Real-Time Processing of Three Dimensional Ultrasound for Intracardiac Surgery

A dissertation presented

by

Paul Martin Novotny

to

The School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University

Cambridge, Massachusetts

May 2007

Thesis advisor
Robert D. Howe

Author
Paul Martin Novotny

Real-Time Processing of Three Dimensional Ultrasound for Intracardiac Surgery

Abstract

Real-time three-dimensional ultrasound has been demonstrated as a viable tool for guiding surgical procedures. This visualization technique enables a range of new minimally invasive techniques in cardiac surgery. Ultrasound permits visualization through the opaque blood pool in the heart, and the advent of real-time 3DUS overcomes difficulties with 3D spatial perception in conventional 2D ultrasound. These procedures eliminate the need for a cardiopulminary bypass and its well documented adverse effects. However, challenges remain before clinical implementation of ultrasound-guided intracardiac surgery can be made practical. 3D ultrasound provides limited spatial perception due to distorted appearance of tissue and instruments. The low quality imaging when combined with the confined dynamic environment of a beating heart makes complex tasks difficult for surgeons to perform.

These challenges are addressed by improving the display of 3D ultrasound with stereo displays, tracked surgical instruments, and introducing robotics for handling the dynamic cardiac tissue. A study is presented that demonstrates stereo displays improve surgical performance, reducing error rates by 50% and procedure times by 28%. The stereo display was found to help surgeons better interpret the low resolution ultrasound volumes. However, a large source of ultrasound distortion is imaging artifacts created by surgical instruments. A detection technique is presented that identifies the position of the instrument within the ultrasound volume. The algorithm uses a form of the generalized Radon transform and passive markers to identify the full position and orientation of the instrument. This detection technique is amenable to rapid execution on parallel computer architectures and as a result real-time instrument tracking was attained. With a tracked instrument location, it is now possible to develop enhanced surgical displays. Instrument tracking also enables robot guidance that is ideally suited to compensate for the highly dynamic environment of beating heart surgery. A system has been designed and tested that uses real-time instrument and target tracking for robot control.

Contents

Acknowledgments

Looking back on this entire process, I realize there is very little that I did without the help of others. New ideas were the result of conversations with the people around me, new directions came from the advice of my mentors. My results would not have been possible without the people who helped design, build, run, or participate in the experiments. After six years of graduate school, I have a lot of people to thank.

Those who helped the most on a day to day basis are the undergraduates I had the fortune to work with. I happened to have the chance to work with three of the best, Dan Kettler, Richard Plowes, and Jackie Stenson. Dan Kettler, a fellow Wisconsin native, was an enormous help in this thesis. I joked he was a 'turbo boost' for PhD students he helped, and he helped many PhD students in our laboratory. His extraordinary abilities as an engineer in design, fabrication, and reasoning touched almost every part of this thesis and I am very thankful to have the opportunity to work with him. Richard Plowes also had a large impact on this work. His enthusiasm and determination were of great use in hunting down doctors and surgeons to obtain much needed ultrasound data. His contribution to characterizing mitral valve motion was astounding given the short time he worked on the project. I also need to thank Jackie Stenson who played a large role in the teaching side of my PhD experience. She was an excellent teaching fellow who helped me with Engineering Sciences 51. I am extremely grateful for her picking up the slack in teaching responsibilities as I struggled to finish my research. She proved to be an excellent teaching fellow.

Fellow members of the Biorobotics Laboratory are what I will remember most about my graduate school experience. I truly learned the most from them. They were the first people I turned to for help with a challenging problem. More often than not, someone would know the answer and would selflessly spend countless hours helping me. From the beginning there was Sol Diamond, Anna Galea, Heather Gunter, Amy Kerdok, and Guy Rachmuth who helped me with classes and adjusting to graduate school life. Ryan Beasley, Aaron Dollar, Peter Hammer, Yuri Ishihara, Petr Jordan, Marius Linguraru, Shelten Yuen, and Chris Wagner, helped me immensely in the later stages of graduate work. I was lucky enough to work closely with one member of the laboratory, Steve Jacobsen. He played a large part in Chapter 2 of this thesis, spending long hours helping set up experiments and running subjects. Special thanks to Doug Perrin, the wisest among us, who always knew how to help a struggling graduate student, whether it be research advice, encouragement, or a friend to have a beer with.

The doctors I worked with during this thesis are extraordinary. Dr. Pedro del Nido, Dr. Gerald Marx, Dr. Jeremy Cannon and Dr. Yoshihiro Suematsu all found time in their busy schedules for my work. Dr. Nikolay Vasilyev especially helped in all of the surgical trials in this thesis. I am thankful for his patience as I fumbled to get my equipment working. Dr. Ivan Salgo, Bernie Savord, and Scott Settlemier of Phillips Medical provided valuable knowledge of the ultrasound machine and without them none of this work would have been possible. I need to also thank Jeff Stoll who I worked with very closely. We were both trying to finish our PhDs and worked side by side on many projects. We had some intense periods trying to meet paper deadlines.

In the end, we produced some work that I am very proud of. I will always remember the days we tried to collect data at the hospital in any room we could, including the nurses' lunch room.

I could not have asked for a better advisor, Professor Robert Howe. I was lucky enough to find an excellent advisor who did everything for his students. He graciously helped with papers late at night or on weekends to meet a deadline and provided that special balance of guidance and freedom that is so important in open ended research. I also need to thank my committee members Professor Roger Brockett, Professor Pierre Dupont, and Professor Todd Zickler who helped shape this work. The meetings I had with them were easily the most productive hours of my graduate career. Their insightful comments and criticism played a large part in the direction of this thesis.

The support of my family and friends played a large part in my personal well being during this process. I need to thank my brother Eric who I spent many vacations with to get away from graduate school. Spending time with him always helped rejuvenate me after a difficult semester or qualifying exams. My girlfriend Kathleen Larney's love and support made the last few years of hard work possible. There were many days I came home frustrated, angry, or disappointed. Within a matter of minutes of being around her, my troubles were quickly forgotten and I remembered what was really important. She was always there to pull me away from school and research to go skiing, camping, or any number of activities. She turned a grumpy graduate student into a very happy person.

The two people who have been the biggest part of this process are my parents. They have been involved in this for almost thirty years and have provided the love, support, and encouragement throughout. Thanks to my mother who more than anyone is always looking out for what is best for me, who rightfully encouraged me to return to graduate school when I left after my masters degree. Thanks to my father, who is my biggest role model. His example has taught me how to be an engineer, teacher, and scientist. My mother tells a story that when I was young I followed my father around with a toy lawn mower while he mowed the lawn. Finishing graduate school is no different, I am still trying to be like my dad.

Dedicated to my parents,
Lynn and Vladimir.

Chapter 1

Introduction

Cardiac surgery has long been aided by the use of the cardiopulmonary bypass (CPB), a process that diverts the blood supply from the heart to an external heart-lung machine. When employing CPB, the heart is disconnected from the circulatory system. The surgeon is then able to work on a relaxed, open, and empty heart while the body continues to receive a supply of oxygen rich blood. As a result, hundreds of thousands of cases that were impossible to treat before are now treated every year [48]. However, the use of CPB is not without its side effects. There are many adverse effects, including delay of neuro-development in children, mechanical damage produced by inserting tubing into the major vessels, stroke, and significant decline in cognitive performance [3, 32, 56]. Murkin *et al.* [32] reported incidence of stroke at 2% of patients and noticeable cognitive dysfunction one week and two months after surgery, at rates of 85% and 45%, respectively. Although these side effects are insignificant compared to otherwise untreatable diseases, an alternative to CPB is desirable.

CPB, developed in the 1950s, brought about a revolution in cardiac surgery [15]. Prior to CPB, attempts were made to perform procedures on the inside of the heart [1, 18, 20]. However, it proved extremely difficult as surgeons operated through incisions in the heart wall while it continued to pump blood. Complications arose due to large amounts of blood loss and a high risk of stroke from the introduction of air into the blood stream. Procedures were further complicated by the limited amount of visual information as the field of view is obstructed by the blood supply.

Recently, surgeons are taking a second look at beating heart surgery because of the attendant risks of CPB. Coronary artery bypass grafting without the use of CPB has received recent attention [4]. This procedure alleviates blocked arteries on the outside of the heart by grafting new arteries to bypass the blockage. Traditionally, the procedure is done by stopping the heart with a CPB so the surgeon can work on a stationary heart. By mechanically stabilizing the heart, CPB is not needed and a corresponding decrease in post operative stroke is observed [40].

For intracardiac procedures, beating heart procedures without CPB are more

complicated and new challenges arise. During intracardiac procedures performed with a CPB, the heart is stopped and drained of blood. The surgeon works on the target site with direct physical and visual contact. This is no longer available during beating intracardiac procedures because of the presence of blood. The blood is opaque so endoscopy is impossible; in addition, physical access is restricted because of the need to limit blood loss.

We argue that image-guided surgery is the key to beating intracardiac procedures. With the right imaging modality, it is possible to peer inside the heart while it continues to function, eliminating the need for CPB. Image guided procedures have many recent successes in other areas of medicine, including neurosurgery [17], orthopedic surgery, and urology. Each of these techniques permits the visualization of structures otherwise inaccessible or that are visualizable only with a greater degree of invasiveness.

1.1 Ultrasound Guided Surgery

Real-time 3D ultrasound is capable of guiding intracardiac beating heart procedures. Ultrasound is a widely used imaging modality in cardiology because of its ease of use, non-ionizing field, and high contrast between cardiac tissue and blood. While ultrasound is traditionally 2D, recent developments have extended ultrasound to a 3D imaging modality. Pioneered by Von Romm *et al.* [44], real-time 3D ultrasound images are created by electronically steering the ultrasound pulse. In much the same way that a 2D probe scans a plane using a 1D phased array of transducers, volumes are scanned by a 2D array of piezoelectric ultrasound transducers. Current commercially available systems, such as the Philips SONOS 7500, use a two-dimensional array of 3,000 piezoelectric elements running at 2-4 MHz to build a full 3D volume. By varying the delays between excitation of the elements, the ultrasound beam is steered throughout the volume, producing a 128 x 48 x 205 dataset from a 7 x 3 x 8 cm volume at 20-25 volumes per second.

Cannon *et al.* [5] found complicated 3D tasks, such as grasping and suturing, are possible with 3D ultrasound guidance. In addition, surgeons have begun to explore 3D ultrasound procedures in animal models [49, 50]. Two cardiac procedures that are particularly promising for ultrasound guidance are atrial septal defect closure and mitral valve repair.

Atrial septal defects are holes between the left and right atria of the heart (Figure 1.1). Typically 5-10 mm in diameter, the defect allows for oxygenated and unoxygenated blood to mix between the left and the right side of the heart. This causes cardiac efficiency to decrease and the patient's heart enlarges to compensate. In some cases heart failure results. Treatment for this defect involves attaching a patch made of pericardium or synthetic material to the atrial septum to cover the defect. Currently, this is done with an open chest surgery where the heart is put on CPB, opened,

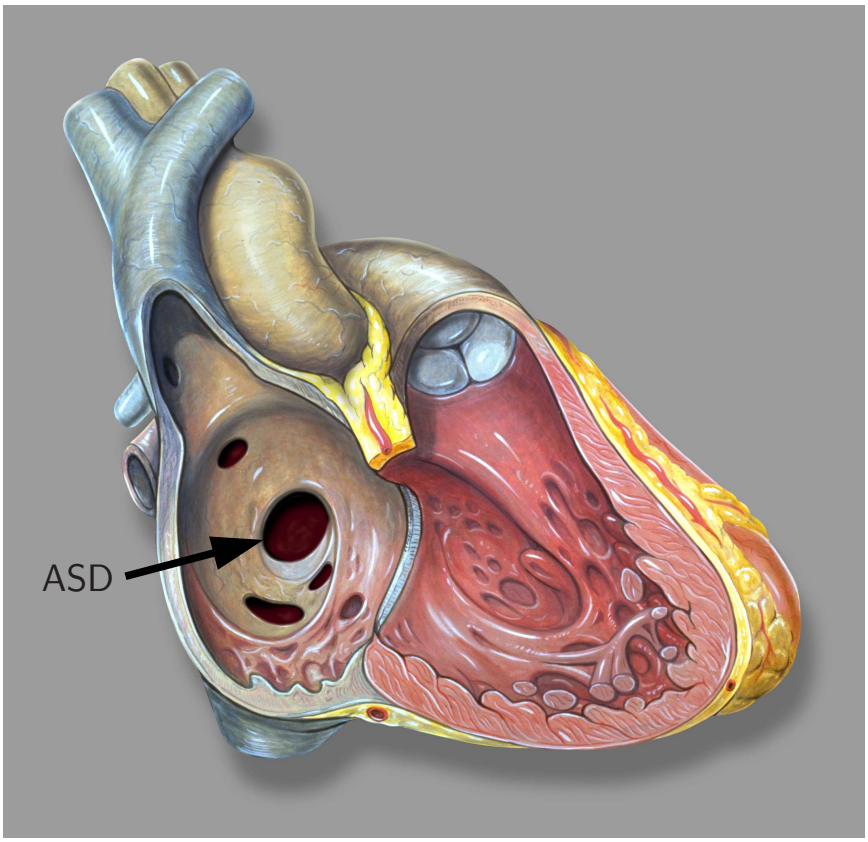


Figure 1.1: Atrial septal defect (ASD) shown from the right side of the heart. (Patrick J. Lynch, medical illustrator; C. Carl Jaffe, MD, cardiologist)

and drained of blood. The surgeon sutures a patch to the atrial septum covering the defect.

