field of view.

Figure 2.6: Blue marker in a cluttered scene

Here, we are only interested in tracking the blue blob at the tip of the surgical graspers. As shown in figure 2.7, we are able to eliminate the unwanted blobs and only extract and track the desired blob that meets our criteria after binarizing the image.

Since the endoscope camera and its light source are co-located, the illumination of the site would not be even throughout the entire field of view and the blobs may not be detected properly as they get farther from the center of the field of view. In order to fix this problem, when converting RGB image to HLS, we only consider hue and saturation values so that the colors of a desired blob to be detected and tracked are only defined on the basis of their wavelength and their purity and therefore regardless of their illumination.

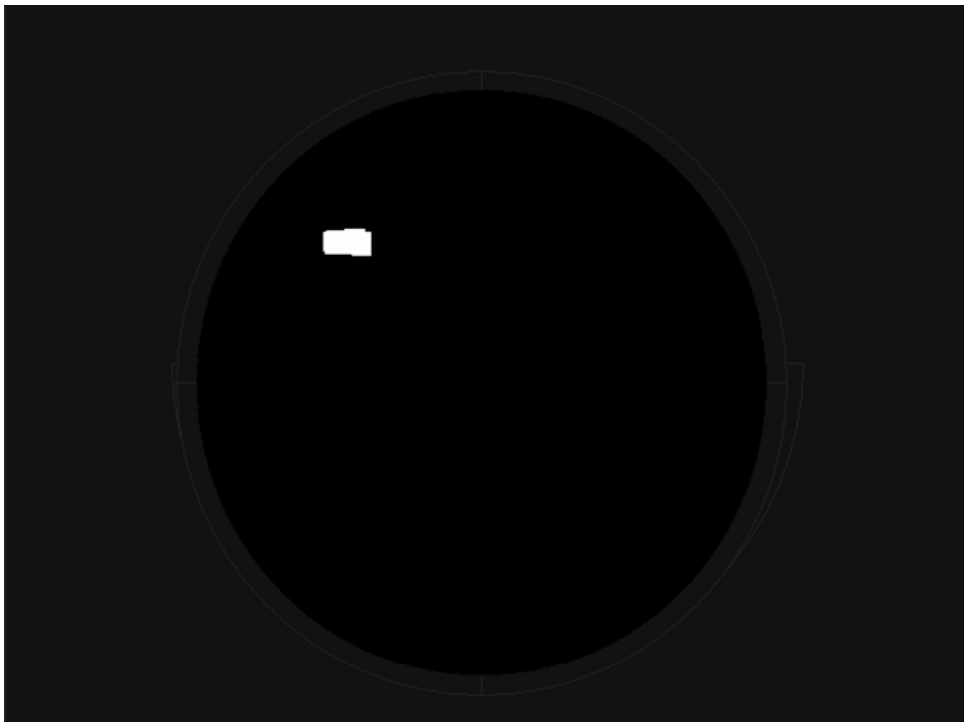


Figure 2.7: Extracting the desired blob form the cluttered scene

2.4 Tracking surgical instruments

Our tracking system uses endoscope calibration parameters to determine the spatial information of the surgical tools' tip and estimate the 3D coordinate of the marker at the tip of the surgical tool. By calculating the marker's diameter, the system can estimate how far the marker is from the endoscope (depth) in millimeters. It then uses the estimated depth to find an appropriate pixel to millimeter conversion factor and calculate the x and y coordinate of the marker at the tip of the surgical tool.

2.4.1 Selecting appropriate blob features for instrument tracking

In many applications, we are interested only in blobs whose features satisfy certain criteria. Since the computation is time consuming, blob analysis is often

performed as an elimination process whereby only blobs of interest are considered in further analysis. The MIL package includes a blob analysis module that can extract a wide assortment of blob features, such as the blob center of gravity (location in the image), blob area, blob breadth (width) and length, blob perimeter, feret diameter at a given angle, feret elongation, and compactness [22].

In order to track the surgical tools in the field of view of the camera, we need to calculate certain blob features which provide us with 3D coordinates of the tip of the surgical tools. Before selecting a feature for calculation, we should take the blob shapes into consideration. Some features are more appropriate for certain blob shapes than for others. For example, if the marker shape is round, the size of breadth or length cannot be used to obtain depth unless the marker is perpendicular to the endoscope. The strip marker chosen for this experiment is similar to the strip marker proposed by [20][21]. Unlike their model which tracked the black marker on a white background, or white marker on a dark background, we used a coloured marker instead. As illustrated in figure 2.8, M is the diameter of the surgical grasper cross section which is 5mm, and d is the width of the marker, which is 2 mm. Therefore the condition $\frac{M}{3} < d < \frac{M}{2}$ is satisfied [20].

Figure 2.8: The strip marker located at the tip of the graspers

The shape of the marker is simple and its color has a good contrast with the background. In fact, the surgical tool's diameter is used to calculate the depth. Regardless

of how the instrument rotates along its axis at any specific depth, there are always two peripheral points on the markers that can represent the tools diameter. Shown in figure 2.9, because of the perspective view, as the marker gets closer to the camera it appears bigger and occupies more number of pixels on the screen and as it gets farther away from the camera it becomes smaller and occupies a smaller number of pixels on the screen.

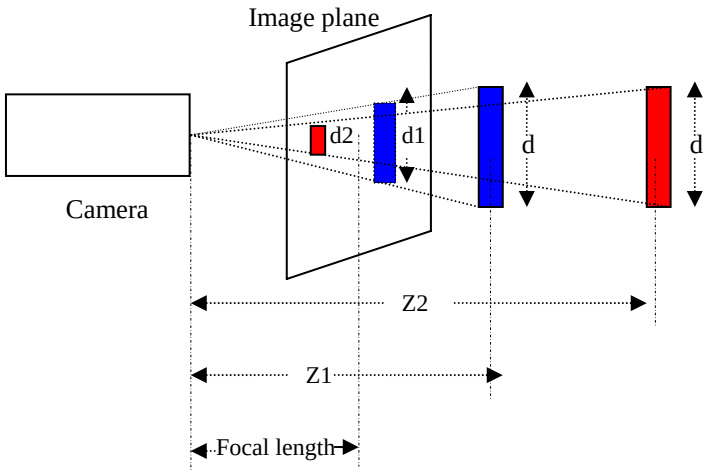


Figure 2.9: Perspective view of the camera, the marker becomes bigger as it gets closer to the camera and vice versa

After finding the focal length of the camera we can calculate the value of the depth (z) from the diameter of the marker. The calculated breadth of the marker (tool diameter) in pixels is the same as $(V_i - V_{io})S_x$ and $(V_j - V_{jo})S_y$ in equation (1) in section 2.2.1. Based on the geometry shown in figure 2.9, we can modify equation (1) and write it based on the marker's diameter, so the new equation will be as follows:

$$\frac{d}{Z_1} = \frac{d_1}{f} \quad \text{or} \quad Z_1 = \frac{fd}{d_1} \tag{2}$$

Where d is the actual diameter in mm, d_1 is its size on the screen in pixel units and Z_1 is the depth in mm. By incorporating the focal length of the camera, our tracking system calculates the real time diameter of the marker and estimates its real time depth from the above equation. The x and y coordinates of the tip of the surgical tool are also estimated by calculating the center of gravity of the blobs in real time.

2.5 Accuracy of the tracking system

As shown in figure 2.10, in order to measure the accuracy of the tracking system, the 3D coordinates of the marker at the tip of the surgical tool was calculated by placing it at the center of 9 adjacent dots from the calibration grid. The 3D coordinates of the dots were estimated at different depths from 30-100 millimetres.

In addition to these measurements, the 3D position of the marker at the tip of the surgical tool was also calculated at the top of 5 rods with different heights as shown in figure 2.11. For this purpose, we placed 5 rods with the diameter of one centimetre but with different heights at different locations on the calibration grid. Since the original program only estimates the z coordinate in real world units and estimates x and y coordinates in pixels, we convert pixels to real world coordinates before we measure the accuracy of the tracking program. The following sub-sections describe pixel to millimetre conversion method as well as the result of estimating 3D coordinates using the developed tracking program.

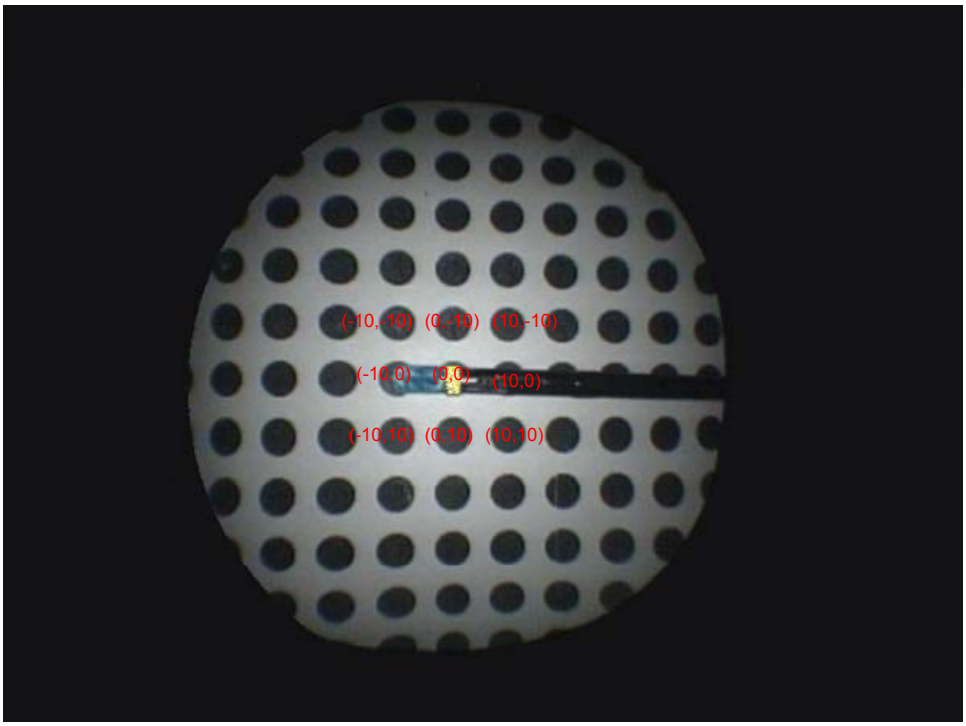


Figure 2.10: The 3D coordinates of the marker were calculated at 9 adjacent dots

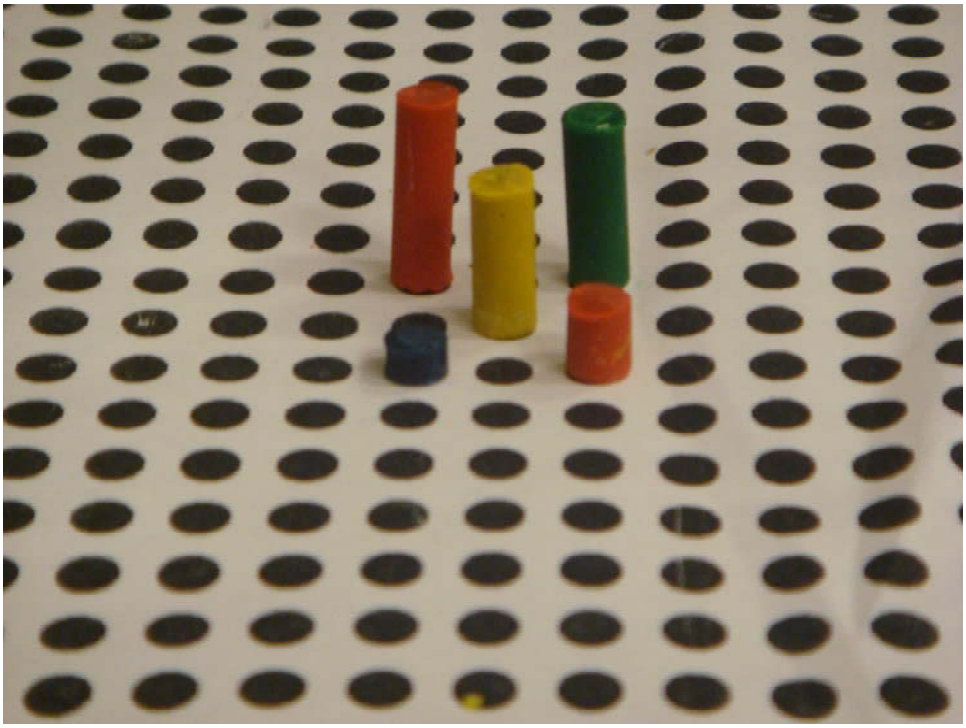


Figure 2.11: The 3D coordinates of the marker were calculated by placing it at the top of each rod

2.5.1 Pixel to millimetre conversion factor

To measure the accuracy of the tracking program in calculating x and y coordinates, we had to convert pixel units to millimetres. Figure 2.9 shows as the marker becomes closer to the camera, its image becomes bigger on the screen, therefore, the number of pixels that it occupies on the screen increases and as it gets farther from the camera the number of pixels that it occupies on the screen decreases. In order to calculate the pixel-to-millimetre conversion factor at various vertical distances from the camera (depths), as shown in figure 2.12, the separations between the centres of two black dots were measured at various depths from 30-100 mm. The centres of the two adjacent black dots were separated by 10 mm on the calibration grid.

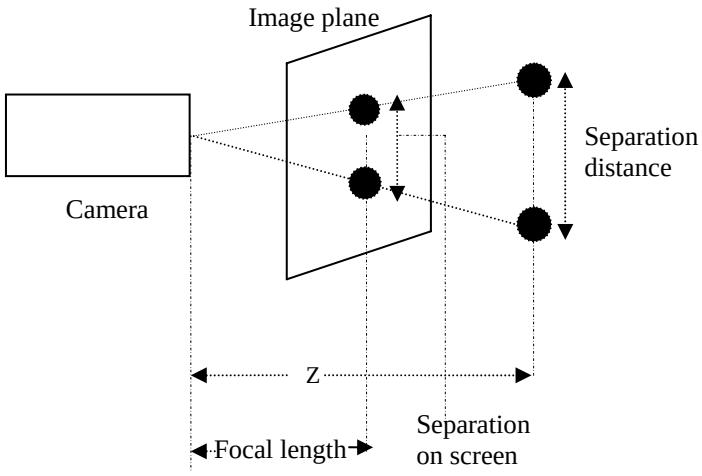


Figure 2.12: Separation between two dots in mm vs. separation on screen in pixels

The number of pixels on the screen representing the 10 mm separation is recorded at each depth. Based on the calculated z value, the program automatically uses the appropriate pixel to millimetre conversion factor using linear interpolation. Figure 2.13 shows the pixel to mm conversion chart.

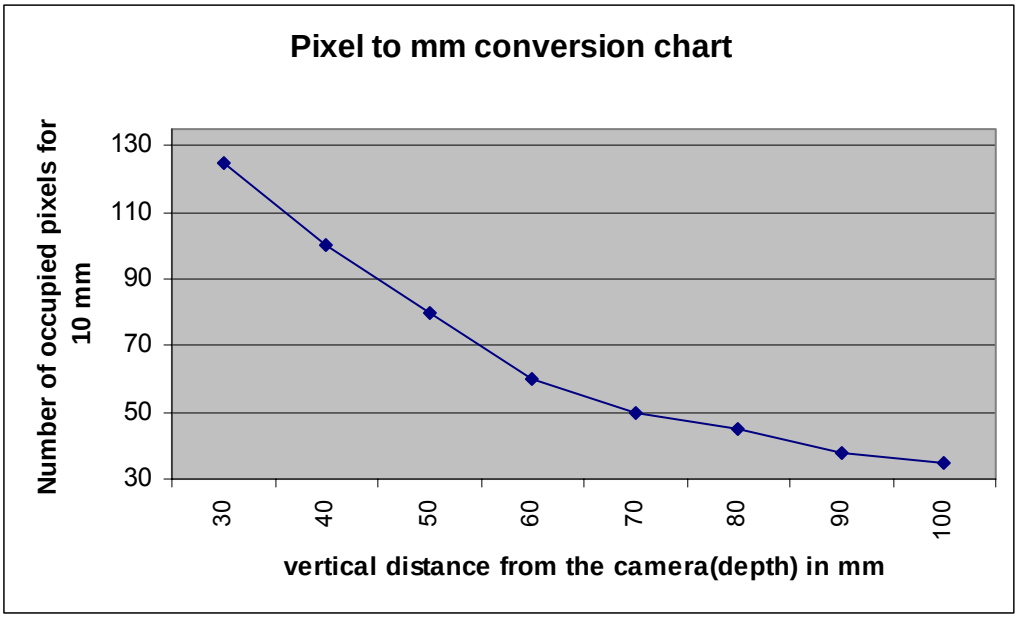


Figure 2.13: Pixel to millimetre conversion chart

2.5.2 Accuracy in calculating coordinates

Table 2.1 shows the result of estimating the x,y and z coordinates of the marker using the tracking program. The estimated x,y and z values of the marker were compared with the actual values of x,y and z of the centers of the dots or rods on the calibration grid and the average errors were calculated for different depths.

Figure 2.14 summarizes the accuracy of the tracking program in estimating the x,y and z values of the marker at the tip of the surgical tool in real world units.