Mitral valve repair typically involves the suturing of a prosthetic ring to the annulus of the valve (Figure 1.2). Diseased mitral valves often have a dilated annulus and as a result the leaflets are unable to fully close. This causes blood to leak back into the atrium during ventricular contraction. Mitral valve repair is performed under CPB with an arrested stationary heart. The surgeon enters through the left atrium and attaches a prosthetic ring to correct annulus dilation by tightening the circumference of the valve.

1.2 Thesis Overview

Initial animal trials highlighted many challenges to the goal of ultrasound guided intracardiac surgery [49, 50]. This thesis focuses on the issues created by using 3D

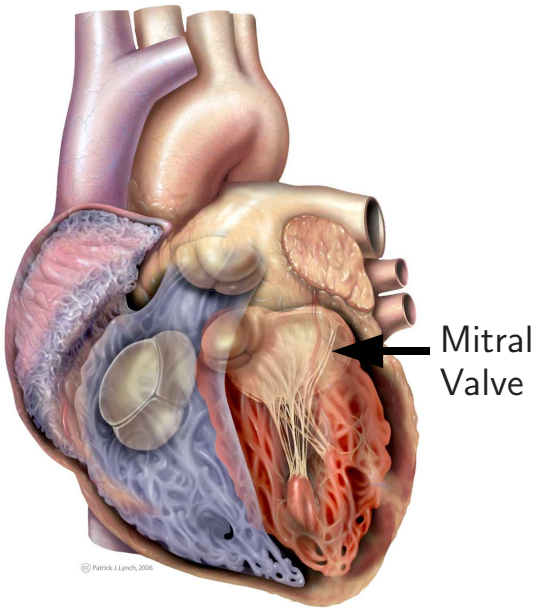


Figure 1.2: View of the mitral valve. The mitral valve ensures unidirectional blood flow between the left atrium and left ventricle. (Patrick J. Lynch, medical illustrator; C. Carl Jaffe, MD, cardiologist)

ultrasound to guide intracardiac procedures. The challenges arise from two main areas. First, 3D ultrasound produces noisy low resolution images that are difficult to interpret. In particular, surgical instruments look incomplete and distorted, making it difficult to identify and navigate. Second, working in a beating heart is challenging for the simple fact that it is moving. To work in this environment, the function of the heart can not be disturbed as the heart supplies blood to the rest of the body. In addition, because the heart is beating, there are many fast moving structures that must be avoided, or interacted with.

Chapter 2 addresses the first challenge, the noisy low resolution ultrasound images. While these images are difficult to interpret, they are truly 3D volumes. Traditionally, ultrasound displays are 2D monitors that do not show the depth in the volumes. Chapter 2 looks at using stereo displays to help surgeons interpret the images. The effect of stereo displays on performing surgical tasks is examined with both tank trials and *in vivo*.

Chapter 3 studies the tracking of surgical instruments to help visualization during surgery. Instruments, when imaged by ultrasound, are especially susceptible to imaging artifacts that make comprehension difficult. This chapter introduces a method to track surgical instruments in real-time. The accuracy and robustness of the technique is tested in tank studies and *in vivo*. Applications of the instrument tracking

technique are proposed for improving visualization or enabling robot guidance during cardiac surgery.

Chapter 4 addresses the challenge of performing surgery on fast moving cardiac structures with a robot. During cardiac procedures, such as mitral valve repair, it is necessary to safely interact with dynamic structures. A robot is used to compensate for this high speed motion. The instrument tracking algorithm developed in Chapter 3 is used to track an instrument held by a robot. The robot's ability to synchronize with a target moving at mitral valve speeds is tested.

While the techniques presented here are developed in the context of beating heart intracardiac surgery, they have applications to image guided procedures in other surgical specialties, including liver, neural, and fetal interventions.

Chapter 2

Stereo Displays for 3D Ultrasound

Almost all medical images are displayed on a flat two dimensional computer monitor. Unlike the natural environment around us, this presentation removes the depth information, or the three dimensional qualities that help us comprehend shapes, distances, and orientations of objects. With the advent of three-dimensional ultrasound, true 3D volumes are produced. However, these images are still displayed on two-dimensional monitors. In ultrasound this is especially detrimental, as removing the depth information removes a key feature to help interpret the noisy low resolution images.

To observe three dimensions, we are equipped with two eyes that view the environment from two slightly offset positions. The brain uses the difference between these viewpoints to infer three-dimensions, or depth. When images are presented on a two-dimensional monitor, both eyes receive identical views of the scene and depth is lost. To restore three dimensions, each eye must receive a slightly offset version of the scene, mimicking what occurs naturally. Several technologies exist today that accomplish this in a variety of ways. Shutter glasses, binocular displays, and stereo LCD displays use different principles to deliver stereo information.

Using stereo displays for endoscopic surgery has had some success, but not attained widespread use. The importance of using 3D displays was noted by Cushieri *et al.* [7] to improve minimally invasive surgery. Satava [42] found that stereo displays would greatly help surgeons accurately move within 3D space. He proposed using dual camera laparoscopic systems combined with a LCD shutter glasses to display the 3D information. Peitgen *et al.* [37] compared stereo and monocular endoscopes with a user trial of 60 surgeons. They found the irrespective of experience, trial times and number of errors decreased by 25% and 50% across all subjects. Although the authors noted that the subjects had no experience with the tasks before the trial. They predicted that 3D displays may not affect experienced surgeons' performance with familiar tasks because they can cope with a 2D display's limitations.

This interpretation was confirmed during a clinical trial by Hanna *et al.* [19]. 4 specialists performed 60 laparoscopic cholecystectomies, 30 trials with a stereo laparo-

scope, and 30 with a monocular laparoscope. Each of the surgeons had significant experience with the procedure, at least 20 clinical cases. It was found that the 3D display had no effect on procedure time or error rate. This confirms that experienced surgeons are able to adapt to the absence of depth by using cues in the 2D image. Kim *et al.* [26] performed a simulated tracking task where subjects followed a target point in 3D space using monocular and stereo displays. With a stereo display, subjects' performance was superior to monocular display. However, when depth cues were added to the display, the performance with a monocular display was equivalent to stereo.

When considering stereo laparoscopic systems, there is a trade-off between using one high resolution camera or two lower resolution cameras for stereo. This is due to size limitations of minimally invasive surgery. With this trade-off, surgeons prefer higher resolution cameras over a stereo display. Van Bergen *et al.* [52] showed that using a high resolution camera improved procedure times, although the results were not statistically significant. Largely due to natural depth cues that infer three dimensions, such as foreshortening, occlusion, and shadows, surgeons are able to compensate for the loss of information. As a result, higher resolution and smaller cameras are preferred over stereo displays for endoscopic surgery.

For ultrasound, these factors are not present. A monocular or stereo display use the same ultrasound probe and have the same resolution. In addition, natural depth cues are absent in the low resolution ultrasound images. For these reasons we believe that the addition of a stereo display to three dimensional ultrasound will help surgeons comprehend the images and therefore improve their ability to perform surgical tasks. To test this hypothesis, a stereo renderer was developed for displaying 3DUS volumes on a stereo display. Stereo displayed 3DUS was tested in both tank trials and *in vivo* with animal trials.

2.1 Stereo Rendering of 3D Ultrasound

To produce stereoscopic images of 3DUS data we developed a custom real-time renderer. For real-time visualization, the system must handle and render 30 MB of data every second. This is accomplished by harnessing the computational power of consumer-level graphics processing units (GPU). Driven by entertainment applications, the computational GPU capacity is moving beyond current CPU capacity, allowing visualization of large volumetric data sets at interactive frame-rates [28]. The fundamental advantage of programmable GPUs is their ability to execute highly paralyzed user routines (pixel shaders). Our implementation uses pixel shaders to cast rays through the volumetric data set in a ray-per-pixel fashion.

The volumetric rendering is produced by first loading the 3DUS volume into a 3D texture on the GPU. In order to obtain the intensity of each rasterized pixel, the data set is sampled along a projection ray through the data set. The intensity (i)

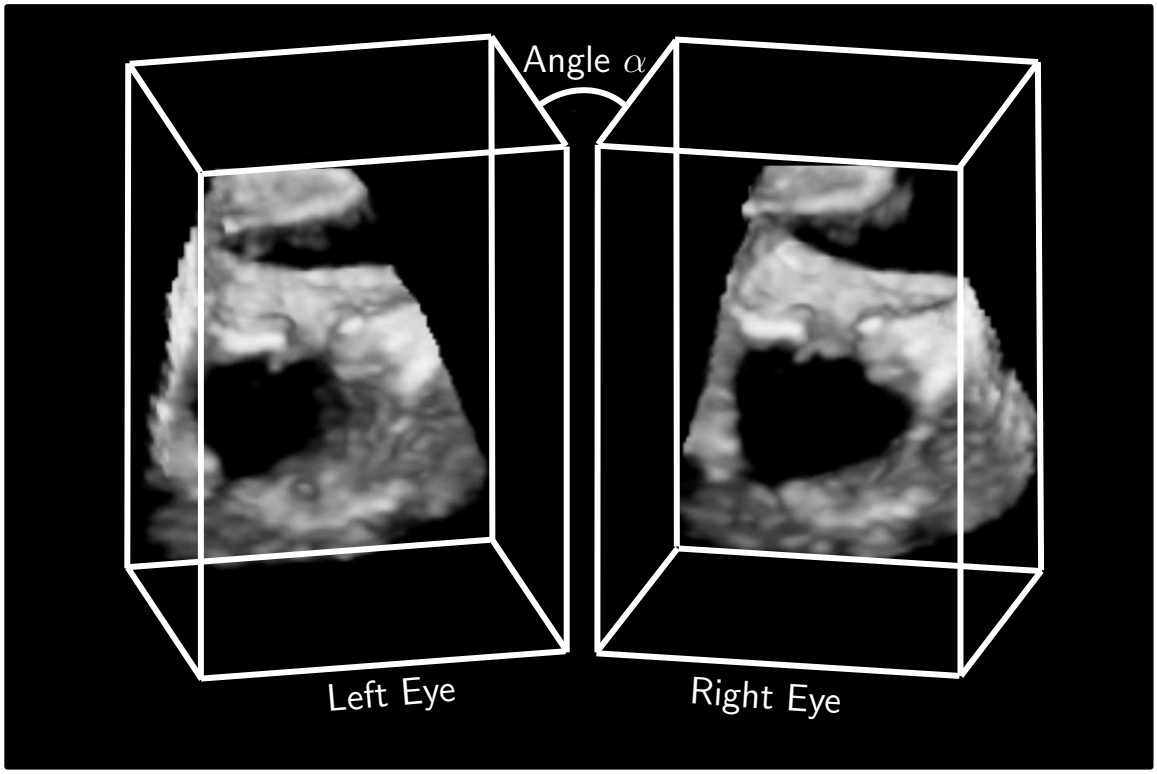


Figure 2.1: Stereoscopic views of a 3D ultrasound volume. Left and right eye views are rendered with an angular separation α to simulate binocular vision.

and opacity (o) of the pixel is compounded in a front-to-back fashion. The initial sampling is performed at the intersection of the projection ray and the data set. Subsequent samples of the ultrasound image (I_k) are obtained along the projection ray, with a sampling distance of one voxel to prevent aliasing artifacts. The sampling and compounding iteration may be summarized as

$$i_{k+1} = i_k + (1 - o_k)I_k, \tag{2.1}$$

$$o_{k+1} = o_k + (1 - o_k)I_k. \tag{2.2}$$

The projection rays are terminated as the projection ray leaves the data set.

For stereoscopic viewing, two 7800GT (nVidia Corp, Santa Clara, CA) GPUs render the ultrasound volume in parallel. Left eye and right eye views are separately generated by rendering the 3DUS volume from two viewpoints, mimicking a left and right eye (Figure 2.1). When rendering typical 3DUS volumetric data sets in full-screen mode (640x480 screen resolution), the renderer maintains highly interactive

frame rates of 70 fps and above.

The separation of the two views is set such that the ultrasound volume appears 0.5 m from the subject's eyes. This distance corresponds to the actual distance between the subject's eyes and their hands, providing a natural interaction.

2.2 Tank Study of Surgical Performance

To study the effect of stereo displays with 3DUS guided procedures, a rigorous tank study was performed with subjects with varying levels of surgical experience. The stereo rendered ultrasound data was used in conjunction with the high quality stereo display of the Da Vinci surgical robot (Intuitive Surgical Inc., Sunnyvale, CA) (Figure 2.2). By testing subjects in a controlled tank environment, repeatable measurements of performance are possible. As a result, the variable conditions present in surgery is eliminated, and the effect of stereo displays is fundamentally studied.

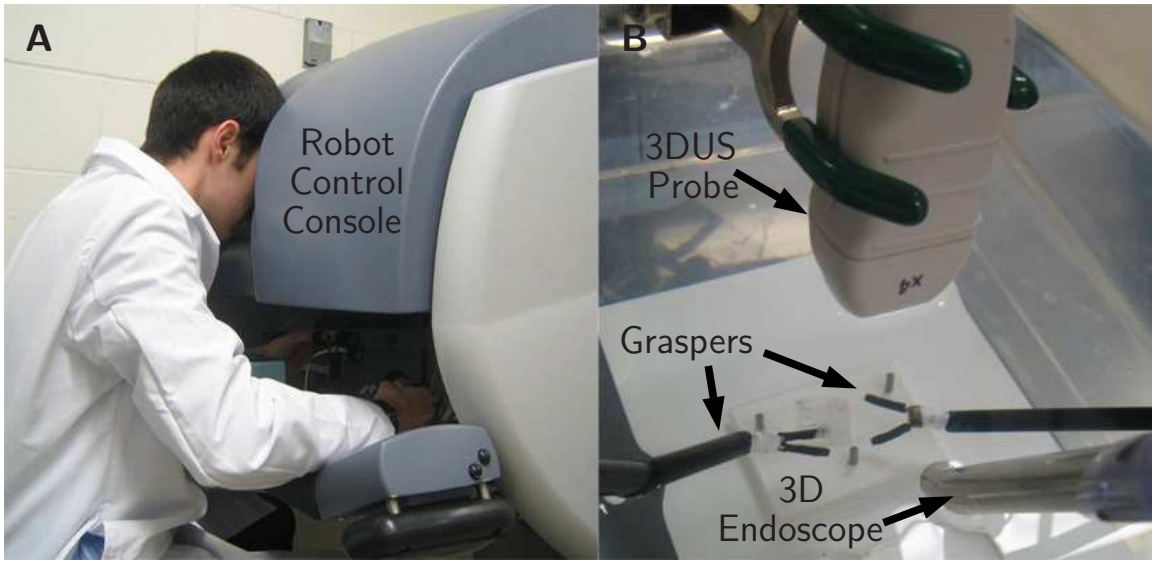


Figure 2.2: (A) The stereo 3D view is shown on a binocular display on the robot control console. User motions at the controls are mapped to two surgical graspers shown in (B).

2.2.1 System Overview

The ultrasound volumes are produced by a SONOS 7500 3DUS machine (Philips Medical Systems, Andover, MA). The 3DUS volumes, typically 128 x 48 x 204 voxels, are created at 25 Hz and sent over a 1 gigabit TCP/IP network to a personal computer

running the rendering algorithm. As the data is received from the ultrasound machine, it is loaded to both GPUs through a PCI-Express bus. Each GPU renders the volume and produces an image from either the left or right eye viewpoint. This image is then passed to the binocular display of the surgical robot through an S-Video connection. As a result, surgeons use the stereo rendered ultrasound data for guiding a surgical procedure as they control the robot from the console. The console also contains all of the controls necessary for the surgeon to control the movements of the robot during surgery.

The surgical robot is a 7 degree-of-freedom robot, with 10 mm laparoscopic instruments mounted on each manipulator (Figure 2.2). The surgeon's movements at the master console are mapped to two surgical graspers controlled by the robot. The workspace of the robot consisted of a 45 x 60 x 15 cm tank of water over which the robot manipulators, the ultrasound probe, and a 10 mm, 0 degree stereo endoscope were mounted. Due to the high acoustic reflectance of metal, surgical graspers such as those used in this study appear distorted and incomplete in ultrasound. To minimize distortion and improve their appearance, the metal graspers were coated with electrical heat shrink tubing and Teflon tape to cover their highly reflective surfaces. An absorptive nickel-impregnated rubber mat was placed in the workspace to reduce ultrasound reflections from the bottom of the tank.

2.2.2 Performance Metrics

To quantitatively evaluate surgical performance, sixteen test subjects performed a pegboard and rope pass tasks using each vision system. These two tasks were selected from among a number of laparoscopic training tasks proven to correlate well with laparoscopic surgical skill [6, 8, 43]. The tasks were chosen for their emphasis on depth perception as well as their suitability to the limiting factors of the ultrasound's imaging characteristics and field of view. The test subjects were required to complete these tasks guided by both monocular-displayed and stereo-displayed 3DUS as well as by a stereo surgical endoscope which provided a benchmark for comparison.

Pegboard (Figure 2.3A): In this task, the test subject picked up a plastic collar sitting around one peg, passed the collar between manipulators and placed the collar around a second peg. To minimize the effects of any learned muscle motion, the initial and destination pegs for the collar were rotated during every trial. The pegs were arranged in a triangular pattern on a 5.0 x 6.5 cm acrylic base. The acrylic collar used was 1.3 x 1.3 x 1.2 cm with a 1.2 cm hole. This relatively large hole made the collar easier to grasp with the surgical manipulators. This emphasizes the amount of time required to maneuver the manipulators into the correct position to grip the collar, the task portion most affected by depth perception. The test subject could use the left and right manipulators in either order to move the collar depending on their own preferences. One trial of the pegboard task consisted of repeating the basic grasp, pass, and replace motion once under one vision system. Subjects performed

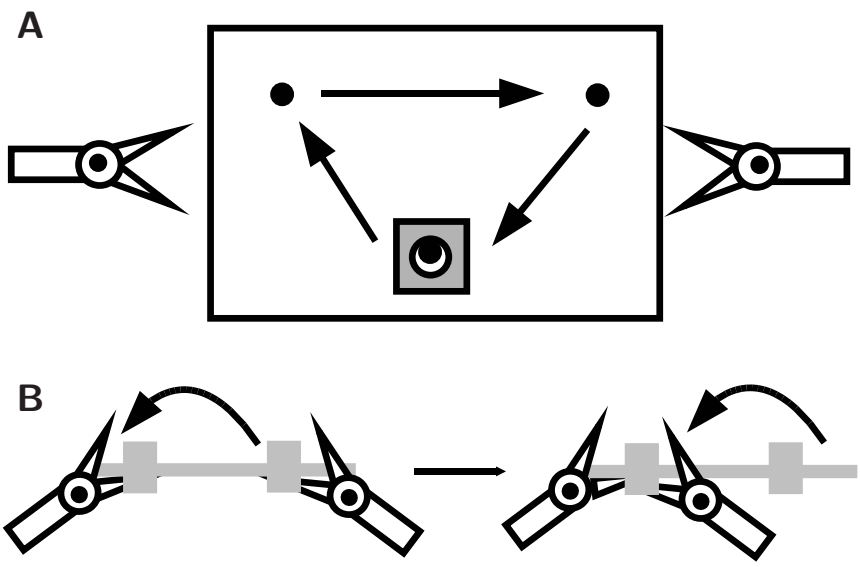


Figure 2.3: (A) Diagram of the peg board task. Subjects were instructed to pick up the collar from one peg, pass the collar to the opposite grasper, and place the collar on the next peg. (B) Diagram of the rope pass task. In this task, subjects pass the rope from grasper to grasper. Subjects were instructed to first grab the rope just to the left of the next knot, then grab just to the right of the same knot with the right grasper. In this fashion they moved along the rope until they passed five knots.

nine trials, three trials for each of the three vision systems.

Rope Pass (Figure 2.3B): The rope pass task consisted of passing a knotted rope from the left to the right manipulator. A 0.5 cm diameter nylon rope with five knots on 3.0 cm centers was used. The test subject started each trial gripping the rope to the left of the first knot with the left manipulator. The rope was passed one knot at a time by gripping the rope with the right manipulator to the right of the knot currently held by the left manipulator and then moving the left manipulator to the left of the next knot. One trial of the rope pass task consisted of passing the rope five knots under one vision system, starting from the same knot under each vision system. Nine trials were performed, three trials for each of the three vision systems.

2.2.3 Procedure

The 3DUS vision system was evaluated by three groups of test subjects. The first group consisted of eight graduate students with no prior surgical experience. These subjects' relevant previous experience ranged from research involving ultrasound imaging and surgical robots to no familiarity with the systems being tested. The second group consisted of four medical doctors with one to ten years of surgical

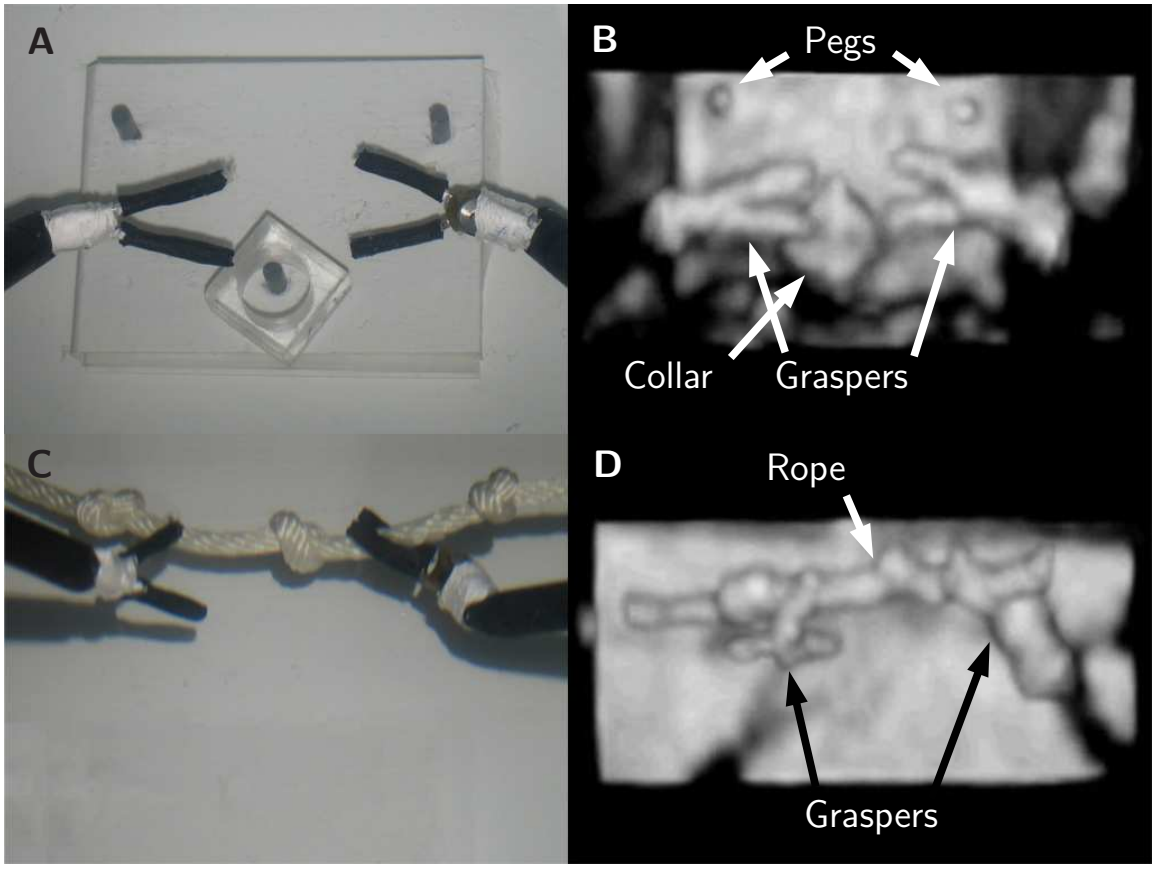


Figure 2.4: Image (A) and corresponding 3DUS image (B) of the peg board task. (C) and (D) show the image and 3DUS image of the rope pass.

experience, but without experience with surgical robotics. The final group consisted of four surgeons with 7-10 years of surgical experience and at least 1 year of clinical experience with surgical robotics.