Actual z (mm)	Calculated z (mm)	z error (mm)	x error (mm)	y error (mm)
30	33	3	1.5	1.5
40	43	3	1.5	1.5
50	53	3	1.5	1.5
60	63	3	1	1
70	72	2	1	1
80	82	2	1	1
90	87	-3	1.5	1.5
100	97	-3	1	1

Table 2.1: Error in estimating the x,y and z coordinates

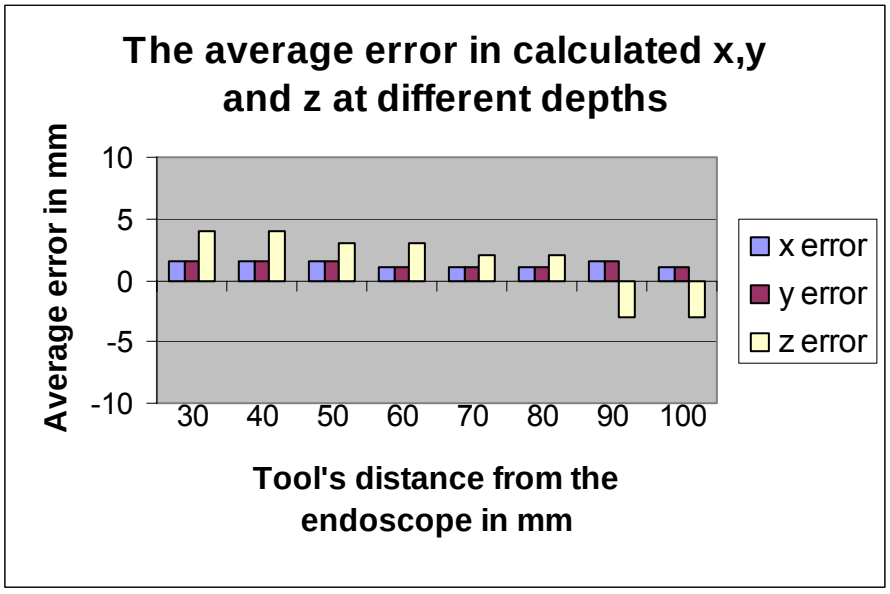


Figure 2.14: Average error in estimating the x, y and z values of the tip of the surgical tool

These results shows that our tracking program can estimate the position of the tip of the surgical tools with an average accuracy of about a millimetre in both x and y directions while the average error in estimating the z coordinates is slightly over 2

millimetres. The major cause of these inaccuracies is the error in estimating the breadth (width) of the marker. Factors such as uneven light reflection from the surfaces, level of illumination, camera distortion and depth could affect estimating the width of the marker. Due to the camera distortion, the error increases as the marker becomes farther from the centre of the filed of view. It also increases as the vertical distance of the marker becomes too close or too far from the endoscope. It is very important to track the surgical tools correctly since the accuracy of the proposed segmentation method in chapter 3 and also the usability of the proposed augmented systems in chapter 4 highly depend on it. Depending on the type of application, these errors might be within the reasonable range. For example, the accuracy of the tracking system might be suitable for laparoscopic surgery but may not be accurate enough for some neurological operations.

Chapter 3

Segmentation of live images using surgical tools

In this chapter, we propose a novel method for segmentation of live endoscope images. Section 3.1 briefly explains image registration and segmentation methods. In section 3.2, we propose a new segmentation method which allows users/surgeons to use the surgical tools as a 3D input device and select the seed points along the boundaries of the desired organ. Then, by applying a cubic spline algorithm to the coordinates of the selected points, the contour of a desired object will be extracted. In section 3.3 we conduct user studies and evaluate the accuracy of the result of segmentation using the proposed method. The results of the performed user studies have raised some important questions which have led to further improvement of our design. In the remainder of the chapter, we bring in two major improvements to our preliminary design. In order to facilitate the segmentation procedure, in section 3.4, we replace the computer mouse with a wireless switch which is attached to the surgical tool handle. Then in section 3.5 we introduce the cube overlay system which could be used to correct the extracted boundaries and fit them closer to the actual boundaries of the organ of interest. At the end of this chapter, we discuss possible applications of the cube overlay system in the registration of pre-operative scans with live endoscope images.

3.1 Image segmentation and registration

Image registration is finding an optimal transformation between two or more images which are taken at different times, with different modalities or different points of view. Image registration is a fundamental task in image processing and is a vital

component of a large number of applications. Clinical applications of image registration are not only within diagnostic settings, but prominently in the areas of planning, carrying out and evaluating surgical and radio-therapeutical procedures. Since the information obtained from different medical imaging modalities are complementary, proper integration of useful data taken from separate images is often desired and registration of the images from practically any combination will benefit the surgeon [23].

Based on the nature of registration basis, registration methods are classified into two categories: extrinsic and intrinsic methods. Extrinsic registration methods are based on artificial objects introduced into the image space. These objects should be designed to be well visible and accurately detectable in all of the imaging modalities. Due to the invasive nature of this type of image registration, it would not be an option to register pre-operative scans with intra-operative images of the endoscope. Intrinsic registration methods are on the other hand based on the image information as generated by the patient. Registration can be based on a limited set of identified salient points (visible anatomical landmarks) or on the alignment of segmented binary structures or segmentation basis. Anatomical landmarks are not always available, for example, in cases such as laparoscopic surgery and detecting them for image registration is not always feasible, especially when the tissue deforms. If there are not enough visible landmarks in both images, segmentation based registration could be used where anatomically the same structures are extracted from both images to be registered and used as a sole input for the alignment procedure. Therefore, segmentation based registration could be an option to register pre-operative scans with live laparoscope images.

Segmentation is a process that separates objects in an image. A large class of registration methods refer to as feature based methods which require some features to be segmented. These features may be detected using simple methods of edge detection or segmented using higher level methods that are customized for specific anatomical structures. Feature based registration methods consist of the four following steps: feature detection (i.e. closed boundary regions, edges, contours and etc), feature matching, transform model estimation and image re-sampling and transformation. The accuracy of the first step (feature detection) is very important for the result of registration and if it's not performed accurately, problems can arise in the following registration steps [24].

Manual segmentation is often used for evaluation of automatic or semi-automatic segmentation. It is often believed to be very accurate if the specialist segmenting the object is highly experienced. Therefore the manual segmentation by a physician or a radiologist is often used for the evaluation of automatic segmentation algorithm [25].

Proper segmentation of specular images is a very challenging and complicated procedure. Robust and fully automatic method of extracting contours from structures in noisy or cluttered images is very difficult and there are still limitations with automatic segmentation of specular surfaces [26]. Figure 3.1.a shows a typical image from inside the abdominal area in laparoscopic surgery. Strong shading which causes inhomogeneous appearance of the images in endoscopic surgery and also specular reflection of the surfaces, interfere with segmentation of the live images from the endoscope. Figure 3.1.b shows the result of unsuccessful automatic segmentation of figure 3.1.a which has resulted in false contouring effect. Figure 3.1.c shows the desired contour of the organ in figure 3.1.a which is traced manually in figure 3.1.b.

Figure 3.1.a: The laparoscope image from inside the abdominal area

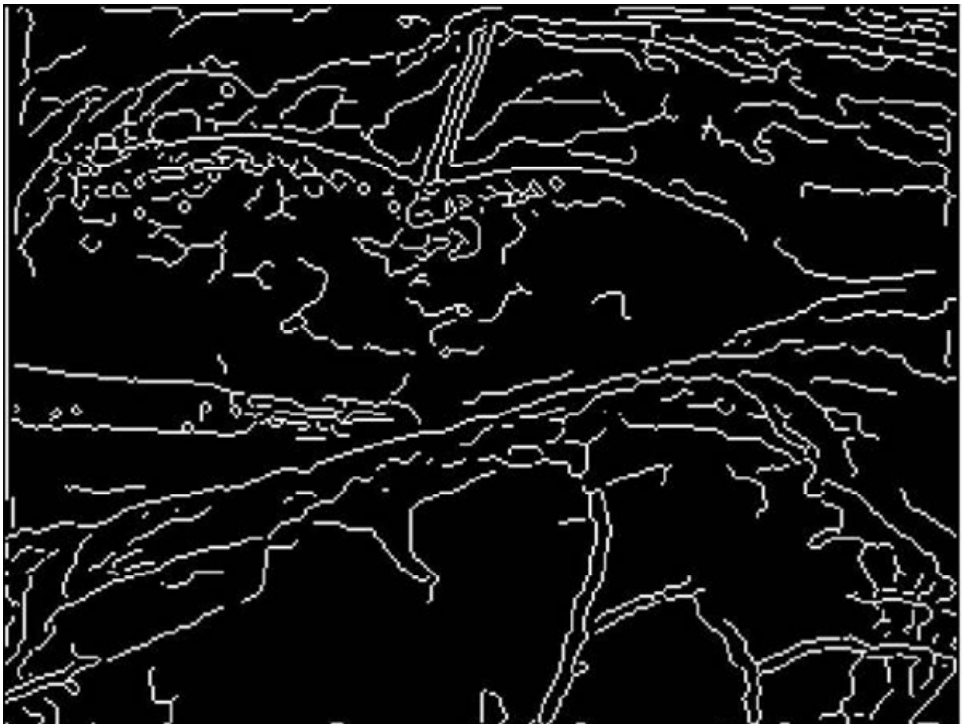


Figure 3.1.b: Unsuccessful automatic segmentation of figure 3.1a

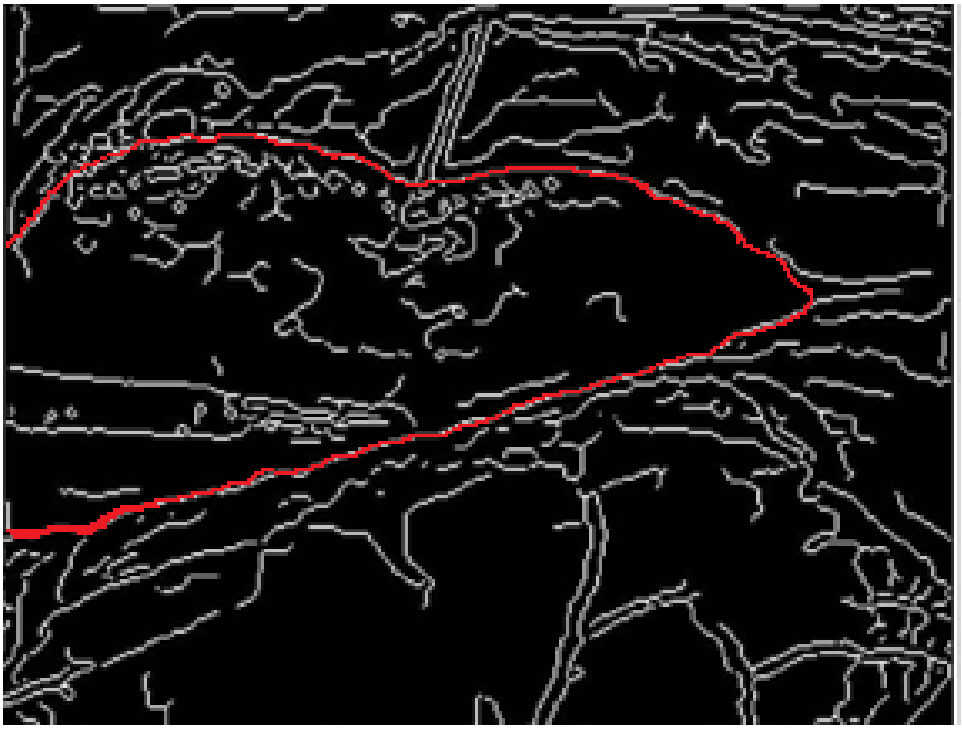


Figure 3.1.c: The desired contour of the organ in figure 3.1.a is manually traced in figure 3.1.b

Hence, it is favourable to have a segmentation method which could eliminate these effects. In the following section, we propose a new method of segmentation for live endoscope images that could be used toward registration of live images with pre-operative scans during surgery.

3.2 Segmentation of live endoscope images

In this section, we propose a novel method for segmentation of live endoscope images using tracked surgical tools. First, the seed points are selected by a tracked surgical tool, then the space between the seed points are interpolated and the seed points are connected using cubic splines to extract the contour of the object.

3.2.1 Selecting the seed points along the boundary

The first step in extracting the contour of an object is to select some points (seed points) all along the boundary of the desired organ which seem closer to its actual boundary. The developed tracking system which was described in chapter 2 tracks the position of a marker located at the tip of the graspers and estimates its real time 3D coordinates. As described in section 2.4.1, we track the position of a strip marker because it represents the tool diameter while its width remains constant regardless of the tool orientations. The shape of the tip of closed graspers is almost cylindrical; therefore its diameter in the image remains constant at different tool orientation. Since we use closed graspers to select the seed points, we can directly calculate the coordinates of the tip of the surgical tool instead of the strip marker.

To allocate the seed points along the boundary of an organ, a program has been developed in C++ using Matrox Imaging Libraries (MIL), which tracks the 3D position of the tip of a graspers with the same method as described in chapter 2 and allows the user to select a desired point by clicking on a computer mouse. Therefore, the user can select the seed points by touching a specific location along the boundary of the desired organ with tracked graspers and confirming the specified point with a mouse click. As shown in figure 3.2. a-c, a “cross” shows the surgical tool’s tip centre and it will be overlaid on the image upon a mouse click and this procedure continues until the intended number of seed points are selected along the boundary of the desired organ. The coordinates of the selected points are also saved for future use. The next step is to smoothly connect the seed points and extract the contour of the object. The following subsection describes the method used to extract the contours.

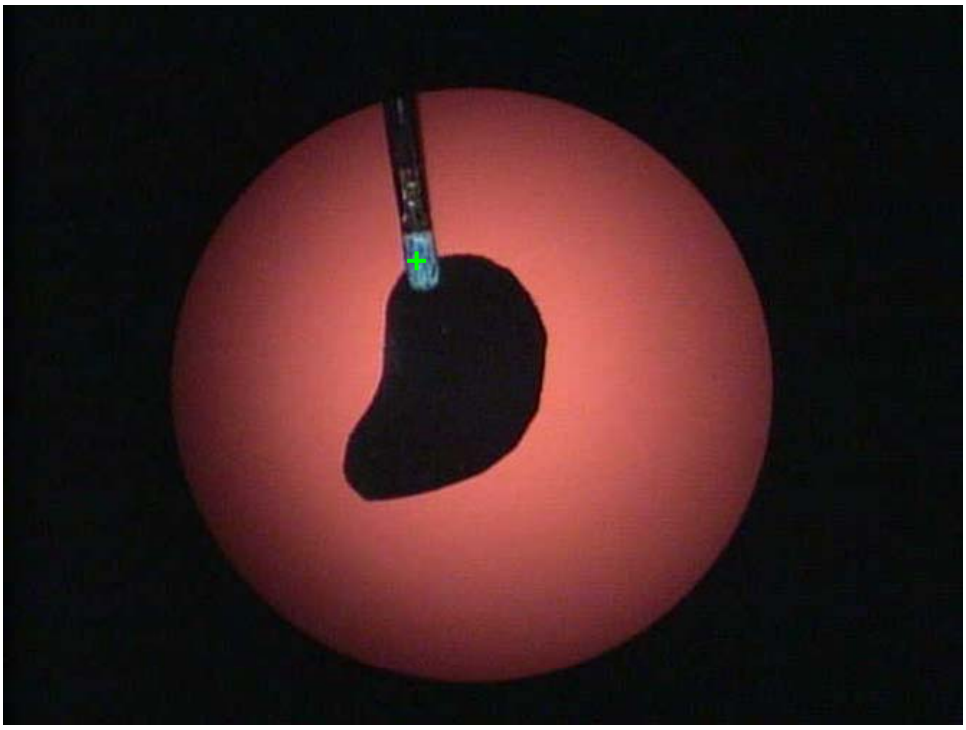


Figure 3.2.a: The first seed point is selected

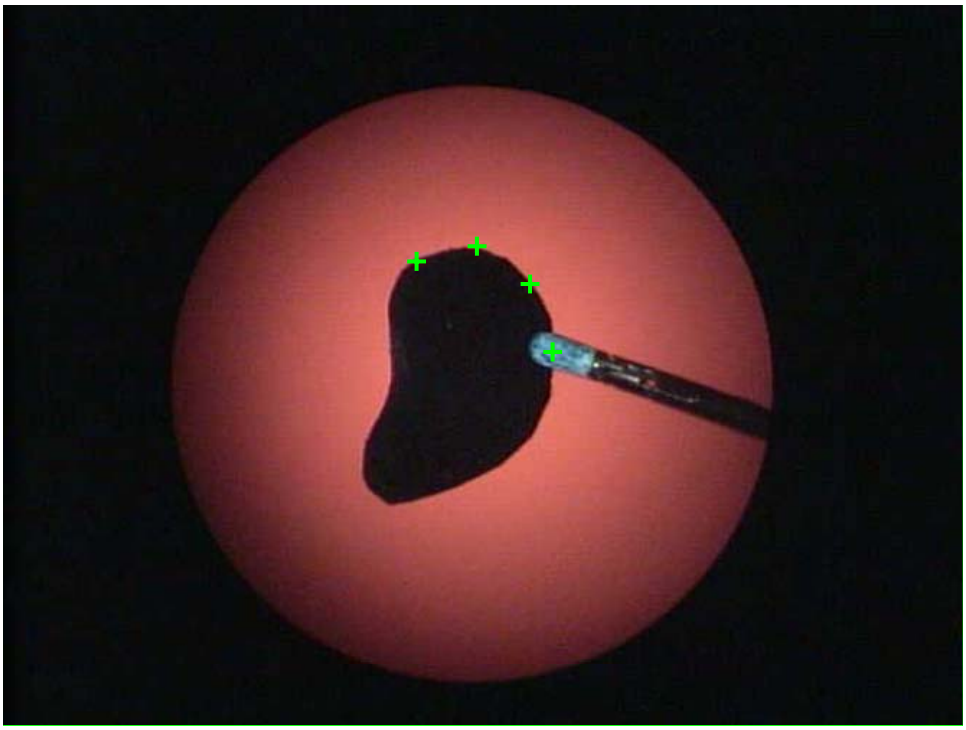


Figure 3.2.b: The seed points are selected along the boundary of the organ

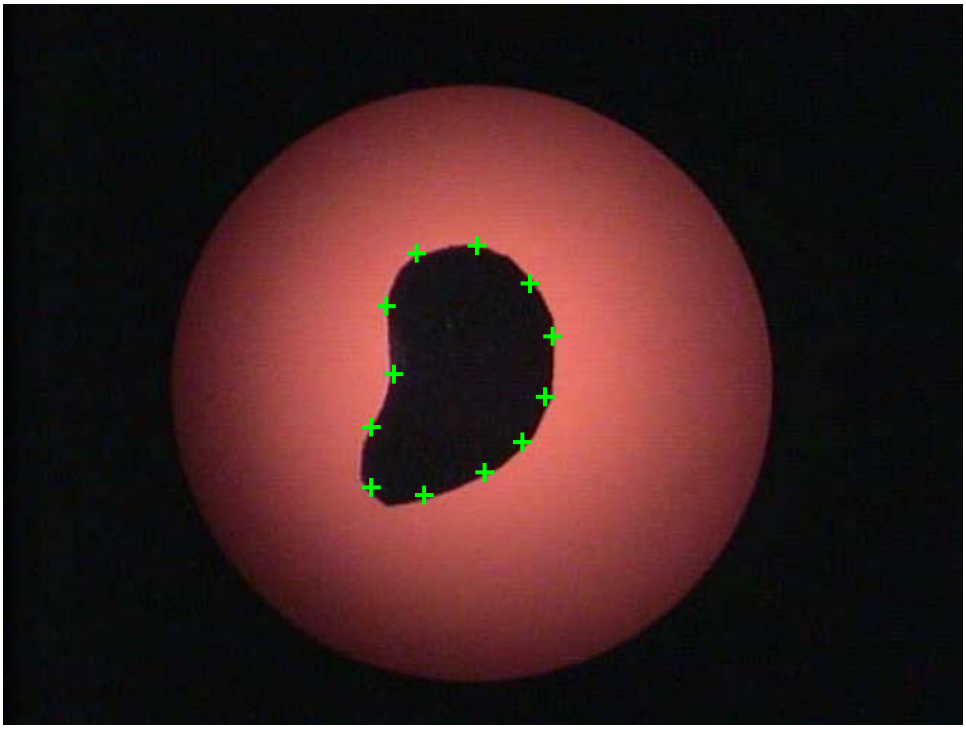


Figure 3.2.c: All seed points are selected along the boundary of the organ

3.2.2 Extracting the contour of a desired object

After selecting all the seed points along the boundary of the object and recording their coordinates, we need to smoothly connect the seed points and interpolate between them and extract the contour of the object. For this purpose, we have used cubic splines since they are drafting aids which are able to draw smooth and continuous curves and connect widely spaced points. Cubic splines are popular models in computer graphics applications because of the simplicity of their construction, their ease of use and accuracy of evaluation, and their capacity to approximate complex shapes through curve fitting and interactive curve design.

3.2.2.1 Cubic splines theory

Suppose we have a set of numeric seed points [x_i, y_i] for $i=0,1,\ldots,n$ for the function $y=f(x)$. Therefore, we have "$n+1$" seed points and "n" intervals between them. The cubic spline interpolation is a piecewise continues curve passing through each of these points. In addition to interpolating all the data, the functions and their first and second derivatives must be continuous across the intervals and equal at the endpoints. The fourth derivative of all equations must be zero.

There is a separate cubic polynomial for each interval, each with its own coefficients:

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x\ )^2 + c_i(x - x_i) + d_i \quad \text{for } x \in [x_i, x_{i+1}]$$

Together these polynomial segments are denoted $S(x)$, the spline. Since there are "n" intervals and four coefficients for each we require a total of "$4n$" parameters to define the spline $S(x)$. We need to find "$4n$" independent conditions to solve for them. We get two conditions for each interval from the requirement that the cubic polynomial match the values of the set at both ends of the interval:

$$S_i(x_i) = y_i; S_i(x_i + 1) = y_{i+1}$$

We still need "$2n$" additional conditions. Since we would like to make the interpolation as smooth as possible, we require that the first and second derivatives to be continuous:

$$S'_{i-1}(x_i) = S_i'(x_i), S''_{i-1}(x_i) = S''_i(x_i)$$

Cubic splines are popular because they are easy to implement and produce a curve that appears to be seamless. However, if the application is sensitive to the smoothness of derivatives higher than second, cubic splines may not be the best choice.

3.2.2.2 Connecting the seed points

Due to the limited graphics capabilities of Matrox Imaging Libraries, this part of the study was performed off line and outside the Matrox and C++ environment. A program was written in Matlab environment which takes the 2D coordinates of the selected seed points and creates a plot by smoothly connecting the points using a cubic spline function. In order to get a smoother contour, the interpolations between the points were made at 0.01 intervals.

Figure 3.3.a shows how smoothly cubic spline can connect the points compared to figure 3.3.b where seed points are connected just with straight lines.

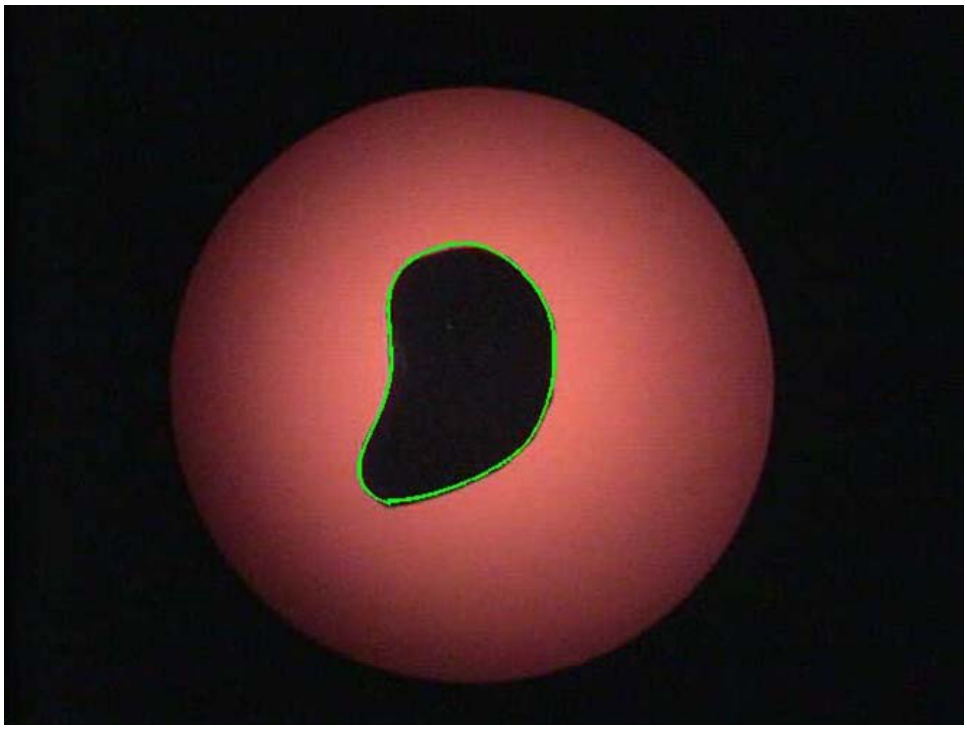


Figure 3.3.a: Contour of the object is extracted by applying cubic spline to the seed points

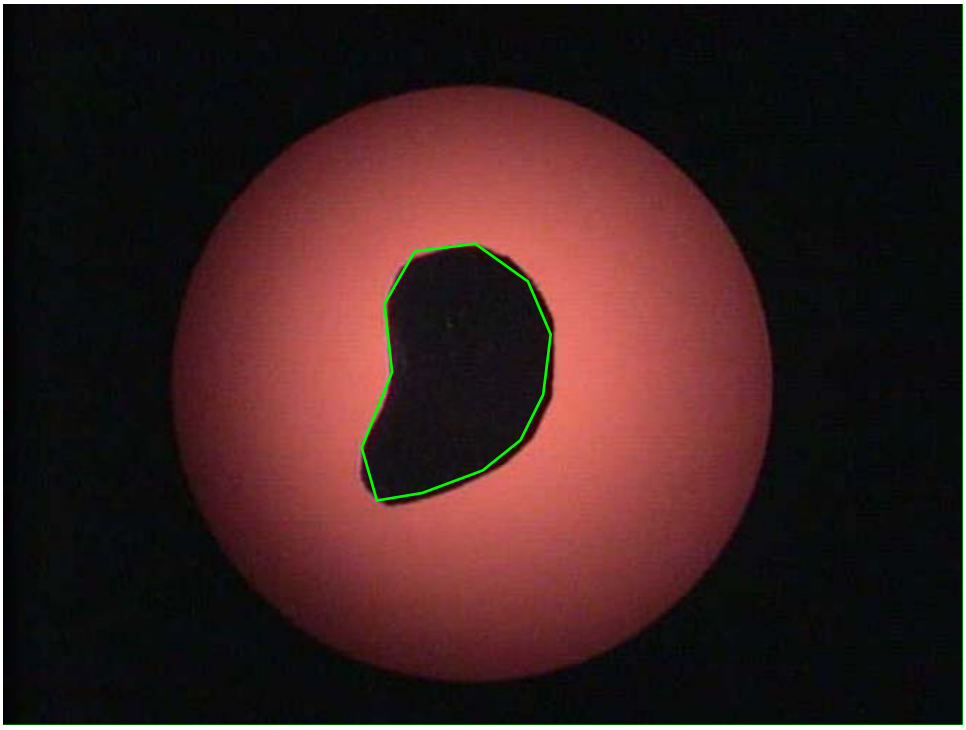


Figure 3.3.b: Contour of the object is extracted by connecting the seed points with straight lines

3.3 Accuracy of image segmentation with the proposed technique

To test the practicality and accuracy of the segmentation method, we have conducted some user studies. However, before performing any user study, we need to define an evaluation criterion to test the accuracy of segmentation results, and also run a pilot study to test the practicality of this method.

3.3.1 Evaluation criterion

The criterion for evaluating the segmentation result is the number of pixels captured by a particular object at a particular depth. Therefore, in order to evaluate the accuracy of our results, we compare the number of captured pixels inside the extracted contours with the actual number of pixels captured by the same object at the same depth.

For this purpose, a program was written in C++ and MIL, which calculates the actual number of pixels occupied by an object. The object of interest was placed on a clear background and its actual numbers of pixels were calculated for different depths. These numbers were used as base values to evaluate the segmentation results. In order to count the number of pixels inside the extracted contours, the area inside each plotted contour was filled with a particular colour to make the pixel count possible. All created plots and images of the object had the same resolution of 480x640.

3.3.2 Preliminary results

Before recruiting the subjects to perform the segmentation task, a preliminary study was conducted to ensure the feasibility of the proposed method. In general, the abdominal area is insufflated during the actual laparoscopic surgery which creates about 3-4 litres of empty space inside the patient's abdomen. This procedure not only facilitates the performance of the surgical task but also reduces the chances of injuring the surrounding tissue during the surgery. To mimic the actual surgical procedure, this study was performed inside a cardboard box with the dimensions of 20x12x15 cm^3. For this study, the experimenter performed the segmentation task at different depths and with different numbers of seed points. Contours of a unique shapeless object were extracted at different depths and for different numbers of seed points and the number of pixels inside the extracted contours were calculated. These numbers were averaged over several trials and compared with the actual number of pixels occupied by that object. Table 3.1.a-c and figures 3.4.a-c summarize the result of pixel count inside the extracted contours for three sets of experiments for 12, 10 and 8 number of seed points, respectively. The results were

averaged over 5 trials at each depth, and their mean values, average error% and standard deviation were calculated.

Depth	Calculated # of pixels for 5 trials					Mean	Actual # of pixels	Error %	Std
50mm	39984	39785	38749	39876	39678	39614	40935	3.23	496.77
60mm	30986	31167	30163	30536	30983	30767	31812	3.28	410.12
70mm	22456	22876	22434	23003	23019	22758	23479	3.07	290.80
80mm	17983	17648	17529	17325	17547	17606	18170	3.10	240.97
90mm	13879	13985	14238	13694	13798	13919	14375	3.17	207.96
100mm	11456	11567	11098	11687	11760	11512	11880	3.08	259.65

Table 3.1.a : Calculated number of pixels inside the extracted contours using 12 seed points

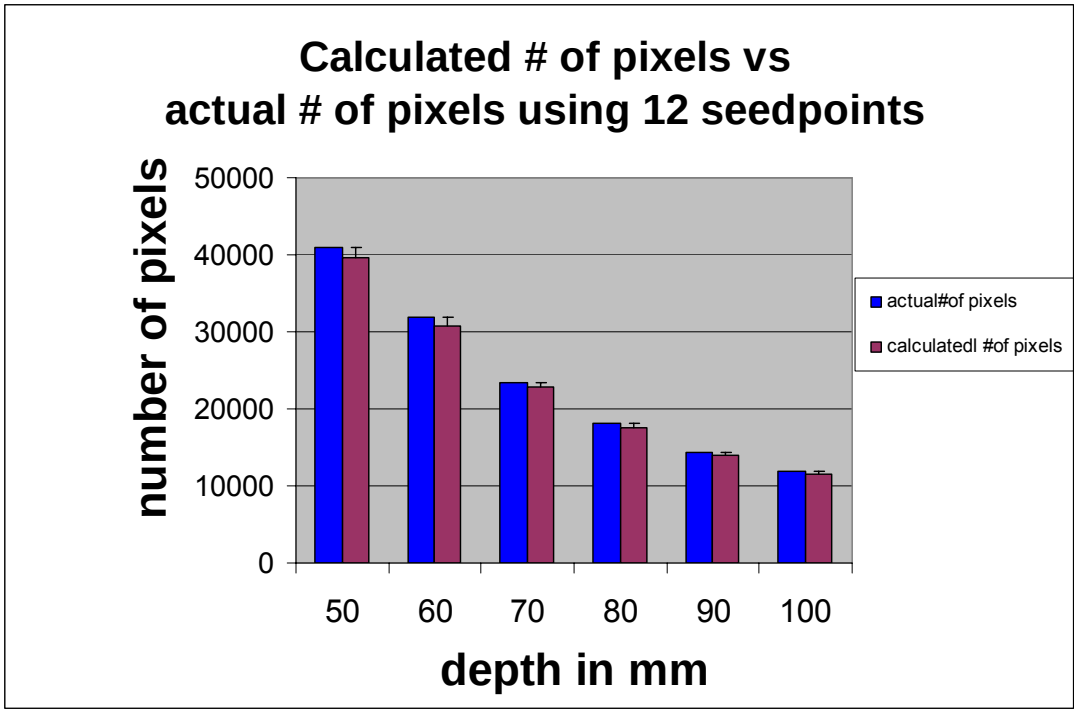


Figure 3.4.a : Number of pixels inside the extracted contours using 12 seed points vs. the actual number of pixels

Depth	Calculated # of pixels for 5 trials					Mean	Actual # of pixels	Error %	Std
50mm	38203	39128	37620	38760	38267	38396	40935	6.20	575.41
60mm	29345	30123	29217	30381	29961	29805	31812	6.31	503.63
70mm	22009	22457	21654	22098	21817	22007	23479	6.27	304.75
80mm	16873	16549	17013	17786	17067	17058	18170	6.12	454.26
90mm	13476	13234	12873	14005	13785	13475	14375	6.26	446.51
100mm	10984	11238	11450	10678	11345	11139	11880	6.24	310.44

Table 3.1.b : Calculated number of pixels inside the extracted contours using 10 seed points

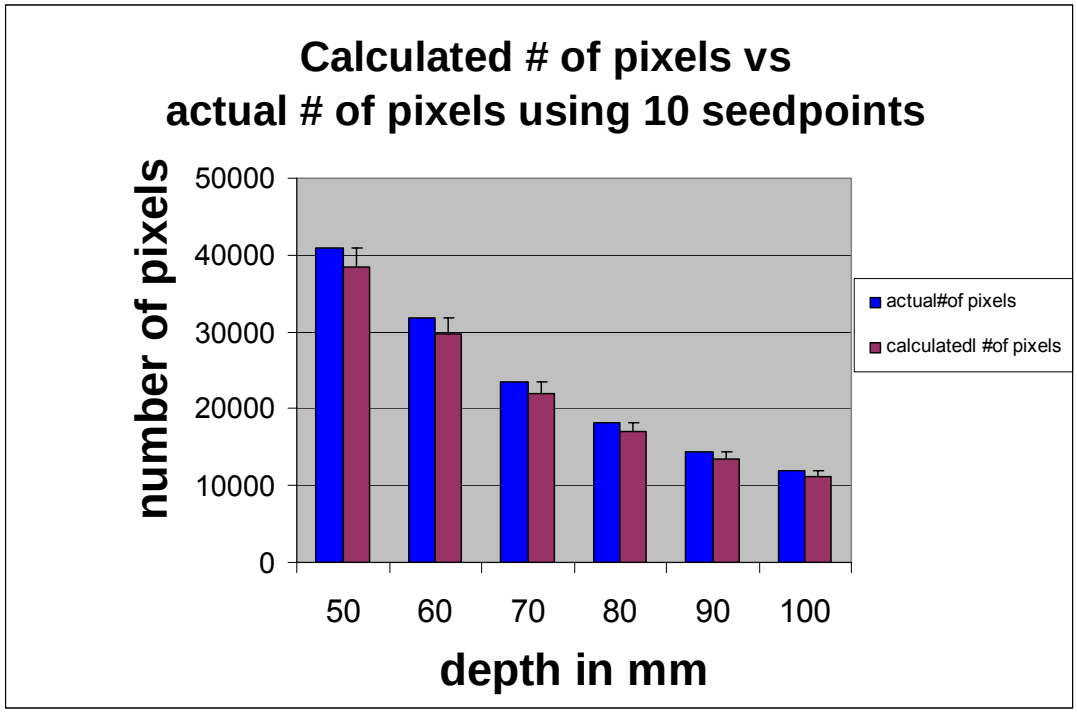


Figure 3.4.b : Number of pixels inside the extracted contours using 10 seed points vs. the actual number of pixels

Depth	Calculated # of pixels for 5 trials					Mean	Actual # of pixels	Error %	Std
50mm	37218	37123	38675	37112	36986	37423	40935	8.58	704.84
60mm	29650	28794	28045	28354	28450	28659	31812	9.91	615.22
70mm	21563	21256	20409	22345	21985	21512	23479	8.38	742.02
80mm	16241	16319	16348	16942	17342	16638	18170	8.43	482.57
90mm	13476	12973	12654	13832	13005	13188	14375	8.26	464.37
100mm	11003	9986	10904	10897	10783	10715	11880	9.81	414.68

Table 3.1.c: Calculated number of pixels inside the extracted contours using 8 seed points

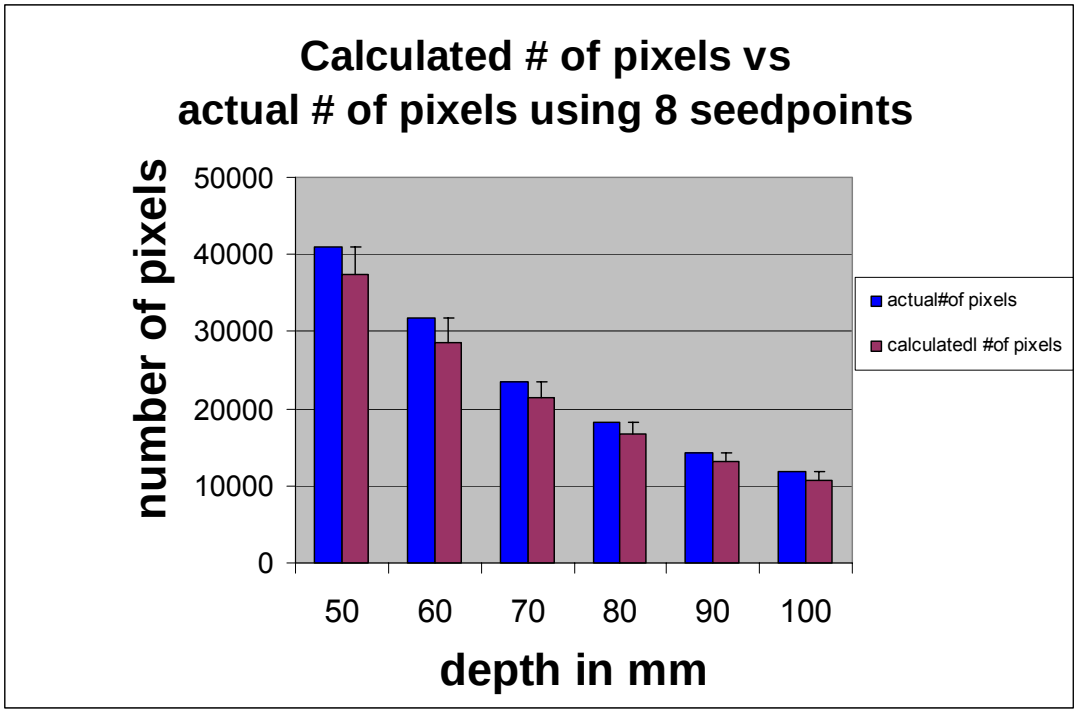


Figure 3.4.c : Number of pixels inside the extracted contour using 8 seed points vs. the actual number of pixels

The results of the preliminary study show that although the segmentation method is feasible, the accuracy of the result is highly dependent on the number of seed points. As the number of seed points increases, the accuracy of segmentation increases at each

depth while the variability of the results decreases. The variability also seems to be decreasing as the depth increases. Averaging over three experiments, the most accurate results (not by far) obtained at 7cm depth from the camera. It is obvious that in addition to the number of seed points, the locations of the selected seed point are extremely important. Therefore, to obtain better results, we need to distribute the seed points evenly along the boundaries of the object. Considering the results of the preliminary study, we performed a user study among university students, which is described in the following section.