Before performing the actual trials, the test subjects were required to complete a short practice program. This practice session was intended to bring the subjects to a standard level of ability and to limit learning effects during trials. Practice typically took 15 minutes and was divided into two sections. In the first, the test subject worked under the guidance of the stereoscopic endoscope and manipulated the rope and collar that would be used in the two tasks. This familiarized the subject with the controls of the robot and the objects used in the tasks. The practice session continued until subjects were able to demonstrate proficiency by picking up and passing the collar between graspers five times and passing the rope five knots. In the second practice section, the subjects completed trials of each task with each vision system to become familiar with guiding the robot under the trial conditions. These final practice trials

were run like the actual data collecting trials and provided an opportunity to become comfortable with the tasks and testing procedure.

Following the practice program, the test subjects completed the actual tests. The tests consisted of nine trials of the pegboard task followed by nine trials of the rope pass task, for a total of three trials for each task-vision system combination. During the actual tests, the order of the vision systems used for each trial was counter balanced in order to further remove any effects due to learning.

At the end of the session, the subjects completed a short survey about their experiences with the three vision systems. They were asked to rate the ease of use of each vision system on a scale of 1 to 5. In addition, the survey asked the subjects to rate their confidence and level of mental fatigue while using the three vision systems.

2.2.4 Analysis

Trial time and number of errors were used to measure task performance. In the peg board task, errors were defined as dropping the collar during the transfer between the hands, or incorrectly placing the collar. In the rope pass, each time the subject dropped the rope was considered one error. Statistical analysis was used to determine if there is a significant effect of the stereo display on subjects' performance. Comparisons were done to determine the effects of stereo-displayed 3DUS on mean and variance of the performance metrics. An F-Test was used to determine if the variances of the distributions are significant ($p < 0.05$), and a Student's t-test to determine any significant difference in mean ($p < 0.05$).

2.2.5 Results

The results of the study show that stereo-displayed 3DUS improves subjects' ability to complete simulated surgical tasks with a surgical robot. As shown in Figure 2.5, the vision system had an effect on the number of errors that the users committed while performing the tasks. Subjects experienced almost no errors while using endoscopic guidance. In the case of the stereo-displayed 3DUS, the average number of errors increased to 0.41±0.16 (mean ± standard error) errors per trial. However, this error rate was not as high as with 2D-displayed 3DUS, where subjects either dropped the collar or rope at a rate of 0.81±0.19 times per trial, a 100% increase over stereo-displayed 3DUS. A Student's t-test confirms that using a stereo display significantly decreases the number of errors during the trials ($p < 0.05$).

The vision system also had an effect on task completion times for both the rope pass and pegboard (Figure 2.7). Trial times decreased when stereo-displayed 3DUS was used compared to monocular-displayed 3DUS. Across all subjects the completion times for the pegboard task were, 10±1 s (mean±standard error) with the stereo endoscope, 36±8 s with the stereo 3DUS, and 56±18 s with mono-3DUS. For the rope pass the completion times were 16±1 s with the stereo endoscope, 40±4 s with

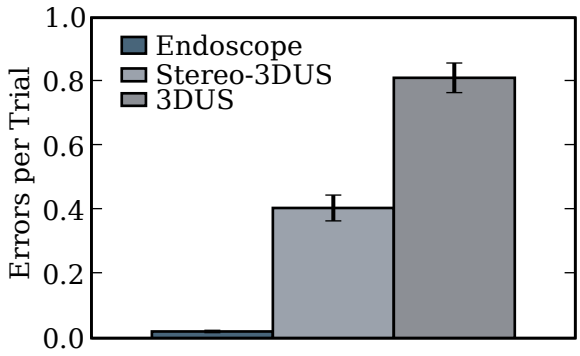


Figure 2.5: Average number of errors per trial. With 3DUS, the use of a stereo display reduced the average number errors per trial. Error bars indicate standard error.

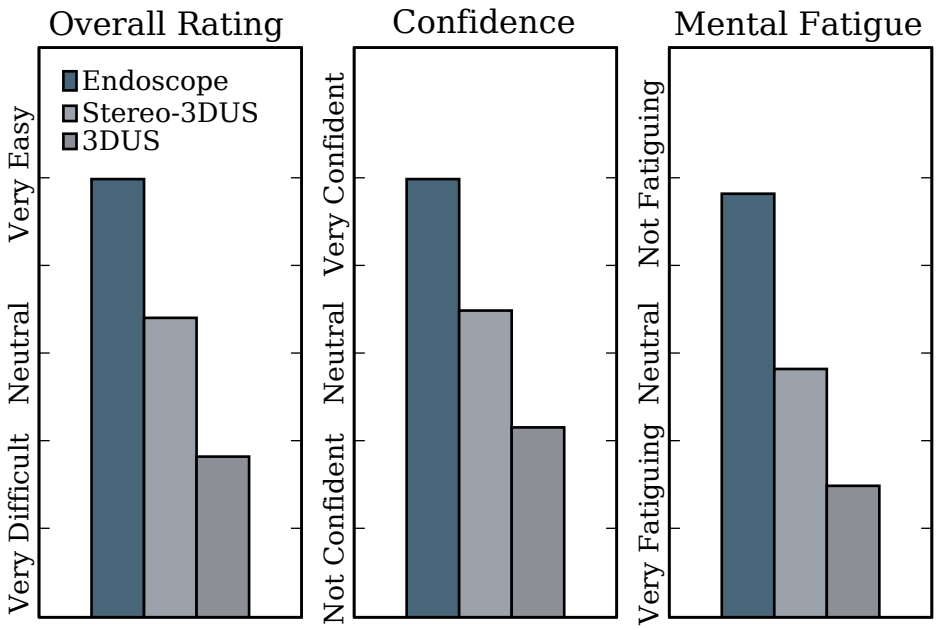
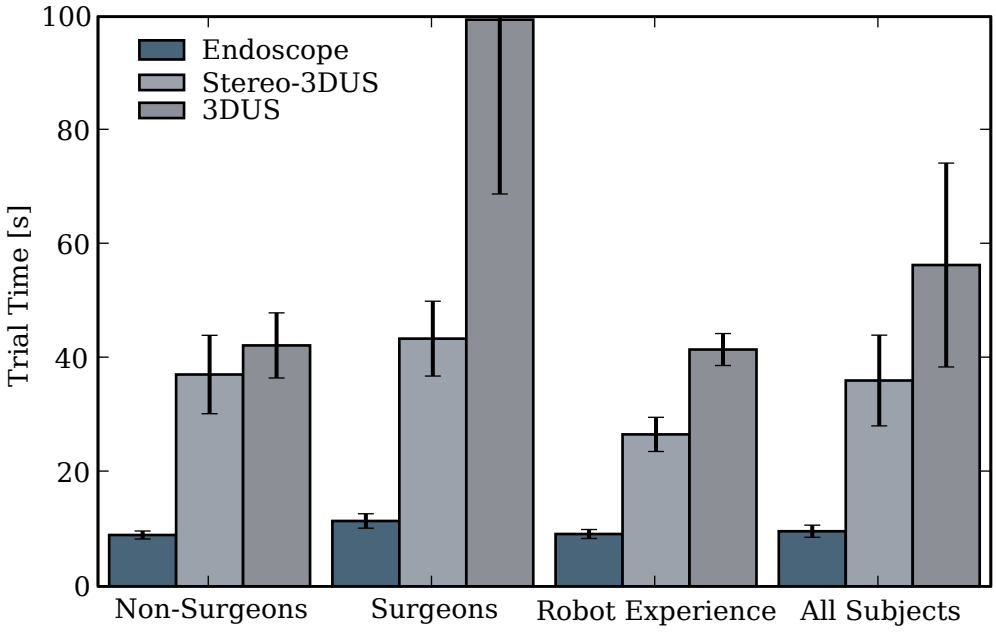
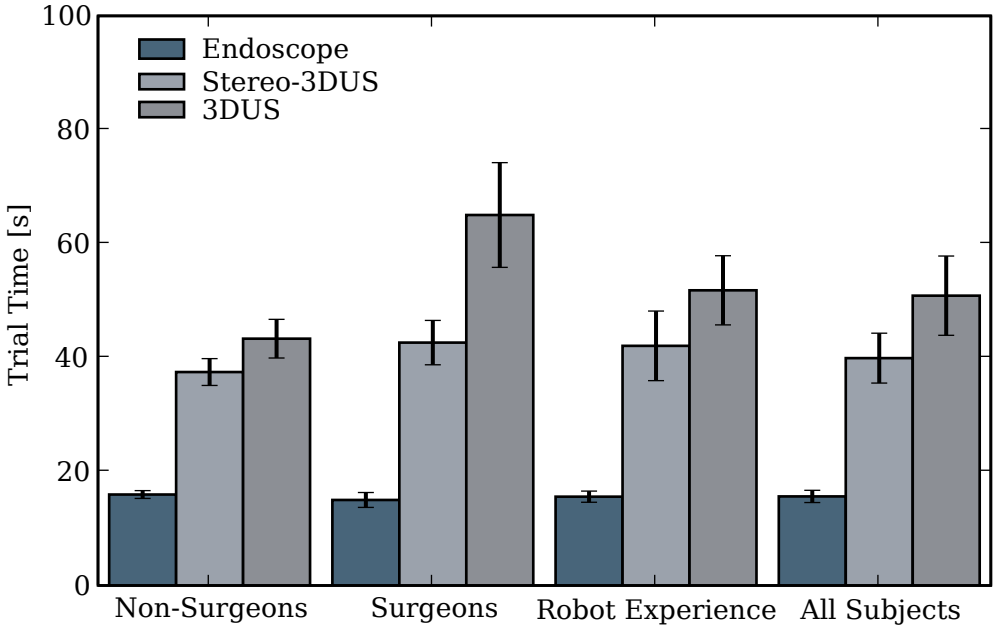


Figure 2.6: Stereo-displayed 3DUS was subjectively rated easier to use, more confidence inspiring, and less fatiguing than mono-displayed 3DUS.

the stereo 3DUS display, and 51±7 s with the normal 3DUS display. Statistical analysis investigated the stereo display's effect on the standard deviation and mean of the completion times. An F-test demonstrated that the standard deviation of



(a) Pegboard task



(b) Rope pass task

Figure 2.7: Trial times for surgical tasks. Error bars indicate standard error.

completion times for both rope pass and pegboard were significantly decreased with the use of a stereo display with 3DUS ($p < 0.05$). Mean times were also significantly decreased ($p < 0.05$), by 35% for the pegboard and 20% for the rope pass.

Analysis of the survey responses showed subjects' preference for stereo-displayed 3DUS over 2D-displayed 3DUS (Figure 2.6). When asked to rate the three vision systems on a five point scale from very difficult(1) to very easy(5), subjects rated the stereo endoscope, stereo 3DUS, and 3DUS as 5.0, 3.5, and 1.8, respectively. Similar responses were given when subjects were asked to rate their confidence with each vision system from not confident(1) to very confident(5). Subjects rated 3D endoscope at 5.0, stereo-displayed 3DUS at 3.5, and 2D-displayed 3DUS at 2.1. Finally, when asked to rate the level of mental fatigue they experienced on a scale from very fatiguing(1) to not at all fatiguing(5), they rated the stereo endoscope, stereo-displayed 3DUS, and 3DUS at 4.9, 2.8, and 1.5, respectively.

2.3 In Vivo Evaluation

Animal trials were performed to test the utility of stereo-displayed ultrasound for guiding surgical tasks. While the previous study examined the effectiveness in a controlled laboratory environment, it did not recreate all the factors present in surgery. In these trials, a surgeon performed an actual surgical procedure under ultrasound guidance. Performance using stereo-displayed 3DUS was compared to mono-displayed 3DUS.