3.3.3 Results of the user study

A user study has been conducted among university student where accuracy and variability of the segmentation method have been calculated over several trials. The endoscope camera was fixed at 7 cm depth throughout the entire experiment. The subjects were asked to perform the segmentation task for 12 seed points while trying to align the centre of the tip of the surgical graspers with the boundary of the object. They were also instructed to select the seed points consecutively and at relatively even intervals along the boundary of the object. The participants were asked to familiarize themselves with the experimental setup before the start of the first trial. Table 3.2 summarizes the results of the user study for six participants and six trails. Figure 3.5 shows the percent of actual number of pixels calculated with this method. As shown, the average error in calculating total number of pixels inside the extracted contours is about 3.93% and the shape of the segmented object is preserved relatively well in all trials.

Subject	Calculated # of pixels for 6 trials						Mean	Error%	Std
S1	23129	23651	22390	21456	21309	23987	22654	3.52	1123.43
S2	21562	23965	22390	21805	22071	22768	22427	4.48	865.44
S3	22876	22987	23143	22578	21198	23249	22672	3.44	758.57
S4	22011	24017	22431	21970	22102	22678	22534	4.02	775.74
S5	22987	21765	21098	22879	22890	23104	22454	4.37	823.50
S6	22123	22091	21806	23258	22249	23570	22516	4.10	717.21
Total Mean = 22543 pixels, Total Error % = 3.99, Total variability (Std)=796.61									

Table 3.2: Results of user study for segmentation task using the computer mouse to select the seed points

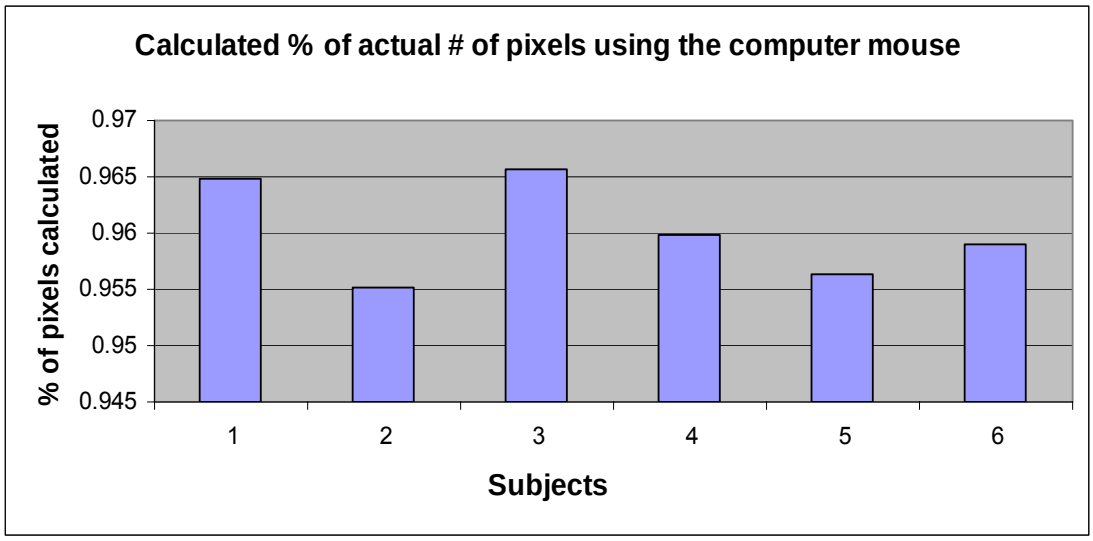


Figure 3.5: Percent of actual # of pixels calculated using the computer mouse to select the seed points

It seems that our proposed segmentation method is practical but there are certainly some flaws and limitations with the proposed method which will be more pronounced in

real applications of the method. We will discuss these problems in the following sections and propose some potential solutions.

In order to evaluate our design, at the end of the user study session, subjects were asked to fill out a short questionnaire. Bellow is the list of questions:

Q1) How do you rate the difficulty of the first trial of selecting the points using the surgical tool from 1-5? (1 is the easiest and 5 is the most difficult)

Q2) Do you think the segmentation task becomes easier for the next trials?

Q3) How do you rate the overall difficulty of the task from 1-5? (1 is the easiest and 5 is the most difficult)

Q4) Can you think of any particular complication in selecting points with this method?

Q5) Can you suggest a method which might make performing the task easier?

Q6) Do you think that having a wireless switch on the tool handle instead of clicking with a mouse can make the task less difficult?

From participants' answers to the questionnaire, the difficulty of the first trial was rated at 2.75 on average, the overall difficulty was rated at 2 and they all agreed that the task becomes easier after the first trial. Two of the participants thought that having the seed points distributed evenly makes the task somewhat complicated. Three subjects were concerned about having both hands required for the segmentation task and all the participants believed that having a switch on the surgical tool handle will definitely help making the task easier.

3.3.4 Discussing the results of the user study

The results of the user study showed that although the proposed segmentation method works relatively well in extracting the boundaries of the organ from specular

surfaces, some improvements need to be made since the accuracy of the extracted contours are highly dependent on the selection of the original seed points. In the above study, subjects were instructed to select certain numbers of seed points. That certain number of seed points had already resulted in a relatively good accuracy in the pilot study. What's more, they were instructed to select the points at equal intervals, which have also added to the accuracy of the segmentation. However, this might be a major issue for the surgeons and cause more fatigue and exhaustion during the surgery. Another important fact that we would need to take into consideration is that in actual surgery there are always chances of soft tissue deformation.

Now the question is what happens if the users do not select the seed points properly? Is there any way that they can compensate for their mistakes? How can they manipulate the extracted contour if the patient moves or the internal organ deforms during the surgery? Does the user have to repeat the procedure over and over again until they finally extract the best fitted contour? Although there is no doubt that the users (surgeons) are required to be trained before they can actually use this system in a real surgery, there is always a possibility of making mistakes, especially in stressful situations when time is a critical issue. Moreover, it is always possible that internal organs deform during the surgery or the patient changes position.

The other issue that could help us to improve our design is to consider the convenience of the user in our future design. It seems that all the participants prefer to perform the task with one hand rather than two hands. It might also be easier for the surgeons to perform the segmentation task unimanually and not to lose the contact with

the second surgical tool because every time that the surgical tool leaves the field of view of the camera, it will take some time to bring it back to the field of view.

In the following sections, we propose two new systems to augment the proposed segmentation method. In the first system, which is explained in section 3.4, we replace the computer mouse with a wireless switch and conduct another user study to compare the results of segmentation between the computer mouse and the wireless switch. Then, in section 3.5, we introduce an interactive system which could eventually allow the user to use surgical tools like a 3D input device, modify the extracted contours and fit them closer to the real boundaries of the organ.

3.4 Adding the wireless switch to the experimental set up

In this section, we introduce a small wireless switch to our experimental set up which could be used as an alternative to the computer mouse. Placing the wireless switch on the surgical tool handle can allow surgeons to use tracked surgical tools like a 3D input device. Results from the user study, discussed in section 3.3.3, suggested that replacing the computer mouse with a wireless switch which is located on the tool handle could make selecting the seed points in the segmentation task easier. In chapter 4, we will discuss some other potential applications of the wireless switch on surgical tool handles. Figure 3.6 shows how the small Zigbee wireless switch is located on the surgical tool's handle which could be easily pressed by the surgeon's finger when required. Using the wireless switch eliminates the need for other input devices such as key board or mouse. Moreover, the wireless switch in our experimental setup is inexpensive and can be used as a disposable unit.

Figure 3.6: The wireless switch is attached to the handle of the surgical tool

3.4.1 Wireless switch hardware and software

The wireless switch hardware consists of two PAN802154HAR00 transceiver modules. One is connected to the frame grabber and is called the robot module and the other one is connected to the small switch on the surgical tool handle and is called the master module. Figures 3.7 and 3.8 show how the wireless switch has been set up for this project. For more information about the modules and the wireless switch, please refer to appendix A at the end of this thesis.

These modules support the full Zigbee stack; however, in this case, the software used is called Simplified MAC or SMAC. SMAC implements neither the full ZigBee stack nor the complete 802.15.4 layer, but it is a simple and easy to use protocol. SMAC is basically a software that is installed on the MCU, and acts like a driver for the

MC13193 wireless chip on the module. For more information on SMAC sample code, please refer to appendix B at the end of this thesis.

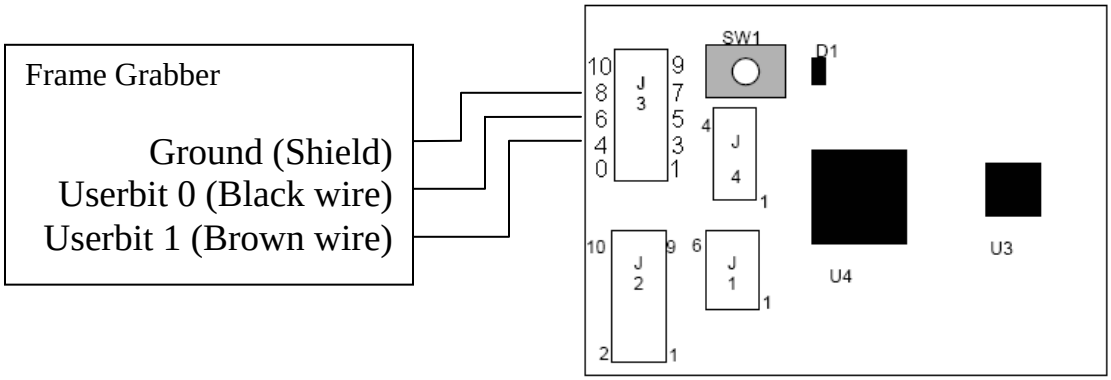


Figure 3.7: The robot module is connected to the frame grabber

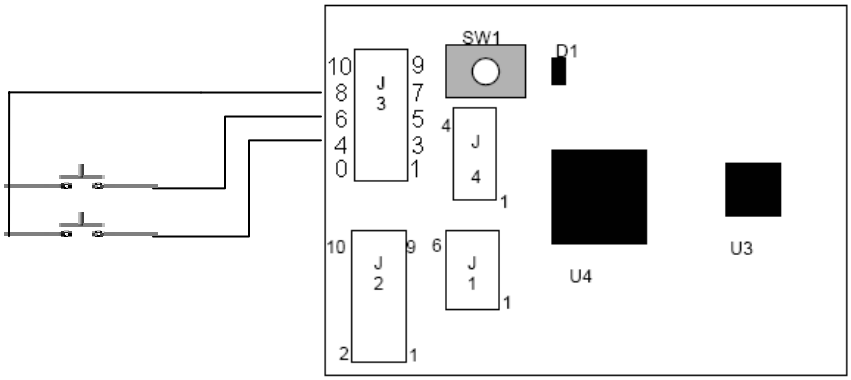


Figure 3.8: The master module is connected to the switches on the surgical tools

Using the functions available in SMAC, two simple programs have been written; one for the module connected to the surgical tool and one for the module connected to the frame grabber. The program on the module that is connected to the surgical tool, periodically (60 times/second) reads the status of the switches and sends a packet which contains a byte of data that has the status of the switch encoded in it. The program on the

module that is connected to the frame grabber receives that packet, extracts the byte of data then decodes it to get the status of the switches, and sets its output pins accordingly. So, if Switch 1 is depressed it sets pin 6 on J3 to high, and if Switch 2 is depressed it sets pin 4 on J3 to high. They remain unchanged until the next packet arrives.

3.4.2 Comparing the wireless switch with a computer mouse

After adding the wireless switch to our experimental setup, we conducted another user study on the segmentation task. A new group of participants was recruited for this study. The experimental set up and segmentation procedure was the same as the first user study described in section 3.3.3. The only difference was that this time, participants confirmed the location of the seed points by pressing the wireless switch located on the tool handle instead of clicking with the mouse button. Table 3.3 and figure 3.9 show the results of the segmentation process using the wireless switch to select the seed points.

Subject	**Calculated # of pixels for 6 trials**						**Mean**	**Error%**	**Std**
S1	23129	22093	22624	23264	22309	22560	22663	3.47	456.22
S2	22875	22109	23091	22323	23071	22548	22670	3.45	407.45
S3	22152	22963	23129	22117	22861	22456	22613	3.69	432.08
S4	23113	22564	22672	22962	22008	22678	22666	3.46	382.35
S5	22290	22671	23173	22383	22549	23541	22768	3.03	489.01
S6	22156	22078	22981	22763	22129	22348	22409	4.56	376.35
Total Mean = 22631 pixels, Total Error % = 3.61, Total variability (Std)=409.57									

Table 3.3: Results of user study for segmentation task using the wireless switch to select the points

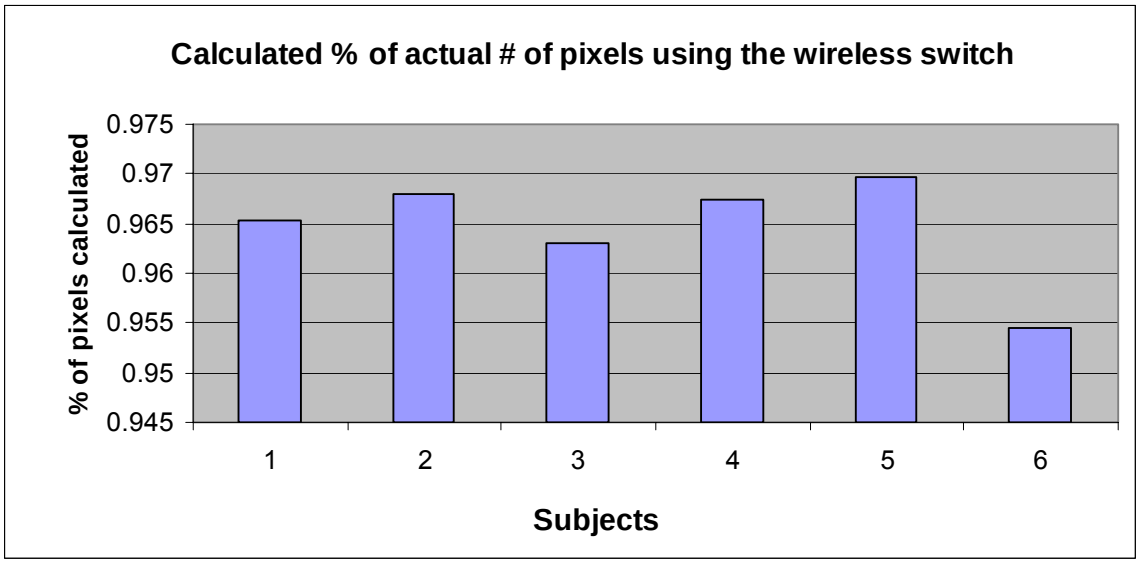


Figure 3.9: Percent of actual # of pixels calculated using the wireless switch to select the points

Comparing the results of this user study with the ones explained in section 3.3.3, it appears that there is a slight improvement in the subjects' performances. The average error has been reduced from 3.99 to 3.61. As well, the variability of performing the segmentation task has been decreased from 796.61 to 409.57. These results suggest that in addition to ease of use, replacing the computer mouse with the wireless switch can not only increase the accuracy of final results but also increase the consistency of performing the segmentation task.

It seems that the wireless switch attached to the surgical tool handle meets not only the visibility principle of a good design [27] but also acts as a good conceptual model for users since it predicts the effect of their action. Users can easily press the wireless switch instead of clicking on a computer mouse to confirm their actions. Therefore, it appears that there is a good mapping between the users' control and the result of their actions in real world.

3.5 Introducing the cube overlay system and its applications

The cube overlay system is developed to simulate a potential system which superimposes a predefined geometrical shape on top of a physical object, this can be used for various image guided surgical diagnostics and procedures. It can also be used as a 3D marker or graphical entity to guide the surgeon or be used to access and display information. This design could eventually allow users to stretch, shrink, scale and deform the extracted contours and compensate for their inaccuracies during the seed point selection stage. To simplify the calculations, we assume that the extracted contour is surrounded by a 3D cube. The cube overlay system could have many possible applications in the field of MIS and image registration.

In addition to manipulating the extracted contours, the cube overlay program could be considered as a potential system to assist with registering pre-operative scans with live endoscope images. Furthermore, it could potentially introduce a solution for some of the vision related problems with MIS, including lack of depth and colour contrast. This system could potentially allow surgeons to superimpose pre-operative 3D scans of a patient's body on top of live images of the endoscope camera to enhance the visualization during MIS. Pre-operative scans and live endoscope images of the same organ are not necessarily taken form the same viewing perspective. Moreover, the organ usually deforms during the surgical procedure and it is difficult to relate pre-operative scans with live endoscope images of the same organ. Assuming that the 3D pre-operative scans are surrounded by the virtual cube, by manipulating the cube, the system would allow surgeons to translate, rotate, deform and resize the 3D images before aligning them with the live images. The cube

overlay system could also be used as a system to measure the depth or the size of an internal organ.

3.5.1 Cube overlay program

The cube overlay program is written in C++ using the Matrox Imaging Library (MIL08). It uses MIL graphics module to draw a cube in the overlay buffer which is superimposed on the live images from the endoscope camera. As shown in figure 3.10, the program allows the user to grab one corner of the virtual cube with the surgical tool and translate, drag and resize or rotate it and ultimately superimpose it on a physical object while projection of the cube is preserved. The cube is originally drawn centred at the centre of the screen (x=0, y=0) and 100mm away from the camera (z=100). The original cube side length is set to 100mm with zero rotation around each axis. The centre of the cube and its 8 corners (indices 0-7) are identified with their live coordinates being calculated in the camera reference frame.

When the surgical tool gets close to the vicinity of the center or one of the corners of the cube, the word “Here” will appear on the selected corner or center beside its index number. This means that the surgical tool is close enough to that specific corner or center therefore the cube could be grabbed by the surgical tool by pressing the wireless switch which is located on the surgical tool handle. Pressing and holding the switch basically transfers the coordinates of the tip of the surgical tool to that particular location of the cube. From this moment, the live 3D coordinates of the tip of the surgical tool are shifted to the grabbed corner of the cube, and the 3D coordinates of 7 other corners of the cube will change accordingly.

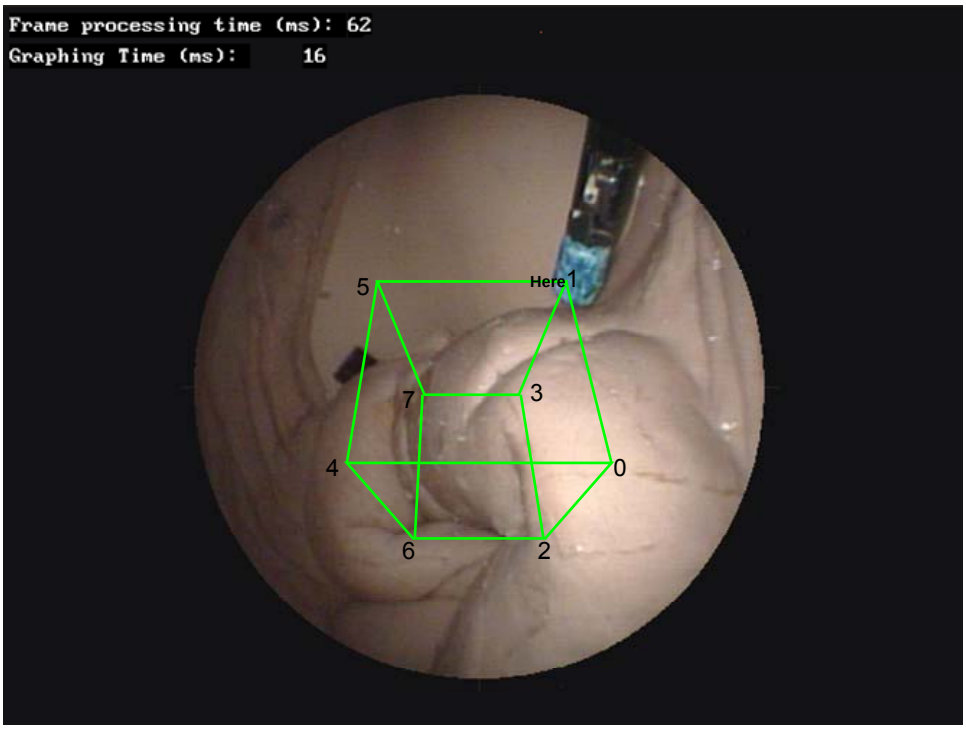


Figure 3.10: Corner of the cube is grabbed and manipulated by a surgical tool

Considering the rigid body transformation, if we know the 3D coordinates of the grabbed corner, which is the same as the 3D coordinates of the tip of the surgical tool, we can calculate the 3D coordinates of the 7 other corners. When the cube manipulation is no longer required, the user can release the cube by releasing the wireless switch.

3.5.2 Translation and rotation of the virtual cube using surgical tools

In order to have a better control of cube manipulation, the cube movement is broken down into 4 sub-movements namely translation and rotation around the x, y and z axes in the camera reference frame. The centre of the cube is assigned for its translation along x,y and z axis. Corners 0-2 are assigned for rotation around the x,y and z axes

respectively. In addition to rotation, the cube could also be dragged toward or away from its centre and resized when grabbed with these corners.

Suppose we have a cube with a side length of *a* and centered at *C*. Figure 3.11 shows the new location of the cube after being gabbed by its *corner #2* and dragged and rotated for θ degrees around Z axis. In order to apply changes to the original cube, there are three factors that need to be calculated: new side length (if the cube is dragged inward or out ward), direction of rotation and angle of rotation.

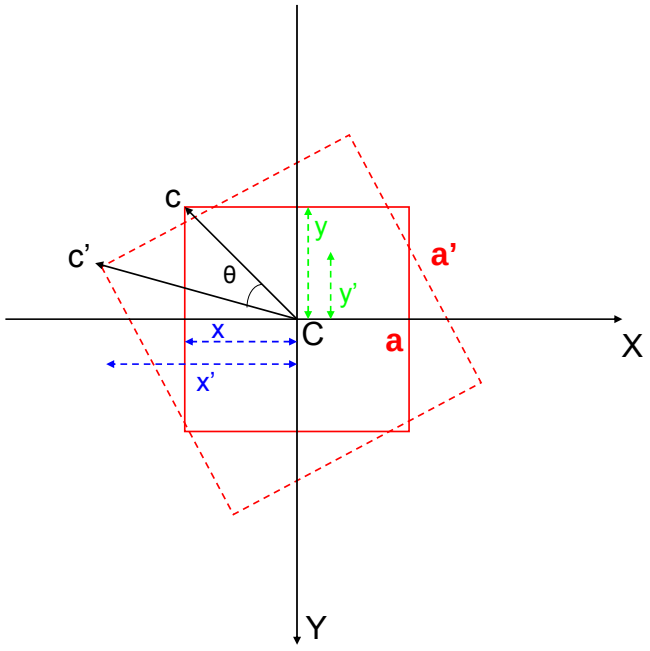


Figure 3.11: Cube rotation around z axis

To calculate the new side length of the cube from figure 3.11, we can write:

$$Cc = \sqrt{x^2 + y^2} = \frac{\sqrt{2}}{2}a \qquad \text{and} \qquad Cc' = \sqrt{x'^2 + y'^2} = \frac{\sqrt{2}}{2}a'$$

Therefore the new side length can be calculated based on the original side length, original and the new coordinates of the grabbed corner of the cube as follows:

$$a' = \frac{Cc'}{Cc} a = \frac{\sqrt{x'^2 + y'^2}}{\sqrt{x^2 + y^2}} a \, .$$

The above equation shows that if the cube is not dragged, then $Cc' = Cc$ and the side length remains the same.

The direction of the rotation angle which could be clockwise or counter clockwise is determined using the external multiplication of vectors. For example, to find the direction of rotation around the z axis, we have:

$$RD = (Cc' \times Cc)_z = xy' - yx'$$

If $RD < 0$, then the direction of rotation is clockwise and if $RD \geq 0$, the direction of rotation is counter clockwise.

The *cosine* of the angle of rotation is calculated using the internal multiplication of vectors. For example, for rotation around the z axis, we have:

$$Cc'.Cc = |Cc'||Cc| \cos\theta \Rightarrow \cos\theta = \frac{xx' + yy'}{\sqrt{x^2 + y^2} \times \sqrt{x'^2 + y'^2}}$$

Here is a sample code that shows how rotation and resizing of the cube around the z axis are calculated:

```
/***********rotation about z*****************/

if(cornerIndex == 2)
{
        centerToCorner = sqrt((corner[cornerIndex].x*corner[cornerIndex].x)      +
        (corner[cornerIndex].y*corner[cornerIndex].y));
        centerToNewCorner = sqrt((newCoor.x*newCoor.x) + (newCoor.y*newCoor.y));
        reSizeFactor   = centerToNewCorner/centerToCorner;
        newSideLength = reSizeFactor*sideLength;

    double directionVariable_z, rotationDirection_z;
    //External Vector multiplication
```

```
    directionVariable_z = corner[cornerIndex].x * newCoor.y - corner[cornerIndex].y *
    newCoor.x;

    if (directionVariable_z < 0)                    //CW
    {    rotationDirection_z = -1;}

    else if (directionVariable_z >= 0)              //CCW
    {    rotationDirection_z = 1;}

    //Internal Vector Mulplication
    double cosOfAngle_z = (corner[cornerIndex].x * newCoor.x
    +corner[cornerIndex].y*newCoor.y )
    /(sqrt(corner[cornerIndex].y*corner[cornerIndex].y +
    corner[cornerIndex].x*corner[cornerIndex].x) *sqrt(newCoor.y*newCoor.y +
    newCoor.x*newCoor.x));

    if(cosOfAngle_z>1)              //preventig errors
    {    cosOfAngle_z = 1;}
    double      zRotation = acos(cosOfAngle_z)*rotationDirection_z;

    rotateZ(zRotation*(180/Pi));
}
```

3.5.3 The perspective view of the cube

In order to have a natural way in which objects appear to the eye, that is a perspective view, we preserved this type of view of the cube in the cube overlay system. To have a natural view of the cube with an increase or a decrease in depth, we have applied the basis of 3D to 2D projection in our program. This means that the cube becomes smaller as it gets farther and it becomes larger as it gets closer to the camera. The user's point of view (camera's point of view) is represented by a point in the 3D space. From this point, the user has a certain view in the 3D world. This view is drawn below in figure 3.12 as a pyramid. All light rays that pass this pyramid will originate from objects that the user sees. When they pass this pyramid, they also pass the front plane. In the end, the whole 3D-2D projection will be on the plane in front of the pyramid.

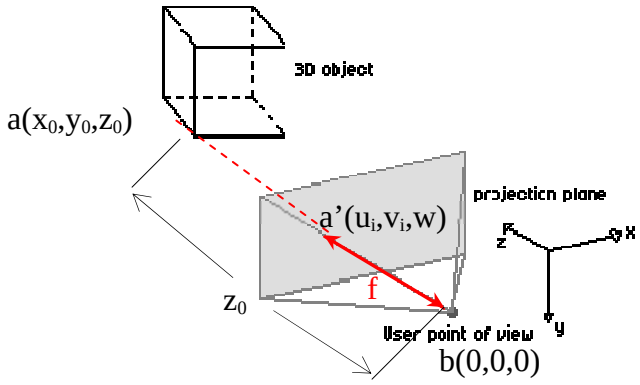


Figure 3.12: The basis of 3D to 2D projection

If we call the image plane *W* and show the lines that pass through the image plane with *[ab],* the intersections of *W* and *[ab]* give the coordinates of points in the image. We assign a fixed value to the *W* plane, then the coordinate of each point in the image could be written as (u_i, v_i, w). Using the simple geometry, if we know that the distance from the camera to the image plane is given by *f*, which is the focal length of the camera, the image coordinates will be related to the object coordinates using these equations:

$$u_i = \frac{fx_0}{z_0} \text{ and } v_i = \frac{fy_0}{z_0}$$

Since the focal length of the endoscope camera has previously been calculated; therefore the coordinates of the image could be calculated from the above equations.

3.5.4 Testing the accuracy of the cube overlay system

One of the potential applications of the cube overlay system is to measure the physical size of objects that appear in the field of view of the camera for diagnostic

purposes. In order to test the accuracy of the cube overlay system, we ran a test to measure the length of an object using the overlaid cube in both the x and y directions.

As shown in figure 3.13, a ruler was placed in the field of view of the endoscope camera at different depths from 50-100 mm. The length of each side of the cube was set to 30 mm. Using the tracked surgical tool the cube was moved and the farther side of the cube was brought to the same depth as the ruler and overlaid on it.

Figure 3.13: The virtual cube is overlaid on the ruler to measure 30mm of its length

Table 3.4 shows the discrepancy between the 30 mm of the ruler and the side of the cube in X and Y direction and at different depths.

Depth(mm)	Measured length(mm), x direction	Measured length(mm), y direction
50	30	30
60	30	30
70	30	30
80	30	30
90	29.5	29.5
100	29.5	29.5

Table 3.4: Accuracy of the cube in measuring 30mm of length in both the x and y directions

These results show that the cube program is very accurate in measuring the length of an object. Therefore, it could be used as a calibration block to measure the size of internal organs. In order to extend the use of the cube overlay system to manipulate the extracted boundaries or to overlay the intra-operative scans on top of live images of the endoscope, it has to be very accurate and reliable. The result of this test indicates that the cube overlay program can potentially be extended for the mentioned purposes.

Chapter 4

Interactive systems to access medical records during the MIS

Current limitations of Minimally Invasive Surgeries (MIS) such as indirect and limited field of view make this type of surgery a very complicated procedure. In general, it would be desirable for surgeons to have access to patient's medical information during the surgery without taking their hands off the surgical tools, or call an assistant using voice commands. Through image analysis of the state of surgical tools during the surgery, we introduce three intuitive systems which could potentially enhance the efficiency of HCI in the area of MIS.

Our proposed systems use the gesture and position of surgical tools as input media which then enable surgeons to have access to pre-operative medical images through design of graphical overlays. This approach can offer a new paradigm for developing a convenient and natural human-machine interface for this class of surgery.

In section 4.1, we introduce our first interactive system to access a patient's medical record during a MIS. Forming a specific gesture with the surgical tools switches the system to its augmented mode and enables access to a patient's medical record. Section 4.2 describes our second interactive system which allows surgeons to have access patient's pre operative medical records via displaying a ring menu around the field of view of the camera during the surgery. In section 4.3, we explain our third interactive system which has a similar design to the ring menu described in section 4.2. However, the menu activation is accomplished by means of the wireless switch attached to the

surgical tool handle. After conducting user studies with the proposed systems, the advantages and disadvantages of each proposed system are discussed in section 4.4 .

4.1 Augmented surgical environment using gesture recognition function

The idea of gesture recognition of surgical tools was originally proposed by Hsu & Payandeh [21]. This entails recognizing different gestures formed by surgical tools such as graspers or scissors. The objective of the surgical tool gesture recognition function is to recognize tool gestures formed with the long-stem surgical tools and allow surgeons to send commands to the system. A surgical tool gesture recognition function in MIS could provide a convenient and intuitive human computer interface for surgeons in order to retrieve patients' medical information. In the actual surgery, it may not be always possible for surgeons to call an assistant for help or to free one hand and manually press a key on the computer to access patients' pre-operative medical scans. In this section, we propose a system which allows surgeons to obtain patients' pre-operative scans upon forming a particular tool gesture.

4.1.1 Gesture recognition system overview

As shown in figure 4.1, the gesture performed by the surgical tools is recognized in three main steps [21]. In the first step, tool blobs are segmented with morphological filters and thresholding. The simple and efficient nature of mathematical morphology could keep the processing time low, which is important in the time-critical surgery environment. With efficiency and processing time in mind, our selected methods for

target tool segmentation are thresholding and mathematical morphology. In the second step, feature quantities are extracted from tool blobs using blob analysis. These feature quantities constitute distinctive information extracted from the target tool blobs that together allow the surgical tool gesture classifier to classify the surgical tool gestures. The feature quantities are then fed to a neural network for gesture classification and finally, the surgical tool gesture is recognized.

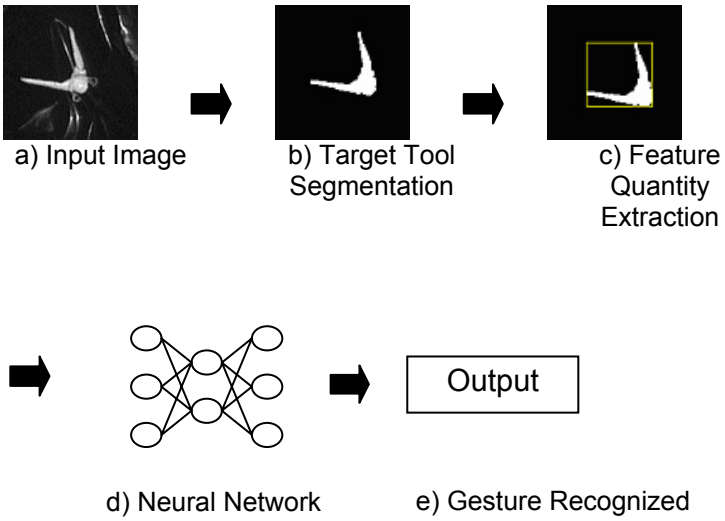


Figure 4.1: Major steps toward gesture recognition of surgical tools

4.1.2 Selection of appropriate gestures

Any time during the surgery when surgeons require more information regarding patient's medical record, they can easily switch to the augmented mode by forming a pre-defined gesture using the surgical tools already in hand. We assume that there would be enough room inside the patient's body to perform a tool gesture without damaging the surrounding tissue.

We have chosen graspers and scissors; the two most commonly used multi-state surgical instruments to form the tool gestures in our experiment. Figure 4.2 shows eleven different gestures formed by graspers and scissors [21].