2.3.1 Protocal

The surgical procedure used for this evaluation was an atrial septal defect closure (Figure 2.8). This procedure involves placing anchors around the perimeter of a patch to secure it to the atrial septum. When properly fastened to the septum, the patch eliminates the flow between the two atria of the heart.

The ultrasound probe was placed epicardially on the right atria to give a view of the left and right atrium. The ultrasound volumes were streamed to the same 3DUS stereo renderer described in Section 2.1. The rendered volumes were displayed on a 19 inch monitor and stereoscopic viewing was achieved with LCD shutter glasses (eDimensional, West Palm, FL). The shutter glasses alternate between blocking the left and right eye. To create a three dimensional display, the rendered ultrasound view on the monitor alternates between left and right eye views, synchronized with the shutter glasses. For monocular views, the monitor was used without the shutter glasses, and the viewpoint remained static.

Four Yorkshire pigs (70-80 kg) underwent the open chest beating heart procedure. For direct access to the beating heart, a median sternotomy was performed. Two purse string sutures were placed on the right atrial appendage for instrument access.

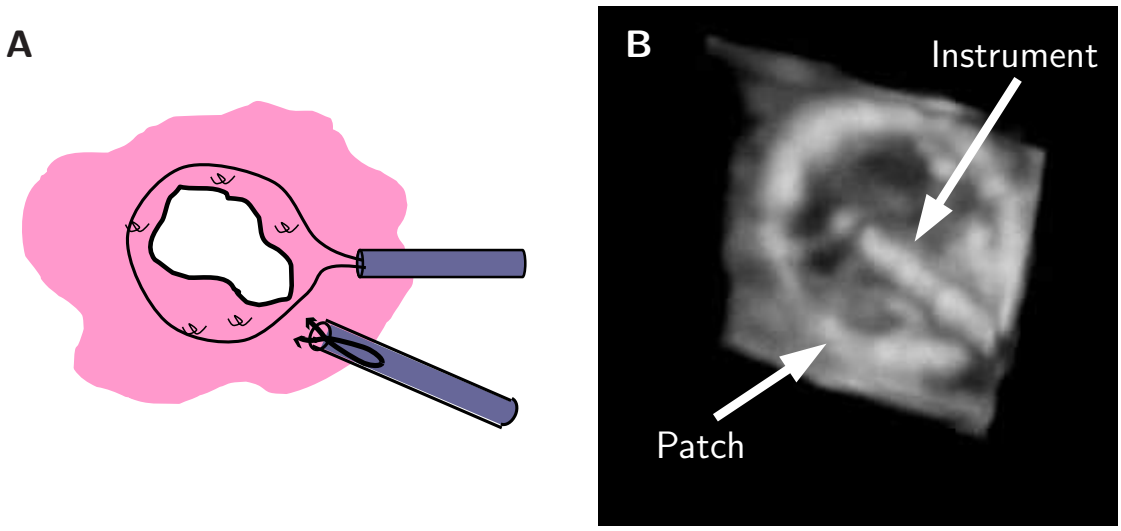


Figure 2.8: Schematic (A) and ultrasound image (B) of an atrial septal defect repair. In this procedure a patch is inserted into the heart to cover the atrial septal defect. An anchor driver is used to attach the patch to the septum.

These purse string sutures allowed insertion and removal of surgical instruments, while minimizing blood loss. To prepare the animal for atrial septal defect repair, an artificial defect was created. A balloon catheter punctured the atrial septum to create a 6 mm hole between the left and right atrium.

Dr. Nikolay Vasilyev (Children's Hospital Boston) performed the procedures. Dr. Vasilyev has significant experience in endoscopic surgery and image guided beating heart surgery. These four trials were split into two groups, two trials for mono and two for stereo-displayed 3DUS. During each study, the surgeon placed 12 anchors around the circumference of a 15 mm diameter ASD patch. The surgeon started with the instrument tip just inside the ultrasound volume, approximately 3 cm above the patch. The surgeon moved the tip toward a position around the circumference of the patch and deployed an anchor. This was repeated eleven more times to place enough anchors to secure the patch to the cardiac tissue and sufficiently block the flow of blood between the left and right atria.

Two metrics were used to evaluate the surgical performance: time to place an anchor and how well the movement followed a straight line. For analysis of the instrument tip trajectory, the tip position was recorded using an electromagnetic tracking system (miniBIRD 800, Ascension Technology, Burlington, VT). The RMS difference between the tip position trajectory and a straight trajectory was used to quantify the linearity of the trajectory. Statistically significant differences between stereo and mono-displayed 3DUS were found using a Student's t-test ($p < 0.05$).

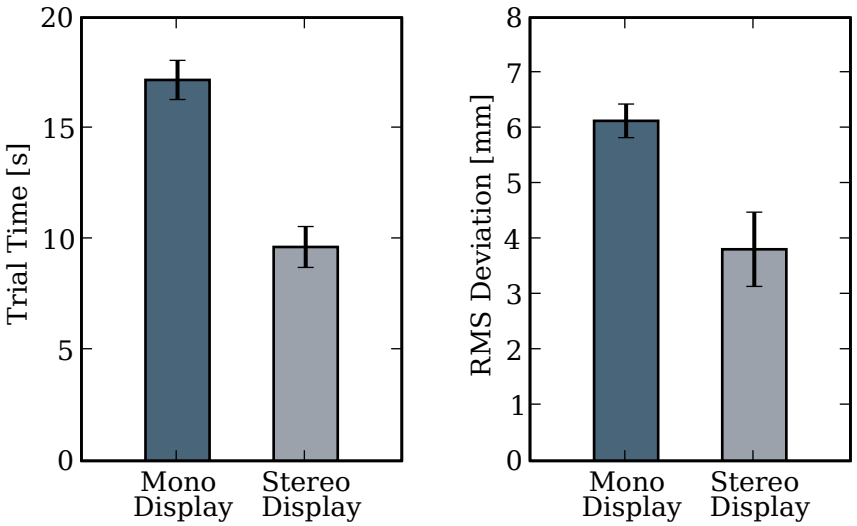


Figure 2.9: Average task completion times and RMS trajectory deviation from a straight line for stereo and mono-displayed 3D ultrasound. Error bars indicate standard error.

2.3.2 Results

Significant improvement was found in trial time and accuracy of tip position movement using the stereo displayed 3DUS. Figure 2.9 shows that anchor placement time decreased from 17.2±0.8 (mean±standard error) to 9.7±0.9 seconds when using the stereo display. Statistical analysis showed this be a significant improvement ($p < 0.05$).

The ability of the surgeon to move in a straight line from the start position to the desired anchor position also improved with the stereo display. When using the monocular display, the subject deviated from a straight line with an RMS error of 6.1±0.1 mm. With the stereo display, this deviation dropped by almost 40% to 3.8±0.2 mm, a statistically significant improvement ($p < 0.05$). Representative trajectories of the tip are shown in Figure 2.11 for stereo and mono-display. These trajectories were chosen as their RMS deviation from a line most closely match the mean for each group. At the end of the procedure, the heart was excised and the patch was photographed for two of the animal trials (Figure 2.10). Qualitatively, the pictures show that the anchors deployed with stereo guidance are more evenly spaced than without.

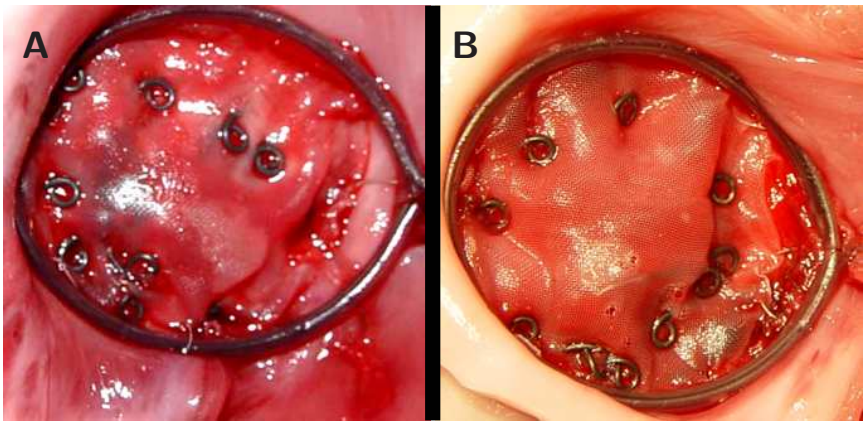


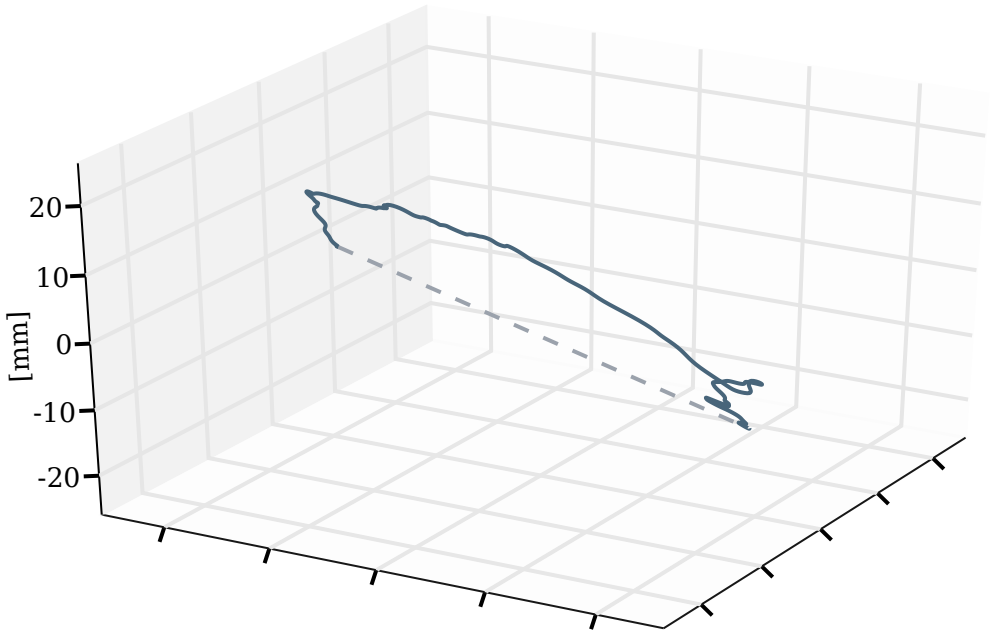
Figure 2.10: Post operative images of an ASD patch anchored to the atrial septum. Image (A) shows the results of using mono-displayed ultrasound for guidance, and image (B) is with stereo-displayed ultrasound.

2.4 Discussion

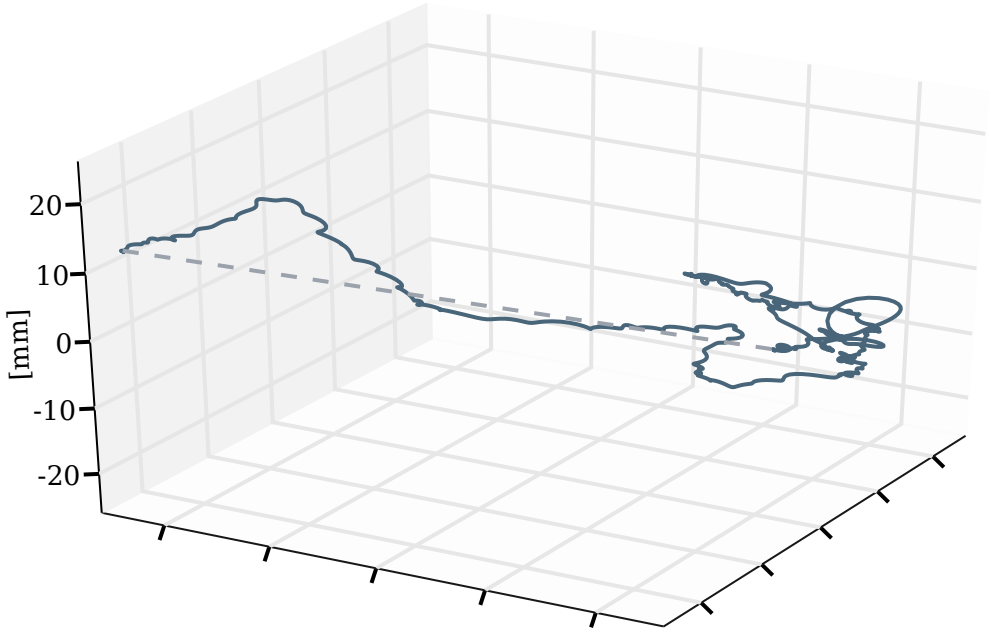
Stereo-displayed 3DUS improves surgical performance as compared to the normal mono-displayed 3DUS. Stereo-displayed 3DUS presents an attractive enhancement of 3DUS by helping improve a surgeon's ability to interpret the noisy ultrasound images. This is especially useful in procedures where endoscopes are not feasible, such as intracardiac and fetal surgery.