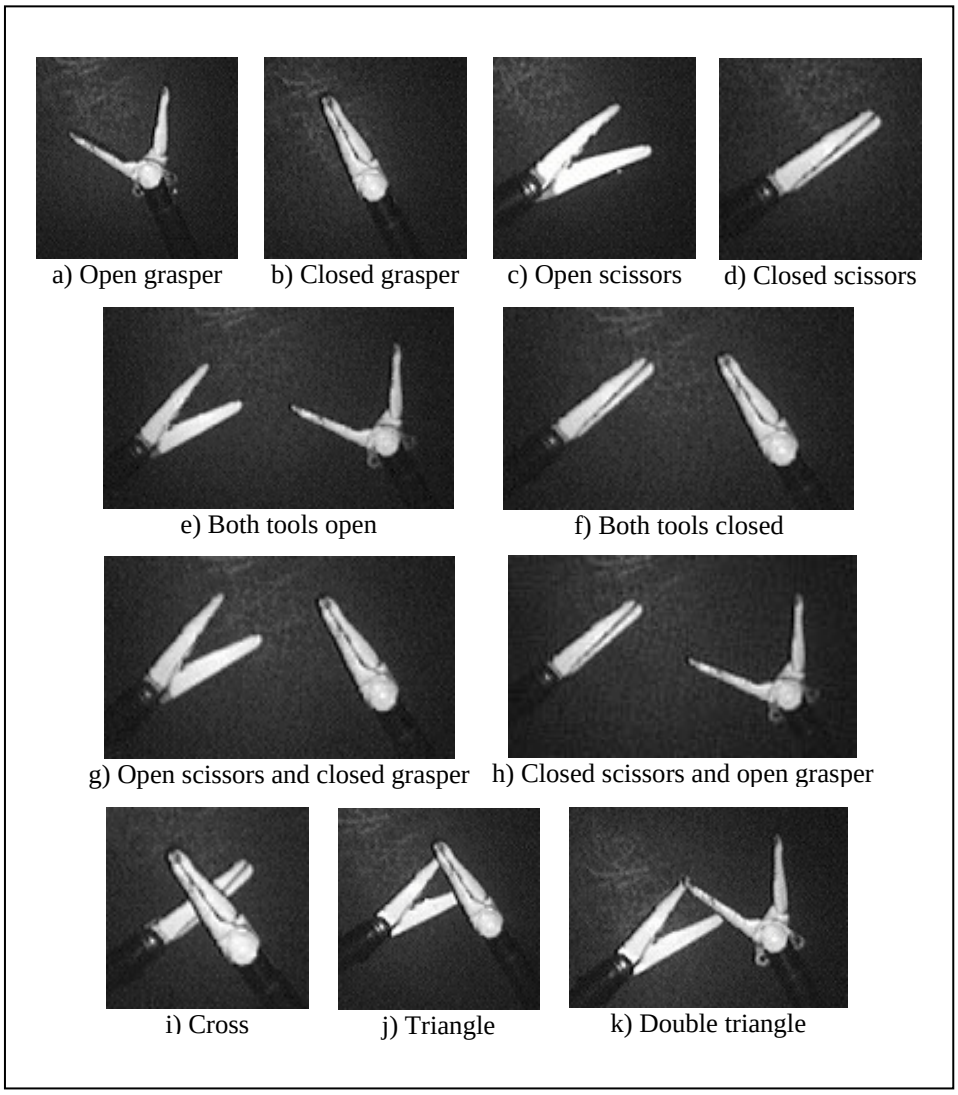


Figure 4.2: Eleven different gestures formed by graspers and scissors

In order to eliminate accidental activation of the system's augmented mode, among all the gestures formed with surgical tools, we have selected one of the tool gestures which is rarely formed during the surgery to activate the augmented mode. For

example, choosing an open/close graspers or scissors for this purpose would not be practical because these gestures are repeatedly formed during surgical procedures. In reality, not all eleven tool gestures would have satisfactory recognition rates, and not all eleven tool gestures would be needed. Since with the current gesture recognition system not all the tool gestures could be recognized at a high rate, therefore, in addition to the uniqueness of the selected gesture, it should also have a relatively high recognition rate by the system. Table 4.1 summarizes the maximum recognition rate for different surgical tool gestures shown in a previous study [28] .

Formed Gesture	**%Maximum recognition rate**
Open graspers	100
Closed graspers	100
Open scissors	90
Closed scissors	100
Cross	80
Triangle	50
Double triangle	0

Table 4.1: Maximum recognition rate for different tool gestures

From the above table, we can see that the maximum recognition rate for the "cross" gesture is about 80 percent. Considering both uniqueness and a relatively higher recognition rate, this tool gesture is selected to activate the system's augmented mode of in our research.

4.1.3 Switching to augmented mode via gesture recognition function

The proposed augmented system is designed in a way that any time during the surgery where surgeons require to look at patients' pre-operative scans, they can switch to the augmented mode of the system by forming a cross gesture in the middle of the endoscope's field of view using surgical graspers and scissors. To avoid accidental activation of the augmented mode, the formed gesture needs to be held for 1 second to be recognized by the system.

As shown in Figure 4.3, upon gesture recognition, the field of view of the endoscope, except the middle section, is going to be divided into 4 regions; open X-ray, open MRI, open CT and Exit. These annotations will be overlaid on the endoscope's field of view. Surgeons can easily choose a specific function by entering the tip of the graspers to the corresponding region.

For this experiment we have used two side by side monitors. The first monitor shows the field of view of the endoscope camera and the second monitor is assigned to display supplementary information. Therefore, selected scans will be opened in the second monitor without cluttering the view of the surgical site.

To exit the interface, a surgeon only needs to select the "Exit" function by brining the tip of the surgical graspers to the "Exit" region of the field of view.

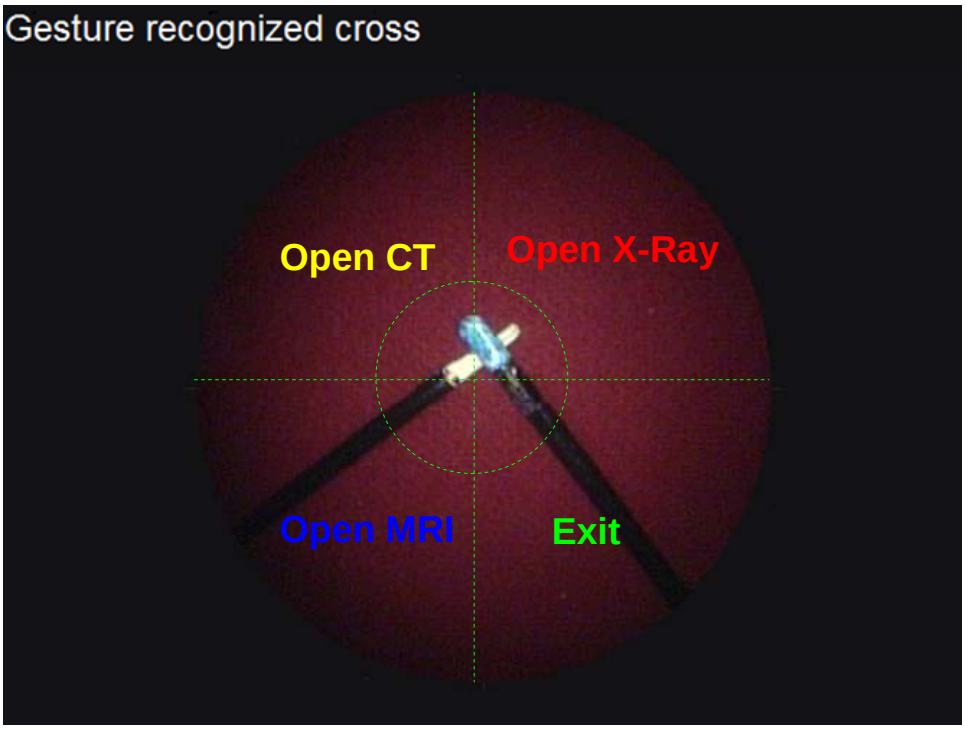


Figure 4.3: Cross gesture is formed and recognized, system switches to augmented mode

4.2 Augmented surgical environment via the ring menu

In this section, we propose an augmented system for accessing the information regarding patients' medical record which is accessible at all the times during the surgery. This is achieved via displaying a narrow ring around the field of view of the endoscope, therefore surgeons work in an augmented environment from the beginning of the surgery. This ring basically acts as a menu which allows surgeons to choose a specific function and have access to a particular medical record of the patient. Considering the point of origin at the center of endoscope's field of view, as shown in figure 4.4, we superimpose a narrow ring around the field of view of the endoscope which is then divided into four (expandable to more) different regions corresponding to different functions, namely open

MRI, open CT, open X-Ray and Exit. When the selected scan is no longer required, surgeon can close that image by simply choosing the exit function from the ring menu.

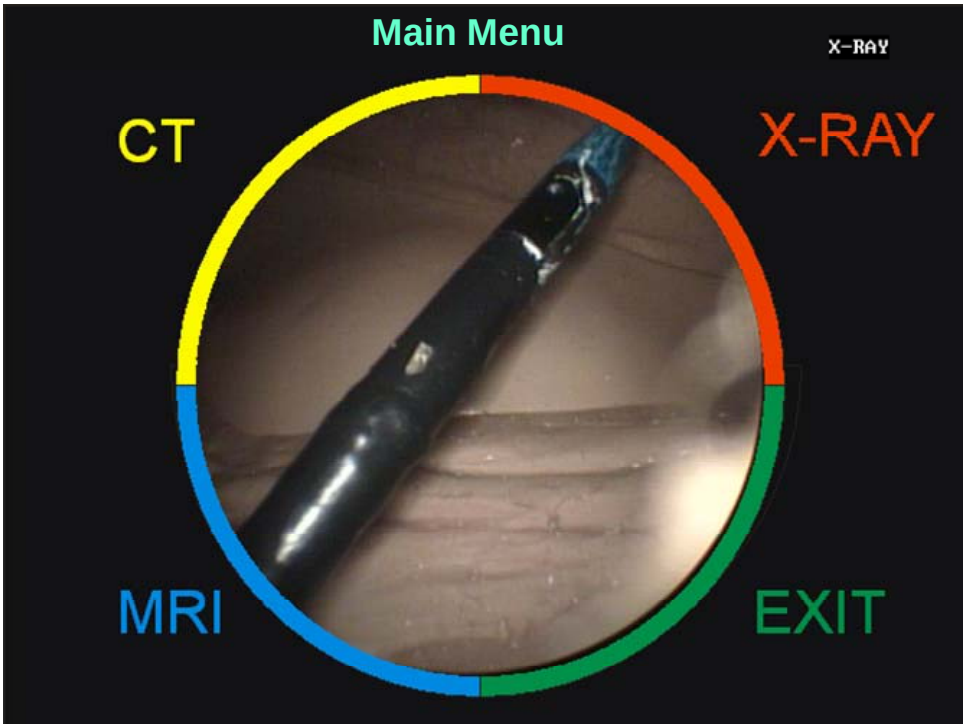


Figure 4.4: The ring menu overlaid around the filed of view of the endoscope

The proposed system allows surgeons to access patients' medical records without losing the contact with the surgical tools. It is designed such that at any time during the surgery, if surgeons require to look at pre-operative MRI, CT or X-ray scans of the patient, all they need to do is select the particular function by touching the corresponding region of the menu with the tip of the surgical tool and hold it there for one second. This delay reduces accidental activation of the scans. As before, two side by side monitors enable opening the preoperative scans in the second monitor, which reduces cluttering the surgeon's field of view.

After activating a particular function, the ring menu changes its contents as shown in figure 4.5 and allows surgeons to choose a subtask of the selected function. For example, if a surgeon activates the X-Ray region from the main ring menu, the subtask ring menu allows the surgeon to choose a previous image, next image or go back to the main menu.

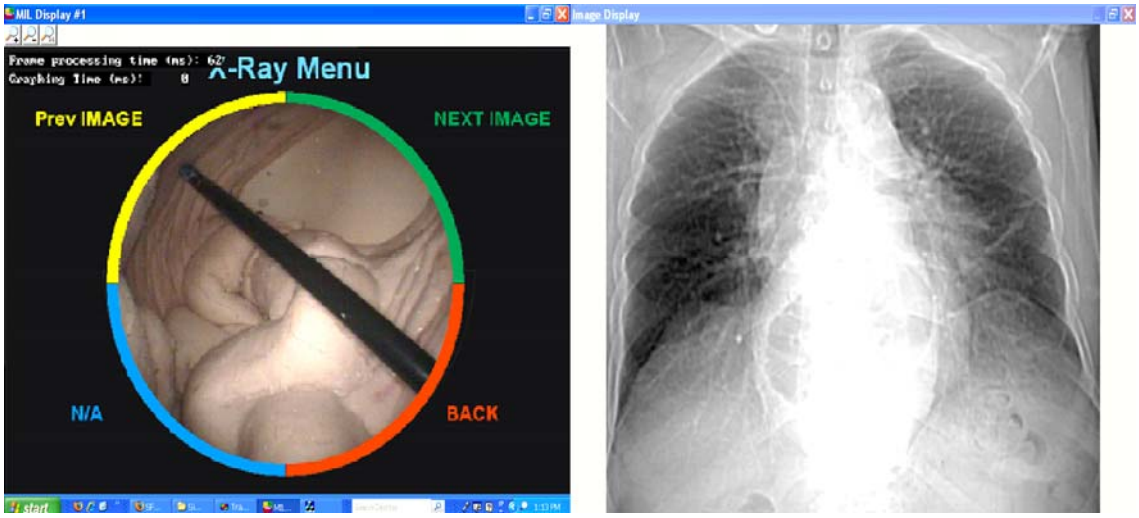


Figure 4.5: In X-Ray menu, user can choose previous or next image or go back to the main menu

4.3 Augmented surgical environment via the ring menu (with wireless switch)

In section 3.4, we added a small wireless switch to our experimental setup. This wireless switch which is attached to the surgical tool handle eliminates the use of any other input devices such as a keyboard or a mouse and could eventually reduce time, voice commands and the number of assistants needed during the surgery. In this section, we propose our third augmented system to access patients' medical record during the MIS. This system is designed similarly to the ring menu, which was

discussed in the previous section. The only difference is that unlike the ring menu mentioned in section 4.2, the scans will not become activated automatically as the surgeon touches a particular section of the menu with the tip of the surgical tool and hold it there for one second. Instead, in order to activate each scan, a surgeon has to press the wireless switch while touching a particular section of the ring with the tip of the tracked surgical tool.

4.4 Comparison of the three proposed systems

In order to evaluate the three proposed systems, user studies were conducted among university students. The participants were asked to activate the CT, MRI, X-Ray and Exit functions using all proposed systems. Before starting each trial, they were told which particular scan should be activated.

In the first set of trials, the participants were instructed to form the cross gesture with graspers and scissors from an arbitrary positions upon receiving the “Go” command from the experimenter. Then they activated a particular scan by pointing at its location with the tip of the gaspers, upon gesture recognition by the system. Table 4.2 shows the total time required form the onset of each trial until the gesture is recognized and a particular function is activated.

In the second sets of trials, subjects started each trial upon hearing the “Go” command and activated a particular function of the ring menu by bringing the tip of the graspers to that location and holding it for one second. Table 4.3 shows the total time from the beginning of each trial until the activation of a particular scan from the menu.

Subject	Time (s) to form the gesture and activate a function for 6 trials						Average
S1	8.43	7.32	9.03	9.67	7.56	8.71	8.45
S2	5.03	7.87	8.23	5.92	6.16	5.21	6.40
S3	8.56	7.24	8.42	7.65	6.97	7.39	7.71
S4	6.29	7.68	8.39	8.91	7.21	5.38	7.31
S5	9.73	8.81	7.62	9.23	7.81	7.39	8.43
S6	8.95	7.83	5.93	6.56	8.64	5.72	7.27

Table 4.2: Total time (s) to form the cross gesture and activate a particular function

Subject	Time (s) required to activate a function from the menu for 6 trials						Average
S1	3.16	4.12	3.05	2.78	2.98	3.27	3.23
S2	3.65	3.11	2.89	2.69	3.56	2.71	3.10
S3	4.29	3.9	3.54	3.43	3.92	3.32	3.73
S4	3.62	2.91	3.21	2.84	2.98	3.08	3.11
S5	2.79	3.26	3.43	3.82	3.15	3.22	3.28
S6	3.43	2.59	3.17	3.36	3.74	3.56	3.31

Table 4.3: Total time (s) to activate a function from the menu

In the third sets of trials, the participants started each trial upon hearing the “Go” command and activated a particular function of the ring menu by bringing the tip of the graspers to that location and pressing the wireless switch which was placed on the

graspers' handle. Table 4.4 shows the total time from the onset of the trial until the activation of the particular function from the menu by pressing the wireless switch.

Subject	Time (s) required to activate a function from the menu for 6 trials						Average
S1	2.14	2.89	1.97	2.65	2.76	2.37	2.46
S2	2.09	1.83	1.98	2.24	2.31	2.59	2.17
S3	3.21	3.34	3.52	4.28	3.45	4.31	3.69
S4	3.09	2.91	3.18	3.28	2.54	2.34	2.89
S5	2.88	2.5	3.37	3.18	2.68	2.96	2.93
S6	3.29	3.64	3.43	2.36	2.82	3.37	3.15

Table 4.4: Total time (s) to activate a function of the menu with the wireless switch

The result from the user study shows that on average it takes about 7.6 seconds for the participants to activate a particular scan by performing the cross gesture with graspers and scissors. It takes about 3.29 seconds to activate each scan by touching a section of the menu and hold the tool for one second and only takes about 2.88 seconds to activate any scan from the ring menu by pressing the wireless switch. Although it seems like performing a particular gesture might take more time comparing to using the ring menu, it is important to remember that the participants in this study had no experience with remote manipulation of long stem surgical tools. Our consultation with the surgeons at Surrey Memorial hospital indicated that forming particular gestures with surgical tools are very simple tasks from their point of view; therefore, we believe that a gesture

recognition function could have potential useful applications in the design of interactive systems for minimally invasive surgeries.

As one would expect, there are advantages and disadvantages for each of the described interfaces in real applications. The major advantage of the gesture recognition function is that surgeons are able to perform surgical tasks without being distracted by any annotations on their field of view of the surgical site given that they can easily switch to augmented mode when required. The other advantage of this system is that it does not require using any external piece of hardware. However, performing the proper gesture will take some time and the system might not recognize the gesture if it's not performed properly.

The main advantage of the augmented ring menu, both with and without the wireless switch is the fact that a patient's medical record could be obtained more easily and faster during the surgery and there is no need to form a specific gesture which has to be recognized by the system. The ring menu is always available around the filed of view; therefore, the surgery is basically performed in an augmented environment rather than switching to the augmented mode when more information is required. A potential problem with the ring menu is that even though it is designed to be very narrow, some users might still find it distracting before getting used to it.

Surgical tools repeatedly enter and leave the field of view of the endoscope during the actual surgical procedures. This means that surgical tools may be changed during the surgery. In order to reduce unwanted activation of scans with the second proposed system, surgeons need to hold the surgical tool at the selected region of the menu for one second before the chosen scan appears on the second monitor. Because of this delay, it takes

more time to activate scans with the second system than the third system which uses the wireless switch instead if this delay time (3.29 vs 2.88 seconds). However, the main advantage of the second proposed system comparing to the third system is that it does not require installation of the wireless switch. If we eliminate this delay, the scans will become activated immediately upon touching the corresponding section of the menu with the surgical tools. Therefore, the average time required for scan activation will become about 2.29 seconds which is even faster than using the wireless switch but there will definitely be a trade off between this time delay time and unwanted activation of the scans. Using the wireless switch in the third proposed system eliminates the accidental activation of the scans during the surgery. Comparing the average time required to touch a section of the menu with the surgical tool (2.29 sec) and the average time required to activate each scan with the wireless switch (2.88 sec) shows that using the wireless switch could not only eliminate unwanted activation of the scans, but also the delay that it adds to the system is less than a second (about 0.6 sec).

Chapter 5

Discussion and future work

In the past decade, many studies have been conducted and specialized devices have been designed to help surgeons to overcome some of the existing problems with MIS. However, not all of these designs and proposed techniques would have applications for real surgical procedures. Unfortunately, if the new technique does not fit properly in the operating room and does not provide the ideal ergonomic environment for surgeons, it would not be useful in real applications. In the following section, we compare our approach to design a new paradigm for surgeon computer interface with some other type of HCI techniques.

In the following section, we discuss our proposed paradigm for surgeon computer interface for MIS. In section 5.2, we compare our proposed interface with some other types of HCI devices and finally in section 5.3, we suggest some methods to extend and improve the proposed designs in this thesis.

5.1 Discussion

It appears that one of the greater challenges to the successful implementation of image-guided surgery and MIS systems in the operating room of the future relates not to the technology, but to the user interface. One of the big disputes in the design of these systems is that the design should be simple to operate by the users. It should also minimize additional technical support and eliminate keyboard and inconvenient switching between mechanisms in the operating room [6].

In this research, we introduced a new paradigm for surgeon computer interface for MIS through image analysis of surgical tools. The potential advantage of our study is that surgeons can obtain more information regarding patients' medical records during the surgery just by using tracked surgical tools as input device which is a natural and convenient way of human machine interface for surgeons.

Our proposed tracking system which was discussed in chapter 2, estimates the 3D position of the surgical tools with an average accuracy of millimetres (see section 2.5). The accuracy of the tracking step is a key step in determining the precision of the segmentation method as well as the gesture recognition system and the proposed systems in chapter 4. This accuracy could be acceptable for many surgical applications such as laparoscopic surgery albeit some surgical procedures require greater accuracy. The image processing method that we have used for the tracking system is blob analysis which is typically performed on grey scale images. Here, we offered blob analysis of coloured images of the endoscope which gives us more control on selecting the desired blobs just by referring to their colours and not by calculating complex shape attributes. It also allows us to simply track multiple blobs if required.

In chapter 3, we introduced a new method for segmentation of live endoscope images during MIS using surgical tools to select the seed points. The result of the first user study showed that if the users follow the given instructions, such as selecting the points sequentially and evenly distributing them along the boundaries of an object, they could extract the contours of the object with an acceptable accuracy. However, the accuracy of this method is highly dependent on the location and number of seed points. This proposed method requires much improvements and it should allow users/surgeons to

manipulate the extracted contour and compensate for their inaccuracies in selecting the seed points. For the second user study, we replaced the computer mouse with a wireless switch which was attached to the surgical tool handle. This substitute made the seed point selection easier by the users; however, the problem of correcting the extracted contours still remains the same. Later in chapter 3, we introduced the cube overlay system as a potential solution to manipulate the extracted contours and fit them more closely to the actual boundaries of the organ. The developed program was successful in manipulating the cube such as allowing for translation, rotation in 3D directions, enlarging and shrinking. Nonetheless, more work needs to be done in order to manipulate the extracted contours.

In chapter 4, we devised three new interfaces to allow surgeons to have access to the patients' pre-operative scans during the MIS. The first system enables access to pre-operative scans by forming a particular gesture with surgical tools which is recognized by the system where the second and third systems allow access to medical records via a narrow ring around the field of view of the endoscope. However, in the second system, the scans get activated automatically when the user touches a particular region of the menu with the tip of the surgical tool whereas in the third system, the scan needs to be activated by pressing the wireless switch located on the surgical tool handle. We conducted user studies for all of the proposed systems, compared them with each other and listed potential advantages and disadvantages of each system. Although the users performed the task faster and easier with ring menu systems comparing to the gesture recognition system, we believe that all systems could have potential benefits for the surgeons during the actual MIS.

5.2 Comparison of our approach with other HCI techniques

Unlike most of the available tracking methods for MIS and HCI applications, neither our tracking system nor the two proposed interfaces to access medical records require an external piece of hardware which could interfere with the surgeon's performance during the MIS.

Commercially available optical tracking systems such as OPTOTRAK and POLARIS have been generally used in MIS and medical robotics researches. As shown in figure 5.1, these tracking systems can estimate the position of the markers which are attached to the instruments or to the human body with a great accuracy [7].

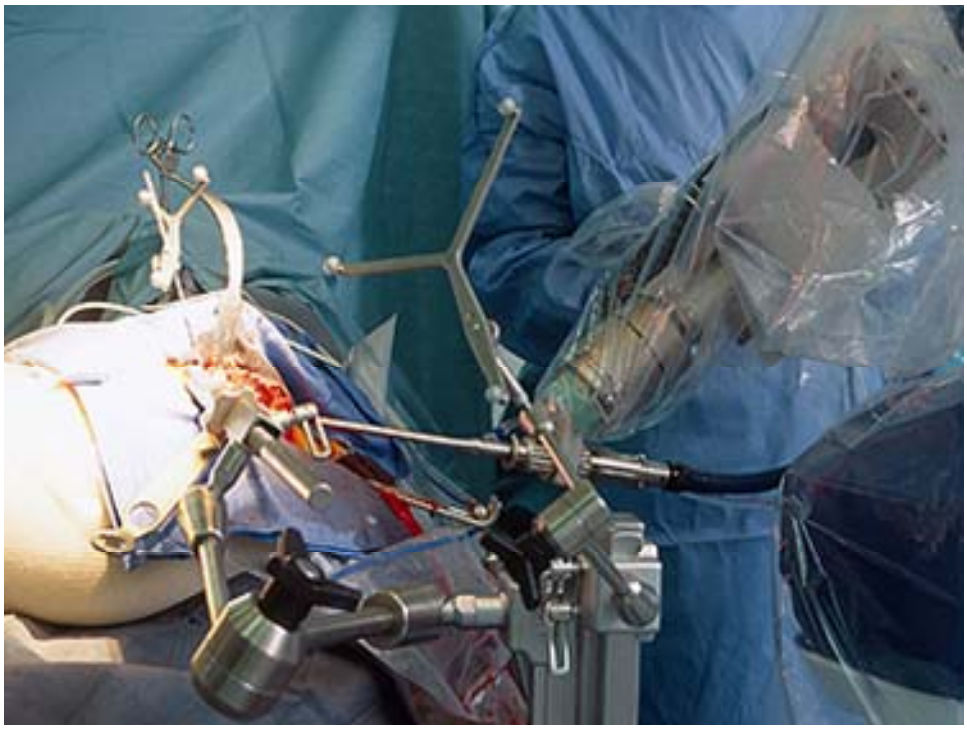


Figure 5.1: LED markers are attached to the instrument during a surgery

Figure 5.2 shows the application of an optical tracking system during an orthopaedic surgical procedure [7].

Figure 5.2: Surgical applications of optical tracking systems

In order to track the 3D position of the LED markers, they have to be present in the field of view of the tracking cameras. However, whether the LED markers used with these tracking systems are wired or wireless, they cannot be tracked by the tracking system if they get obscured during the surgery and they can not satisfactory track the surgical tools inside the patient's body. In addition to this, other possible states of the tools inside of the surgical cavity can not be identified by only tracking parts of the tool which is outside of the cavity. Such states can be the open-close configuration or tool contact with the tissue and organs. Our tracking system can easily track the position of the surgical tools inside the patient's body by processing the images of the endoscope. It

can also recognize the gesture formed by the surgical tool and use it as an interface which allows the surgeon to send commands to the system and have access to patients' medical records.

Employing real time imaging techniques such as MRI, ultrasound and fluoroscopy could solve the problem of tracking surgical tools inside the patient's body. Although intra-operative MRI can provide excellent images of anatomical structures, especially for brain, it also has its limitations. Not only it requires substantial investment for the scanner itself, but also due to the existence of a strong electro-magnetic field it requires equipping the operating room with MRI-compatible instruments. In addition to this, the physical layout of many interventional MRI systems compromises the surgeon's access to the operating field. Intra-operative ultrasound machines, which have little distortion, cost about 10% of a typical MRI machine [29]. In order to track the position of the ultrasound probe, a commercial tracking device such as a Polaris tracker is still required. Our two proposed systems can potentially enable access to patients' pre-operative scans without having an expensive imaging system during the surgical procedure. During the MIS procedures, where the tissue shift is negligible, our two proposed systems to access pre-operative scans would be beneficial to the surgeons. However, if the tissue shifts substantially during the surgery, it becomes difficult to relate the endoscope images to the pre-operative scans. Therefore, proper registration of pre-operative scans with live images of the endoscope will be crucial.

Wearable HCI devices such as data gloves, shown in figure 5.3, have been used for menu selection tasks in aircraft maintenance by Witt *et al.* [30].

Figure 5.3: Data gloves

These researchers conducted user studies for a menu navigation task in aircrafts using different interactive devices. The main focus of their experiment was the missing information about how aircraft maintainers feel and behave when using specific wearable user interfaces that are usually controlled by novel interaction devices beyond those known from desktop systems .The result of their user study showed that the best accuracy in navigation task was achieved with the mouse and in spite of the fact that a data glove was introduced as a novel system, its accuracy was comparable to the mouse and certainly superior to the speech method. The problem with wearable gloves for the menu selection for MIS is that these systems still require external hardware attached to the user's body, which might cause discomfort for the surgeons during the surgery. Unlike data gloves and head mounted devices, with our proposed systems, surgeons can have access to pre-operative scans without having an extra piece of hardware attached to their

body. The use of a small wireless switch on the surgical tool handle does not interfere with the surgeon's hand movement and it could only be pressed when required.

Eye gaze tracking systems have been used for many applications in navigation tasks. Eye gaze studies have an increasingly important role in examining the usability of human computer interfaces [16]. Although the current eye gaze tracking system could eliminate head mounted optics, the current systems only allow the user one square foot of head movement and hence greater accuracy is possible if subject is using a chin rest [18]. In general, MIS procedures have already imposed limited movement and awkward body postures for the surgeons, therefore the existing eye gaze tracking systems might not be the best solution to access patient's pre-operative scans during the MIS. Another reason that makes this type of tracking less suitable for the MIS is the possibility of accidental activation of the scans when it's not necessary. Eye gaze tracking systems might be more useful if they are used as interactive systems for radiology workstations, where the main focus of the user is navigating through medical scans.

In brief, the possible advantages of our approach comparing to some other HCI methods for MIS include:

- No need for expensive external hardware or software.
- Tracking the position of the surgical tools inside the patient's body efficiently.
- No restriction on the surgeon's movements.
- No additional device to be attached to the surgeon's body.
- Allowing for a high control on activation of pre-operative scans.

- Accessing patients' medical records without losing the physical contact with surgical tools.
- Utilizing familiar surgical tools.
- No need for extensive training.

5.3 Future work

This new paradigm for surgeon computer interface that we proposed in this research could offer a great benefit for MIS, image guided surgery, robotics and telesurgery. However, the proposed systems are at their preliminary stages of development and could be improved and extended in many ways. In the following sections we suggest some future studies in this area.

5.3.1 Using active contours to enhance the segmentation result

In this research, we used cubic splines to connect the selected seed points and extract the boundaries of the organ of interest. Cubic splines are simple and fast means for interpolation and approximation, but the accuracy of the extracted contour with this method is only dependent on the location of the selected seed points and it does not take into account any image characteristic. For future studies, we suggest using snakes or deformable models instead of cubic splines to connect the seed points and interactively manipulate the contours. These models take into account other image attributes to extract the contours of the object not just the location of the seed points.

Snakes or active contours, originally proposed by Kass *et al* [31], are powerful image segmentation techniques and are used extensively in computer vision and image

processing applications, particularly in order to locate object boundaries. Snake is an energy minimization technique that uses information both on the objects shape and its image properties to determine the segmentation of the object. The internal and external forces are defined so that the snake will conform to an object boundary or other desired features within an image. The basic behaviour of the active contour or snake is to minimize the energy sum across its length. The energy functional which represents the energy of the snake is given by:

$$E_{snake}v(s) = \int_0^1 (E_{\text{int}}v(s) + E_{ext}v(s))d_s$$

Where vector $v(s) = (x(s); y(s))$ is the position of the contour. The active contour's internal energy term (E_{int}) is the one that enforces the continuity of the contour by placing an energy weight on both the first and second order terms of the function $v(s)$. The internal energy of the snake is given by:

$$E_{\text{int}}v(s) = w_1(s).\left\|\frac{\partial v}{\partial s}\right\|^2 + w_2(s).\left\|\frac{\partial^2 v}{\partial s^2}\right\|ds$$

The first order term relates to the contour's tension or how much distance is allowed between its connected segments. The second order term controls the contours stiffness or rigidity. The weights $w_1(s)$ and $w_2(s)$ control the snake's tension and stiffness. The internal energy imposes internal shrinking and straightening forces on the snake, which make the snake act both as a spring and as a flexible rod. The external energy (E_{ext}) drives the snake towards the relief boundary and localizes the snake. This external energy is given by:

$$E_{ext}v(s) = -\left|G_\sigma(x,y) * \nabla I(x,y)\right|$$

Where $G_\sigma(x, y)$ is the Gaussian of the standard deviation σ and $\nabla I(x, y)$ is the gradient of the image. The external energy should have a minimum at the boundary such that the snake converges towards it.

Deformable models are capable of accommodating the significant variability of biological structures over time and across different individuals. Furthermore, they support highly intuitive interaction mechanisms that, when necessary, allow medical scientists and practitioners to bring their expertise to bear on the model-based image interpretation task [32].

5.3.2 Other criteria to validate the segmentation accuracy

The evaluation criterion used in this research is not accurate enough in order to validate the result, of the segmentation task. Although the numbers of calculated pixels inside the extracted contours were relatively close to the actual number of pixels occupied by the object, it does not necessarily means that the shape and the exact location of the contour is the same as the object. Object shape is one of the most important attributes of a physical object. The shape information is often essential to understanding scenes in a visual perception system. Extracting the 3D shape of an object from images is a main task in computer vision, and has been actively investigated in recent years, which has potential important applications such as navigation, surveillance, object recognition and virtual reality [33]. For future studies, we suggest using some shape matching techniques such as iterated closest point (ICP) [34] as well as quantitative analysis like pixel or voxel count to evaluate the segmentation results.

5.3.3 Validate the overall interface

The standard method for validating a new therapy/surgery is by evaluating its performance relative to standard practice. Almost always, a prospective clinical trial is necessary to validate a new approach [35].

Our user studies showed that our proposed segmentation method as well as the three proposed interfaces to access patients' pre-operative scans could be practical. However, all participants in our studies were university students with no background in MIS and no experience in working with long stem surgical tools. Moreover, the experimental setup used for this research is only a simple simulation of the actual MIS. In order to be able to claim the usability and practicality of the proposed methods in this research, they need to be tested by surgeons and in real surgical environment.

5.3.4 Extend the application of the cube overlay system for registration of pre-operative scans to live endoscope images

If images are to be useful in planning or guiding surgical procedures, they must be registered to the patient, such that the image-space which is defined by the 3D array of voxels can be related directly to the real world coordinate system of the patient. As described in section 3.1, segmentation is a major step for many image registration techniques. We also noted that segmentation of specular surfaces is a very challenging task and for this reason we proposed our segmentation method for live endoscope images. As a part of future studies in this area, we propose using the cube overlay program as an interactive device to manipulate the 3D extracted contours of the object which would later be used for image registration. As shown in figure 5.4, the extracted 3D contour of

the object is surrounded by the virtual cube and the user can manoeuvre the extracted contour by manipulating the cube with surgical tools.

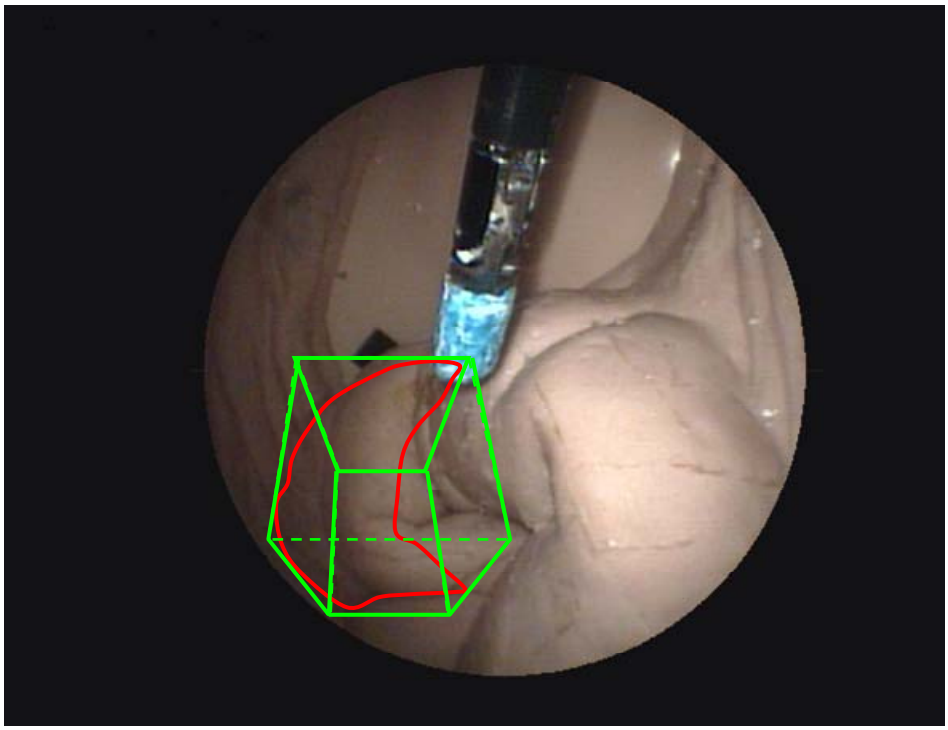


Figure 5.4: The extracted contour is manipulated using the cube overlay system

The proper overlaying of 3D pre-operative scans with live endoscope images could be a potential solution to some of the vision related problems with MIS, such as lack of depth and color contrast. For future studies, we suggest extending the cube manipulation task into a two handed task to register 3D pre-operative scans with live endoscope images. Pre-operative scans and live images of the same organ are not necessarily taken form the same angle. The cube overlay program could be a potential solution to this problem. Figure 5.5 is synthetically created to demonstrate this paradigm. The 3D pre-operative scan of a part of an internal organ is surrounded by the virtual cube.