To study the effects of depth information with 3DUS guided procedures, two manipulation tasks were used to evaluate surgical proficiency in a tank trial. These tasks compared the ability to conduct manipulations with a surgical robot using 3DUS with and without a stereo display. These tasks, however, did not fully model the complete surgical environment, nor fully explore the set of movements or tasks a surgeon performs during a procedure. In addition, the tasks were conducted inside a water tank instead of within a dynamic *in vivo* environment. To mitigate these differences, the tasks were carefully selected from laparoscopic training tasks that are proven to correlate with surgical skill [6,8,43]. As a result, differences in performance with these tasks in a water tank should be representative of differences seen in actual surgical procedures.

The tank trial results were verified with an *in vivo* animal trial. While this trial only involved one subject due to ethical and cost restrictions on animal use, it did provide valuable data on the effects of stereo displays in surgical procedures. As was seen in the tank studies, trial times were significantly reduced with stereo-displayed 3DUS. In addition, the subject found it easier to navigate in three dimensions. This is highlighted by the trajectories seen in Figure 2.11. For stereo-displayed 3D ultra-



(a) Stereo-displayed 3DUS



(b) Mono-displayed 3DUS

Figure 2.11: Instrument trajectory with stereo (a) and mono (b) displayed 3DUS.

.

sound, the surgeon was able to make a direct movement to the target anchor location. With mono-displayed ultrasound, the surgeon did not take a direct path to the target location and searched for the correct location at the end of the trajectory. The surgeon noted a reliance on tactile feedback to find a proper anchoring location with monocular 3D ultrasound.

The performance improvement seen with the stereo display is consistent with results found with 3D endoscopes in laparoscopic [12, 24, 51] and robotic [13] surgery. These studies have demonstrated the improvement of surgical performance when comparing normal 2D endoscopes with 3D endoscopes. Widespread use of 3D endoscopes has not come about due to lower image quality and larger size compared to traditional endoscopes. Furthermore, surgeons are adept at using depth cues, such as foreshortening, occlusion, and shading, that infer depth in endoscopic images. In the case of 3DUS, it is a different story. There is no image quality degradation or need for a larger ultrasound probe. The stereo display uses the same probe and 3DUS volume for both stereo and non-stereo display. In addition, natural depth cues are not present in ultrasound images, making it difficult to infer three-dimensions. As a result, there are few foreseeable disadvantages to adoption of stereo 3DUS displays.

Advantages of using a stereo display with 3DUS are highlighted in this study. Specifically, incidences of errors, task completions time, and repeatability were significantly improved with the stereo-displayed 3DUS. Analysis of trial times showed that subjects were faster and more consistent when using stereo-displayed 3DUS; the mean and variance of the completion time significantly decreased for both tasks in this study. While fundamentally subjects were performing the same movements to complete the tasks regardless of vision system, when things went wrong, the effect of the vision system was most apparent. With the 3D endoscope, a subject's understanding of their movements and the environment is near perfect and errors are quickly corrected. With 3DUS, both stereo and non-stereo, if everything is going smoothly, users were able to complete the tasks without trouble. However, when errors in trajectory or understanding of the environment occur, they are not immediately comprehended by the user. These errors are more quickly recognized and corrected under stereo 3DUS than mono 3DUS. The effect is highlighted by the significant difference in the standard deviation of the trial times. In the worst case, subjects dropped the rope or misplaced the collar. Such errors occurred 100% more often with the 2D-displayed 3DUS.

In addition to the objective results, subjects also expressed a subjective preference for the stereo-displayed 3D ultrasound over mono-displayed 3DUS. Subjects felt more confident in their movements and experienced a lower level of mental fatigue with the stereo-display. Subjects preferred the 3D endoscope in all aspects over 3DUS, but using stereo 3DUS is a more feasible alternative to mono 3DUS in situations where endoscopic guidance is not possible.

The results of this experiment demonstrate the utility of using stereo-displayed 3DUS for improving surgical performance. As 3DUS guided procedures become more

prevalent, these results suggest that clinicians should seriously consider the addition of a stereo display. As we have shown, the adoption of a stereo display lowers error rates, increases speed, and improves consistency, three traits very important for surgical procedures.

Chapter 3

Instrument Tracking in 3D Ultrasound

Ultrasound has proven to be a versatile non-invasive technique for imaging parts of the human body. When used for surgery, surgeons find it difficult to navigate instruments with 3D ultrasound in the dynamic, confined intracardiac space. Tools look incomplete and distorted, making it difficult to distinguish and orient. This is largely due to the incompatibility of ultrasound and hard objects. Ultrasound is optimized for imaging soft tissue and fluids. Hard objects such as metal, plastics, and bone cause reflections, reverberations, and scattering of the acoustic energy that interferes with the imaging.

To address this issue, researchers are developing techniques to localize instruments in ultrasound. Enhancing the displayed position of the instrument allows surgeons to more accurately control the instruments as they perform surgical tasks. In addition, real-time tracking of instruments in conjunction with a surgical robot opens the door for a range of enhancements, such as surgical macros, virtual fixtures, and other visual servoing techniques [27, 36].

Previous work in instrument detection can be broadly separated into two categories: external tracking systems such as electromagnetic and optical tracking [29, 30] and image-based detection algorithms [10, 34, 53, 57]. External tracking systems have suffered from the limitations of the surgical environment. Electromagnetic tracking has limited accuracy and is problematic to implement due to the abundance of ferro-magnetic objects in the operating room. Optical tracking of instruments is complicated by line-of-sight requirements. Both of these systems suffer from errors introduced by improper registration of the ultrasound image coordinates to the tracking coordinate frame. To eliminate such errors, image based algorithms are used to track instruments within the ultrasound image. Most of this work focused on tracking needles [10, 57] and more recently surgical graspers [53] in 2D ultrasound images. As 3D ultrasound systems have become widely available, these 2D techniques have been adapted for implementation in 3D. An appealing approach to instrument localization

is the Radon or Hough transform. These techniques have wide spread use in 2D image analysis for detection of wide variety of shapes. Applications of these techniques have focused on detecting 2D objects in 2D images, however, Hough and Radon based techniques have shown promise in 3D medical image analysis. Most relevant is a needle tracking technique for prostate biopsy [9] that projects the ultrasound volume onto two orthogonal planes. A Hough transform is then performed on the two 2D images to identify the needle.

Beating heart intracardiac procedures pose different challenges and requirements than the 2D breast and prostate biopsy procedures in previous work. For example, the high data rates of 3D ultrasound machines, 30-40 MB/s, require very efficient algorithms for real-time implementation. Previous 2D ultrasound techniques are too computationally costly, or inappropriate for three dimensions. These methods are only appropriate for finding bright objects such as needles in ultrasound images that standout amongst relatively homogeneous tissue. In cardiac procedures, larger instruments such as anchor drivers and graspers do not stand out amongst the surrounding dynamic, heterogeneous environment. To work in this environment, the algorithm must be efficient for handling the large data rates, and capable of distinguishing instruments from fast moving cardiac structures of similar intensity.

In this work we present a technique capable of detecting instruments used in minimally invasive procedures, such as endoscopic graspers, staplers, and cutting devices. In the following sections the interaction of instruments with ultrasound is examined and used to develop a tracking technique. A generalization of the Radon transform is proposed for identifying instrument shafts in 3D ultrasound volumes. Furthermore, we show that this technique can be implemented on a parallel architecture such as inexpensive PC graphics hardware, enabling the detection of instruments in real-time. To achieve full six degree of freedom tracking of a surgical instrument, the Radon based tracking is combined with passive markers for identification of the tip position. The algorithm is tested in both tank studies and an *in vivo* animal trial.

3.1 Surgical Instruments in Ultrasound

In contrast to elastic solids, materials found in the human body such as tissue, blood, and fat, reflect a small amount of acoustic energy. This is one of the underlying assumptions that make medical ultrasound possible. The basic principle behind medical ultrasound relies on time of flight measurements from small impedance variations in the image field. As the signal encounters a change in the impedance of the medium, part of the signal is reflected and part is transmitted. This effect is governed by the impedance of the material or

$$Z = \rho c \tag{3.1}$$

Table 3.1: Density, speed of sound, and impedance values for biological materials compared to engineering materials (at 37C)

Material	**Density** $[\frac{kg}{m^3}]$	**Speed of Sound** $[\frac{m}{s}]$	**Impedance** $[\frac{kg}{m^2 s}]$ **x** 10^6
Air (1 atm)	1.14	353	0.000402
Water	993	1527	1.52
Blood	1060	1530	1.62
Fat	950	1460	1.39
Muscle	1080	1590	1.72
Bone	1200-1800	2700-4100	3.2-7.4
Aluminum	2700	6420	17.3
Steel	7900	5800	45.8

with ρ and c defined as the local density and speed of sound, respectively. At a discontinuity there is a change in the impedance

$$Z_1 = Z_0 + \Delta Z. \tag{3.2}$$

The reflected and transmitted energy for perpendicular incidence with the discontinuity is defined by

$$R = \left| \frac{Z_1 - Z_0}{Z_1 + Z_0} \right|, \tag{3.3}$$

$$T = \left| \frac{2Z_0}{Z_1 + Z_0} \right|. \tag{3.4}$$

For ultrasound interacting with most mediums found in the human body, ΔZ is small and Equations 3.3 and 3.4 are approximated by

$$R \approx \left| \frac{\Delta Z}{2Z_0} \right|, \tag{3.5}$$

$$T \approx 1.0. \tag{3.6}$$

As a result, ultrasound transducers are designed for small amounts of reflected energy while almost all of the energy is transmitted through the tissue. As shown in Table 3.1, this assumption is correct for most materials found in the human body, with bone and air pockets being the major exceptions. When solid instruments are introduced into the tissues, many of the assumptions made during ultrasound design are no longer valid. The density, speed of sound, and impedance of many engineering materials is vastly different from tissue, blood, and other mediums found in the body (Table 3.1). As a result, the approximations made in Equations 3.5 and 3.6, are no longer valid. For this high impedance discontinuity, a significant portion of the energy

is reflected, as well as a much smaller proportion of the energy is transmitted. As a result, clinical ultrasound machines will not display these materials correctly.

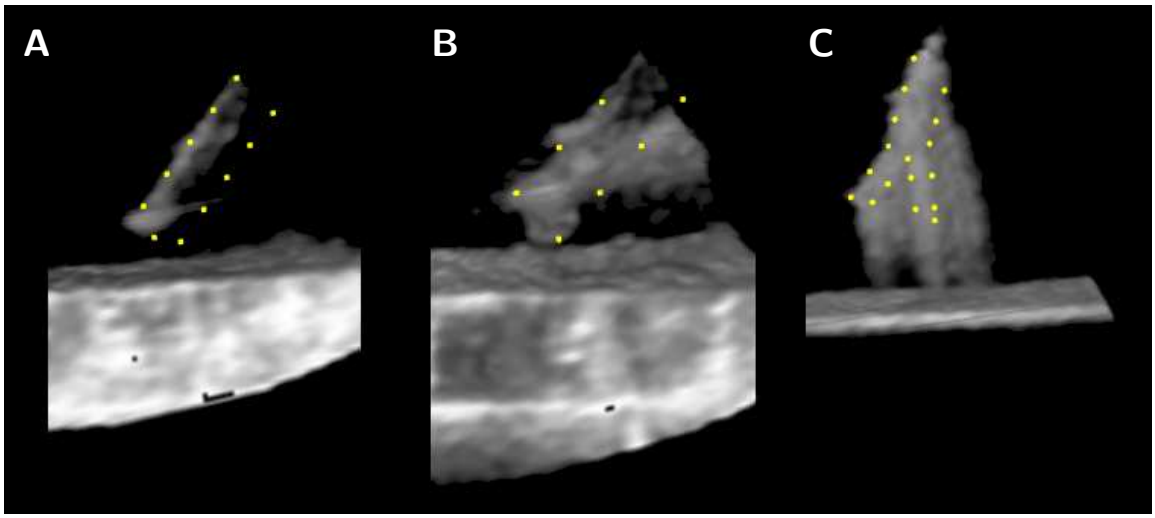


Figure 3.1: 3D ultrasound images of (A) 6.2mm acetyl rod; (B) 6.2mm aluminum rod; (C) endoscopic grasper. Estimated actual shape shown as dotted line. Transducer is at the top aiming down.