Using two tracked surgical tools, the users can bring the virtual cube to the actual location of the internal organ and by rotating, resizing and translating the cube, they can ultimately superimpose the pre-operative scan on top of live endoscope images.

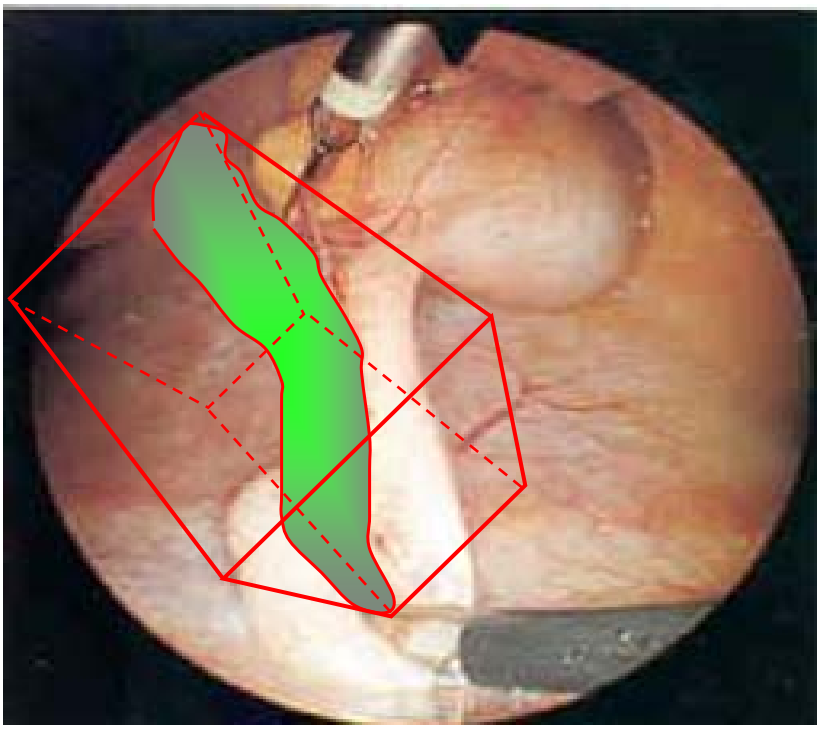


Figure 5.5: The pre-operative scan of an organ is being manipulated by the virtual cube in order to be overlaid on the live images of the endoscope

In order to superimpose pre-operative scans with live endoscope images using the cube program, it might be needed to manipulate the cube with more than pure rotation or translation. In most surgical procedures, soft tissue deforms during the surgery, therefore it is not always easy to relate the pre-operative scans to live endoscope images.

The current cube overlay program only considers the rigid body transformations, namely translate, rotate and resize the cube. For future studies, we can extend the cube deformation to linear and quadratic deformations. The former includes sheer and stretch

deformations while the latter includes twisting and bending. Figure 5.6 shows examples of rigid, linear and quadratic deformations for the cube [36].

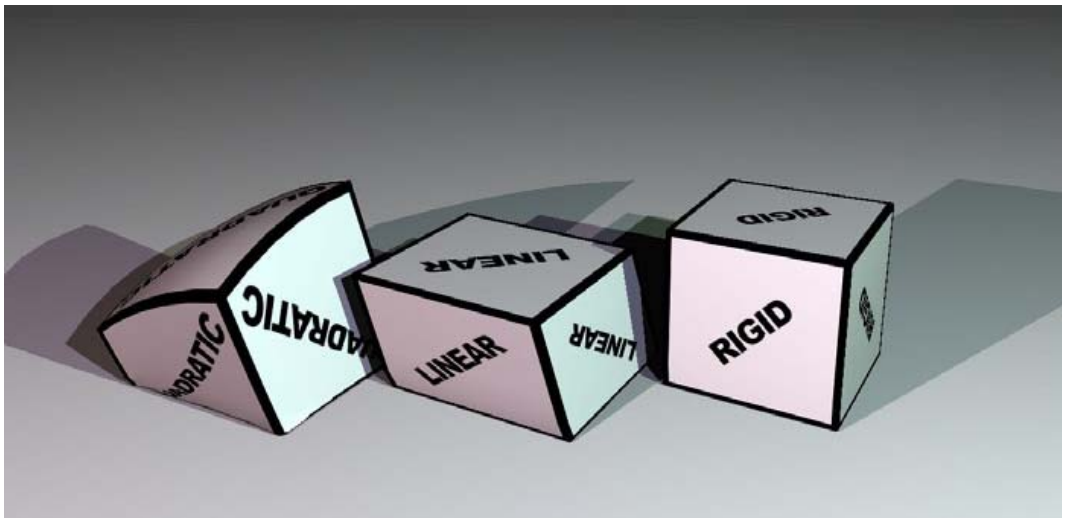


Figure 5.6: Rigid, linear and quadratic deformation of the cube

APPENDICES

Appendix A

Panasonic ideas for life

Zigbee™ Comm Module

IEEE802.15.4

Part No: PAN802154HAR00

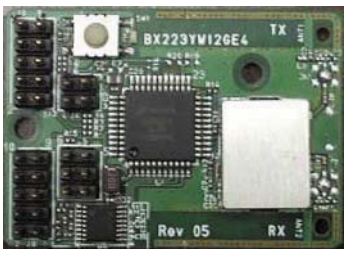
The PAN802154 is a full modem communication module for the 2.4 GHzISM band.It is designed to be in full compliance with 802.15.4 radio standards; and the ZigbeeT protocol layer.

Features:

- RS-232 Serial Interface for digital data interface
- Multiple I/O for sensor inputs and actuator outputs
- Scalable for simple MAC: 802.15.4 MAC; and will be applicable to feuture Zigbee layer
- On board antennas or optional external antennas via connectors
- Switch and LED for customer specific application

Applications:

- Home Automation / Appliance control and monitoring
- Industrial / Security / Asset Tracking
- Commercial / Electronic Labelling
- Medical / Patient Monitor (non-critical)

Actual size 

freescale™ Alliance Member

Industry Canada

Performance specifications, Summary:

Parameter	Description	Remarks
Dimensions	48 x 34 x 3 mm	w/onboard printed antennas
Operating Voltage	2.2 to 3.4 Volts DC	
Operating Temperature	-20 to 70^{o}	
Power Consumption	Tx: 35ma Max Rx: 40ma Max Hybernate: Doze	Dormant is less than 100ua
Output RF Power	0 dB typical 3dBm max	
Receive Sensitivity	-92 dBm typ	
CPU with 4K RAM; 60K Flash	8 bit	
Bit rate	250 Kbits / sec 128 bytes / 4ms	
Interface	RS – 232 Header Header connector for 8 digital I/O; and 2 Analog input	2 ports – 10 bit A/D converter

Panasonic PAN802 154 Module is a low rate/low power communication device based on the Freescale Zigbee Sensor Application Reference Design (SARD) development platform. It operates in the ISM 2.4 standard. The PAN802 154 is shipped ready to be downloaded with Freescale 802.15.4 PHY/MAC layer and the ZigBee protocol layer. The module uses Freescale's 802.15.4 transceiver (MC13193), microcontroller (GT60) and is licenced to use all Freescale ZigBee Protocol stack layer software. Further, the PAN802154 has an on-board RS-232 interface IC and two on-board printed antennas that are etched on both sides of the board optimum RF sensitivity. The entire RF section is shielded to prevent RF leakage and further improve RF performance.

Appendix B

(a) Master module code:

The following is the gist of the code executed by the Master module (the module connected to the switches). This piece of code can be found in main.c.

TMP1_Int_flag is a flag that is set by the timer. Every time a time interrupt happens, this flag is set to 1. The timer is set up in such a manner to give 60 interrupts per second. So the status of the switches is checked 60 times per second.

As one can see, if neither PB1 or PB2 is pressed, it stores a character “0” in the data buffer (gauTxDataBuffer[0]='0'), and sends the package containing that data buffer (MCPSDataRequest(&gsTxPacket)). Then, it resets the timer flag and turns the LED off.

The same scenario applies to other situations; however, for each situation a different character is sent as data.

```
for (;;)
{
    if(( TMP1_Int_flag == 1) && (PB1 != PB_PRESSED) && (PB2 !=
    PB_PRESSED))
    {
        gauTxDataBuffer[0]='0';
        MCPSDataRequest ( &gsTxPacket );  //Sending the Packet.
        TMP1_Int_flag = 0;
        LED = 0;
    }
    else if (( TMP1_Int_flag == 1) && (PB1 == PB_PRESSED)&& (PB2 !=
    PB_PRESSED))
    {
        gauTxDataBuffer[0]='1';
        MCPSDataRequest ( &gsTxPacket );  //Sending the Packet.
        TMP1_Int_flag = 0;
        LED = 1;

    }
    else if(( TMP1_Int_flag == 1) && (PB1 != PB_PRESSED)&& (PB2 ==
    PB_PRESSED))
```

```
      {
          gauTxDataBuffer[0]='2';
          MCPSDataRequest ( &gsTxPacket );  //Sending the Packet.
          TMP1_Int_flag = 0;
          LED = 1;
      }
      else if (( TMP1_Int_flag == 1) && (PB2 == PB_PRESSED) && (PB1 ==
      PB_PRESSED))
      {
          gauTxDataBuffer[0]='3';
          MCPSDataRequest ( &gsTxPacket );  //Sending the Packet.
          TMP1_Int_flag = 0;
          LED = 1;

}
```

(b) Robot module code:

The following piece of code is executed in robot module (the module connected to the frame grabber). It can be found in simple_phy.c in Common directory. MCPSDataIndication() function is called whenever a package is received by the module. The package is a structure and is called gsRxPacket. It contains a member called pu8Data. This member contains the data of the package.

As you can see if data is equal to 48 (character “0”) both OUT1 and OUT2 are set to zero, and the LED is turned off. As one can see, for each data sent by the Master Module there is a corresponding if-statement that sets OUT1, OUT2, and LED appropriately.

```
int MCPSDataIndication(tRxPacket *gsRxPacket)
/* Just a direct return. Main loop will handle it. */
{
      if(  gsRxPacket->pu8Data[0] == 48) //48 is ascii code for 0 (PB1
and PB2 are released)
      {
        OUT0 = 0;
        OUT1 = 0;
        LED  = 0;
      }
      else if(  gsRxPacket->pu8Data[0] == 49) //49 is ascii code for 1
(PB1 is pressed)
      {
```

```
            OUT0 = 1;
            OUT1 = 0;
            LED  = 1;
          }
          else if(  gsRxPacket->pu8Data[0] == 50) //50 is ascii code for 2
      (PB2 is pressed)
          {
            OUT0 = 0;
            OUT1 = 1;
            LED  = 1;
          }
          else if(  gsRxPacket->pu8Data[0] == 51) //51 is ascii code for 3
      (PB1 and PB2 are pressed)
          {
            OUT0 = 1;
            OUT1 = 1;
            LED  = 1;
          }

}
```

References:

[1] Faraz, A. and Payandeh, S. Engineering approach to mechanical and robotic design for minimally invasive surgery(MIS), Kluwer academic publishers, (2000)

[2] http://www.intuitivesurgical.com/index.aspx

[3] http://www.ircad.fr/virtual_reality/robotics.php?lng=en

[4] Taylor, R., H., Lavaiiee, S., Burdea, C., G., and Mosges, R.: Computer Integrated Surgery, *Technology and clinical applications*, pp. 201-203, Massachusetts Institute of Technology, (1996)

[5] Tendick, F., and Cavusoglu, M., C. Human machine interface for minimally invasive surgery, *Journal of Engineering in Medicine and Biology society*. Proceeding of 19th international conference, IEEE/EMBS, vol.6, pp. 2771-2776, (1997)

[6] Peters, P., M. Image-guided surgery: From X-ray to virtual reality, *Journal of Computer Methods in Biomechanics and Biomedical Engineering* OPA Amsterdam, (2000)

[7] Northern digital, http://www.ndigital.com/

[8] Szpala , S., Wierzbicki, M., Guiraudon, G., and Peters, T.M. Real-time fusion of endoscopic views with dynamic 3-D cardiac images: a phantom study, IEEE Transaction on medical imaging, vol.24, No.9, pp. 1207-1215, (2005)

[9] Aquino, S.,L., and Vining, D., J., virtual bronchoscopy*, Journal of Clin. Chest Med*., vol.20, pp. 725-730, (1999)

[10] Konen, W., Schollz, M., Tombrock, S.: The VN. Project: Endoscopic image processing for neurosurgery *, Comput. Aided Surge*., vol.3, pp. 144-148, (1998)

[11] Blackwell, M., Nikou, C., DiGioia, A., M., Kanade, T.: An image overlay system for medical data visualization, *MICCAI'98, LNCS*, pp. 1496, 232-240, (1998)

[12] Chou, W., Wang, T., and Hu, L., "Design of Data Glove and Arm Type Haptic Interface," *11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems,* pp. 422-427, (2003)

[13] Schalk, G., McFarland, D., Hinterberger, T., Birbaumer, N., and Wolpaw, J., "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on Biomedical Engineering*, vol. 51, issue 6, pp. 1034-1043, (2004)

[14] Ko, J., Kim, K., and Ramakrishna, R., "Facial Feature Tracking for Eye-Head Controlled Human Computer Interface," *IEEE Region 10 Conference TENCON*, vol.1, pp. 72-75, (1999).

[15] Nishikawa, A., Hosoi, K., Koara, K., Negoro, D, Hikita, A, Asano, S., Kakutani, H., Miyazaki, F., Sekimoto, M., Yasui, M., Miyake, Y., Takiguchi, S., Monden, M.: Face MOUSe: A novel human-machine interface for controlling the position of a laparoscope, *IEEE Trans. on robotics and automation,* vol.19, no.5, (2003).

[16] Duchowski, A.T.: A breath-first survey of eye tracking applications.*Journal of Behaviour research methods, Instruments and computer*, vol.1,pp. 1-16, (2000)

[17] Lyons, E., Barreto, A., and Adjouadi, M., "Development of a Hybrid Hands-off Human Computer Interface Based on Electromyogram Signals and Eye-gaze Tracking," *23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, pp. 1423-1426, (2001)

[18] Atkins, S., M., Moise., A., and Rohling, R.: An application of eye gaze tracking for designing radiologists' workstation: Insights for comparative visual search tasks. *ACM Transactions on applied perception,* vol.2, pp. 136-151, April (2006).

[19] Atkins, S., M., Fernquist, J., Kirkpatrick, A., E., and Forster, B., B. Evaluating interaction techniques for stack mode viewing. *Journal of digital imaging*, July (2008).

[20] Zhang, X., and Payandeh, S. Application of visual tracking for robot-assisted laparoscopic surgery, Journal of robotic systems, 19(7), pp.315-328, (2002)

[21] Hsu, J., and Payandeh, S. Toward tool gesture and motion recognition on a novel minimally invasive robotic surgery system, IEEE International Conference on Robotics and Automation, pp.631-637, (2006)

[22] Matrox imaging, http://www.matrox.com/

[23] Maintz, J.,B., A., and Viergever, M., A. A survey of medical image registration, Medical Image Analysis, vol.2, no.1, pp.1-36, (1998)

[24] Zitova, B., and Fulsser, J. Image registration methods: a survey. Image and vision computing, 21, pp.977-1000, (2003)

[25] Pole, R., and Toennies, K., D. A new approach for model-based adaptive region growing in medical image analysis, LNCS 2124, pp.238-246, (2001)

[26] Zimmerman, G., Gordon, S., and Greenspan H.: Automatic landmark detection in uterine cervix images for indexing in a content-retrieval system, Biomedical Imaging: Nano to Macro. 3rd IEEE International Symposium on. pp. 1348 – 1351, (2006)

[27] Norman, D., A. The design of everyday things, Perseus Books Group, (2002)

[28] Hsu, J. Toward integration of surgical robotic system with automatic tracking tool gesture and motion recognition. MA.SC thesis, school of Engineering Science, Simon Fraser University, (2006)

[29] Comeau, R., M., Sadikot, A,.F., Fenster, A., and Peters, T., M. Intra-operative ultrasound guidance and tissue shift correction in image-guided neurosurgery. Med.Phys. 27(4), pp.787-799, (2000)

[30] Witt, H., and Kluge, E., M. Domain Expert vs. Layman: Exploring the Effect of Subject Selection in User Studies for Industrial Wearable Computing Applications, Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology, pp.167-174, (2007)

[31] Kass, M., Witkin, A., and Terzopoulos, D. Snakes: Active contour models. *International Journal of Computer Vision 1*, 4, 321–331, (1987)

[32] McInerney, T., and Terzopoulos, D. Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis*,1(2), pp. 91-108, (1996)

[33] Limin, X., Shiwen, G., Xinquan, S., Yaoping, F. 3D surface reconstruction using fuzzy deformable models. Signal Processing Proceedings, 2000,WCCC-ICSP 2000. *5th International Conference on*.vol.2, pp:886 – 889, (2000)

[34] P. Besl and N. McKay. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, pp.239–256, (1992)

[35] DiMaio, S., Kapur, T., Cleary, K., Aylward, S., Kazanzides, P., Vosburgh, K., Ellis, R., Duncan, J., Farahani, K., Lemke, H., Peters, T., Lorensen, W., Gobbi, D., Haller, J., Clarke, L., Pizer, S.,Taylor, R., Galloway, Fichtinger, G., Hata,N., *et al.*

Challenges in Image guided therapy system design, *NeuroImage*, Volume 37, Supplement 1, pp. S144-S151, (2007)

[36] Muller, M, Heidelberg, B., Teschner, M., and Gross, M. Meshless deformations based on shape matching, *ACM Transactions on Graphics*, vol.24, (3), pp.471-478, (2005)