The leading surface of the object is the most complete portion of the instrument image (Figure 3.1). It is a large impedance discontinuity that scatters energy back to the source. As the ultrasound propagates, it encounters the cylinder and reflects a significant proportion of the energy, much larger than would be found in normal muscle-blood boundary.

3.2 Instrument Axis Tracking

Instruments used in minimally invasive procedures reflect most of the acoustic energy from its leading surface. Minimally invasive instruments are typically long and straight with a 3-10 mm diameter. At almost all orientations and positions, these instruments appear as long straight objects, a feature that is not found in cardiac tissue. We use a form of the Radon transform to identify these instruments within the ultrasound volumes. The Radon and Hough Transforms are widely used for detecting lines in two dimensional images. In its original formulation [41], the Radon transform maps an image, I, into the Radon space $\mathcal{R}\{I\}$

$$\mathcal{R}\{I\}(d,\phi) = \int I\left(d\cos(\phi) - s\sin(\phi), d\sin(\phi) + s\cos(\phi)\right) ds. \tag{3.7}$$

d and ϕ define a line in two dimensions by its perpendicular distance to the origin and slope, respectively. s is a free parameter that corresponds to a specific point on the shape. This method transforms a difficult image processing problem into simply identifying maximums in the Radon space. In its original formulation, the Radon transform is unsuitable for instrument tracking in 3D, but has been extended for identification of arbitrary shapes. Using the notation introduced by Luengo Hendriks *et al.* [31], points on a parametric shape are defined by the function $\mathbf{c}(s, \mathbf{p})$ where $\mathbf{p}$ defines the parameters of the shape. In this framework, the Radon transform is rewritten as

$$\mathcal{R}_{\mathbf{c}(s,\mathbf{p})}\{I\}(\mathbf{p}) = \int I\left(\mathbf{c}(s, \mathbf{p})\right) ds. \tag{3.8}$$

A line segment in 3D is parametrized with 6 variables. (x_0, y_0, z_0) defines the center of the line segment in 3D. Two angular parameters, (θ, ϕ), describe its orientation, and L defines its length. More compactly, these 6 parameters are written as

$$\mathbf{p} = [x_0, y_0, z_0, \theta, \phi, L]\,. \tag{3.9}$$

Points lying on a line segment are now defined for $\mathbf{p}$ and s as

$$\mathbf{c}(s, \mathbf{p}) = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} \cos(\theta)\cos(\phi) \\ \sin(\theta)\cos(\phi) \\ \sin(\phi) \end{bmatrix} \left(\frac{L}{2}\right) s, \quad s \in [-1, 1]. \tag{3.10}$$

Combining Equation 3.8 and Equation 3.10 yields a form of the generalized Radon transform for line segments in 3D volumes

$$\mathcal{R}_{\mathbf{c}(s,\mathbf{p})}\{I\}(\mathbf{p}) = \int_{-1}^{1} I\left(\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} \cos(\theta)\cos(\phi) \\ \sin(\theta)\cos(\phi) \\ \sin(\phi) \end{bmatrix} \left(\frac{L}{2}\right) s\right) ds. \tag{3.11}$$

Identifying lines in the 3D volume now becomes a problem of finding local maximums of $\mathcal{R}_{\mathbf{c}(s,\mathbf{p})}\{I\}(\mathbf{p})$, from Equation 3.11, where $\mathbf{p}$ denotes a local maximum or likely instrument position. In other words, we integrate the image volume, I, along a direction defined by (θ, ϕ), through the point (x_0, y_0, z_0) and identify maximums. This scheme is illustrated in Figure 3.2 where integrations are illustrated for multiple directions. Figure 3.2C contains points with a high integral, or simply, the image is brighter than the other four images (Figure 3.2A, 3.2B, 3.2D, and 3.2E). This is a result of the correspondence of the integration direction and the object's axis. As a result, by finding the maximum value of $\mathcal{R}_{\mathbf{c}(s,\mathbf{p})}\{I\}(\mathbf{p})$, the axis of the instrument in 3D space is implicitly defined by the parameters $\mathbf{p}$.

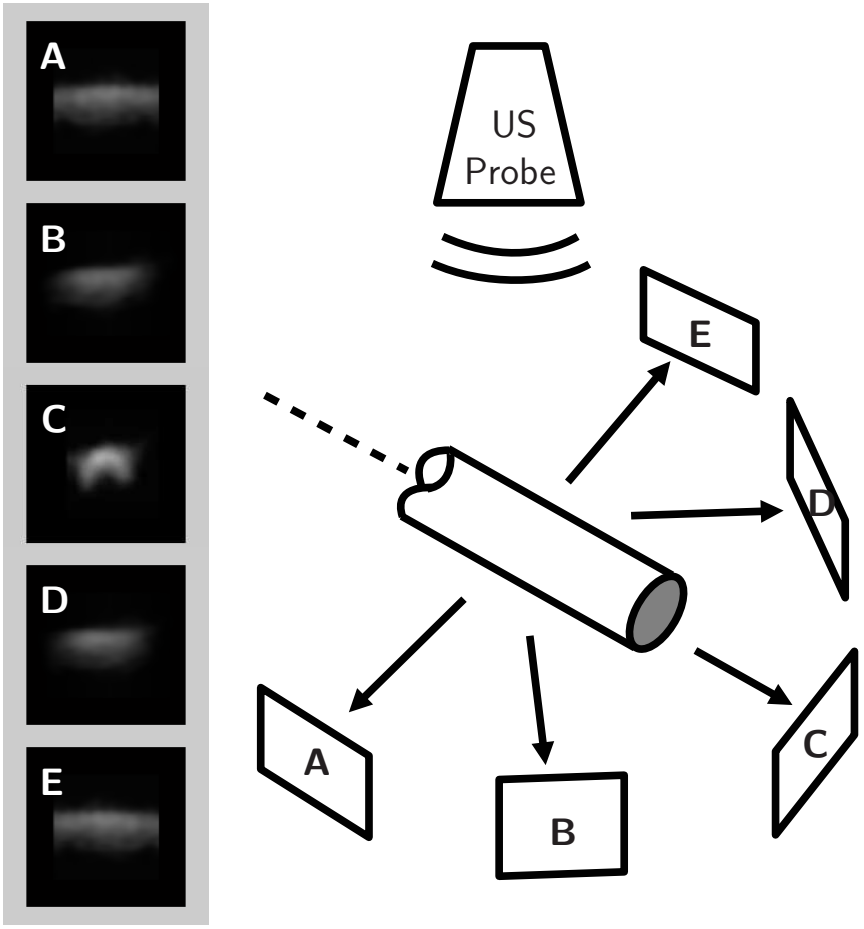


Figure 3.2: Example of the Radon transform detection of the instrument axis. Each image (A-E) is a projection of the ultrasound image along the corresponding direction shown in the schematic. The projection along the axis of the instrument (C) is the brightest. Note that this diagram omits out-of-plane projections that are part of the implementation.

3.2.1 Real-Time Implementation

One of the most promising features of the Radon transform approach described above is the potential for real-time computation. Equation 3.11 can be independently calculated for each $[x_0, y_0, z_0, \theta, \phi, L]$ and is ideally suited for computation on parallel architectures; current personal computer graphics cards are built with features well-suited for this application. Many researchers have shown that highly parallel calculations, when implemented on these graphics processor unit (GPU) cards, show significant performance advantages over CPU based implementations [28, 35, 54]. Using a similar approach, we programed a PC graphics card (7800GTX, nVidia Corp., Santa Clara, CA) to calculate the necessary integrations, $\mathcal{R}_{\mathbf{c}(s,\mathbf{p})}\{I\}(\mathbf{p})$.

Before the algorithm is executed, data is transferred to the GPU: the sequential values of $\mathbf{p}$ to be searched are preloaded into five textures in the GPU memory, where each texture corresponds to each parameter $[x_0, y_0, z_0, \theta, \phi]$. The three-dimensional ultrasound data is loaded onto the graphics card into a three-dimensional texture. Once the data is loaded, the algorithm runs by 'rendering' to an output texture. Sixteen parallel pipelines on the graphics card (programmable pixel shaders) calculate the integral defined in Equation 3.11 by stepping through the volume simultaneously a length of L. The pixel shaders perform the integrations for each input parameter set defined in the input textures, and output the results to the corresponding position in the output texture. Tri-linear interpolation, implemented in hardware, is used as each pixel shader integrates the image intensity in the volume along the direction defined by (θ, ϕ) through the point (x_0, y_0, z_0). The output texture is then transferred back to main memory for post-processing by the CPU. On the CPU, the maximum intensity in the output texture is identified as the instrument axis position.

3.2.2 Elimination of False Positives

It is necessary to eliminate erroneous non-instruments that may produce bright projections, but in fact are not instruments. Large segments of tissue, or even instrument artifacts, introduce areas that could be misidentified. However, the consistent appearance of the cross-section of instruments (Figure 3.2C) is used to eliminate non-instruments in our algorithm. By determining the shape of this cross-section *a priori*, we are able to eliminate candidates that are not consistent with the instrument appearance. In this implementation, a projection along the candidate instrument axis is compared with a template. The normalized cross correlation (NCC) provides a measure of the similarity between the template and the candidate projections

$$NCC = \left(\frac{1}{\sigma_P \sigma_Q}\right)\left(P - \bar{P}\right) \cdot \left(Q - \bar{Q}\right). \tag{3.12}$$

Here P and Q are the vectorized forms of the candidate and template image intensities, with mean $\bar{P}$ and $\bar{Q}$, and variance σ_P and σ_Q. Empirically we found that a

$NCC > 0.85$ was sufficient to distinguish instruments from non-instruments. In addition, this threshold was found to be insensitive to different imaging conditions as it was used throughout the study.

3.2.3 Passive Markers

Once the axis of the instruments is found, it is necessary to detect the final two degrees of freedom of the instrument (tip position and roll angle) to fully define its position and orientation. To this end, we build on work first introduced by Stoll *et al.* [45]. Here we use a new marker design, shown in Figure 3.3A. To produce distinct elements, 800 μm polyurethane foam was wrapped around the instrument shaft. Uncoated metals such as the stainless steel used for surgical instruments are highly reflective in ultrasound. As a result, an 80 μm fiberglass embedded PTFE coating was applied to the instrument in order to improve the appearance.

Finding the markers begins with the image volume already loaded into texture memory from the Radon transform algorithm. Built-in tri-linear interpolation is used to quickly render a slice through the instrument axis with the instrument orientated horizontally (Figure 3.3B). To identify the position of the bumps, a template matching algorithm is used on the ultrasound slice. The algorithm uses the normalized cross correlation (Equation 3.12) between a candidate region of the slice and a template, shown in Figure 3.4A.

The positions of the three best matches found in the slice are used to determine the tip position of the instrument and the roll angle. The tip position is found with a known offset of 3 mm between the two closest markers and the instrument tip. To find the roll angle, the ratio of the distances x_1 and x_2 is used (Figure 3.3A). Since the third marker is wrapped in a helical pattern around the instrument shaft, the roll angle is a linear function of this ratio.

3.2.4 System Overview

For instrument tracking, the ultrasound data is produced by a Sonos 7500 ultrasound system (Philips Medical, Andover, MA). The ultrasound volumes are streamed from the ultrasound machine to a personal computer over a gigibit LAN using TCP/IP. The data stream is captured on the target PC and passed to the instrument tracking algorithm running in separate asynchronous thread. The tracking algorithm calculates the modified Radon transform on the graphics processing unit (7800GTX, nVidia Corp, Santa Clara, CA) using DirectX 9.0c. The entire system runs on a Pentium 4 3 GHz personal computer with 2 GB of RAM.

For the first ultrasound volume, Equation 3.11 is calculated for evenly spaced points throughout the parameter space. Spatially, the volume is sampled at 5 voxel increments in x, y, and z. For angles θ and ϕ, Equation 3.11 is sampled in 10 degree increments. Due to symmetry, the angles are only sampled from 0 to 180 degrees.

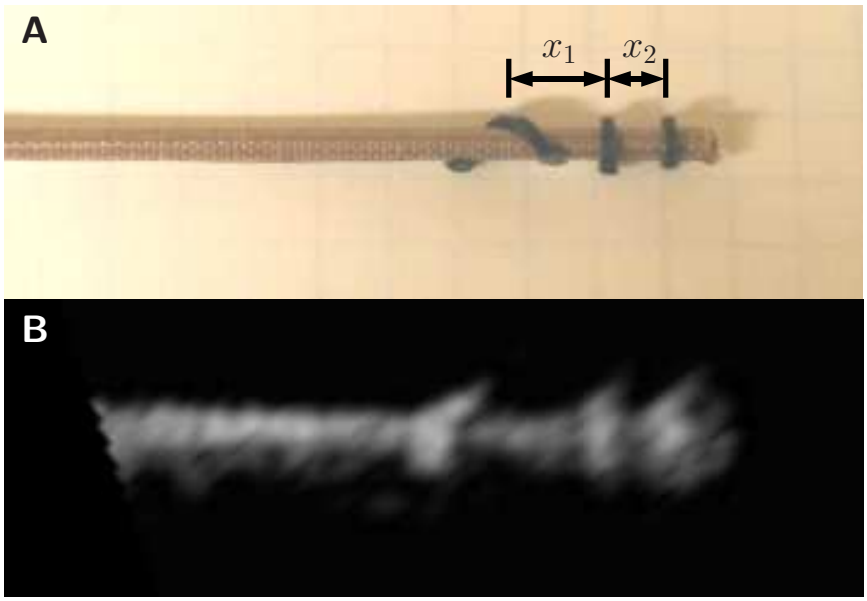


Figure 3.3: Picture (A) and ultrasound image (B) of a minimally invasive anchor driver with passive markers. The instrument tip and roll angle is calculated using the distances x_1 and x_2 .

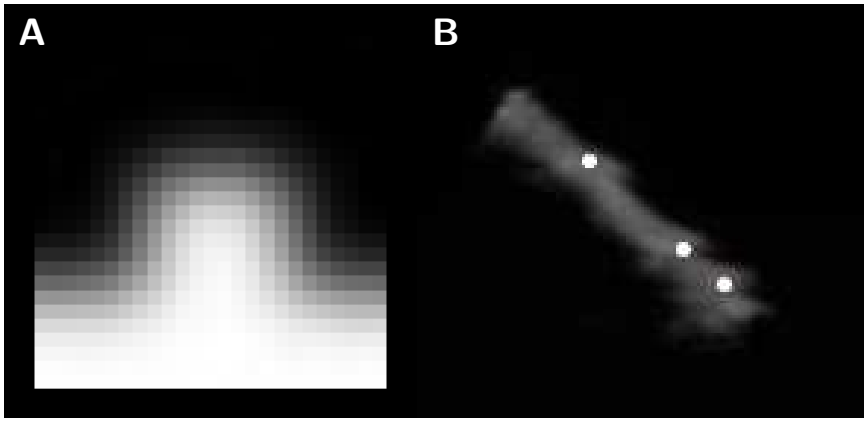


Figure 3.4: (A) Passive marker template used to identify the location of the three markers on the instrument. (B) 3D ultrasound image of a surgical instrument in a water tank. White dots indicate tracked passive markers.

This search constitutes the initialization of the instrument tracking, as the entire volume is searched.

For subsequent frames, the tracking algorithm confines its search space to an area centered on the location found in the previous frame. Since the ultrasound volumes are updated at 25 Hz, this search space can be fairly small. In our trials, we empirically found that limiting the search space to ± 5 voxels spatially in the x, y, and z directions and ± 10 degrees around the angles θ and ϕ found in the previous frame was sufficient to capture typical surgical movements.

3.2.5 Experimental Evaluation

Three sets of experiments were conducted to validate the proposed methodology. First, the Radon based instrument axis tracking was tested in a controlled tank environment. This study examined the effect of different material types on accuracy. Second, the combined Radon and passive marker algorithm was tested with a surgical anchor driver in the same tank trials. With the addition of passive markers, the full six degree of freedom tracking is characterized. While these two studies carefully characterized the accuracy of the method, it does not reflect the target conditions for the algorithm, detecting instruments within a beating heart. As a result, a third study was necessary to validate the technique *in vivo*.

The accuracy of the proposed method was measured with a system that precisely positioned and oriented the instruments within the ultrasound field. The instruments were connected to a three-axis translational stage with 1 μm resolution, and 2 rotational stages with a resolution of 5 minutes. Registration to the ultrasound coordinates was done by using a flat-plate phantom and two-wire phantom [38]. Registration accuracy was 0.4 mm for positions and 0.6 degrees for angular measurements.

Table 3.2: Materials used in tank study.

Material	**Coating**	**Diameter**	**Inside Diameter**
Stainless Steel	-	5.2 mm	4.4 mm
Stainless Steel	tape	5.2 mm	4.4 mm
Stainless Steel	acrylic	5.2 mm	4.4 mm
Acetyl	-	6.3 mm	-
Wood	-	6.3 mm	-

The instruments used in the procedure were chosen to mimic minimally invasive instruments. Multiple material types and coatings were used in order to compare their effects on the appearance of the instruments in ultrasound and identify compatible materials with the tracking algorithm. The instruments consisted of cylindrical rods 5-6 mm in diameter with five different material types and coatings. Shown in Table 3.2,

a stainless steel tube was tested with 3 different coatings: no coating, covered in polymer coated fabric gaffer's tape, and coated in an acrylic based dip (0.55 mm thickness). In addition to the stainless steel, cylinders made from solid acetyl and wood were used.

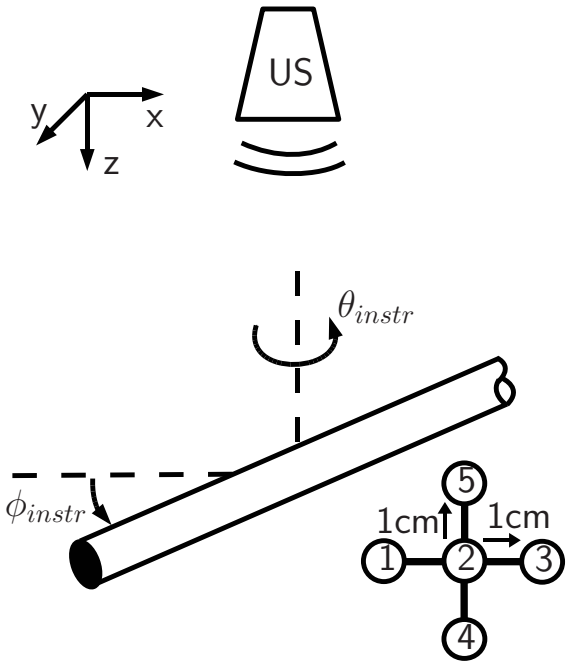


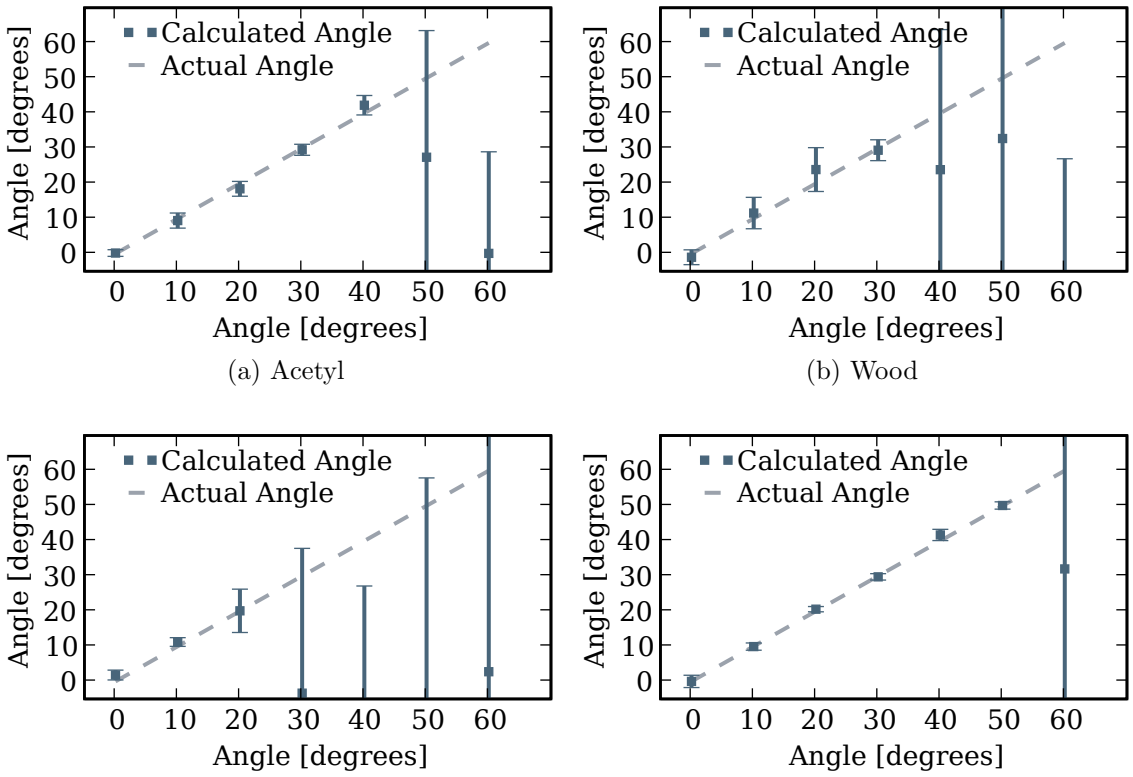
Figure 3.5: In the tank study instruments were imaged at different ϕ_{instr}. Images were taken of the instrument at five positions for each orientation.

A series of ultrasound images were taken by varying the orientation angles of the instrument (ϕ_{instr}). As shown in Figure 3.5, this angle refers to rotations about the y axis of the ultrasound image, identical to ϕ defined in the modified Radon transform (Equation 3.11). ϕ_{instr} ranged from 0 to 60 degrees in 10 degree increments. It was not physically possible to image the instrument beyond 60 degrees because of the size of the field of view and dimensions of the ultrasound probe. This constraint exists in surgical situations also and therefore orientations beyond 60 degrees were not considered in this study. At each angular orientation, the instrument was imaged in five different positions within the ultrasound field. The first position had the instrument tip in the center of the image, and the other four positions were each 1 cm from this initial center position in the axial direction lateral directions (Figure 3.5). As a result, 5 images were taken for each ϕ_{instr}.

For analysis with an actual surgical instrument, a minimally invasive anchor driver was tested (Figure 3.3). This anchor driver consists of a 14 gauge needle through which a push rod forces an anchor out the tip. Passive markers were added to this

instrument to test the combined Radon transform and passive marker six degree of freedom tracking. The accuracy of the hybrid method was measured with the same tank setup described above. The instrument was attached to the three-axis translational and rotation stage and imaged at various positions and orientations. With the new hybrid instrument tracking method it is possible to evaluate the positional error of the tracking algorithm. As a result, the position of the instrument tip was measured with the three-axis stage.

In vivo validation was performed by tracking an instrument within a beating porcine heart. The anchor driver was imaged inside a porcine heart during an open chest beating heart procedure. The instruments were inserted through ports in the left atrial wall and secured by purse-string sutures. The ultrasound probe was positioned epicardially on the left ventricle to give a view of the right and left ventricle. The surgeon was instructed to move the instrument inside the left ventricle.



(a) Acetyl

(b) Wood

(c) Stainless Steel with acrylic coating

(d) Stainless Steel with tape coating

Figure 3.6: Tank results for acetyl (a), wood (b), stainless steel coated in acrylic (c) and stainless steel coated in tape (d) for angle ϕ_{instr}. Errors bar indicate standard deviation.

3.2.6 Results

In the tank trial, the algorithm had varying degrees of success depending on the instrument material and coating. In the best case, shown in Figure 3.6d, tape coated stainless steel was detected accurately for a wide range of angles. For ϕ_{instr} of 0-50 degrees, the mean fell within 0.5 degrees of the measured angle, with a standard deviation of 1 degree. However, beyond 50 degrees the algorithm was unable to accurately find the instrument within the ultrasound image with errors over 20 degrees.

Acetyl, wood, and acrylic-coated stainless steel instruments (Figure 3.6) also were successfully detected, but for a much narrower range of orientations than the tape-coated stainless steel instrument. For acetyl, ϕ_{instr} estimates began to deteriorate for angles greater than 45 degrees, while wood and acrylic coated stainless steel suffered above 30 and 20 degrees, respectively.

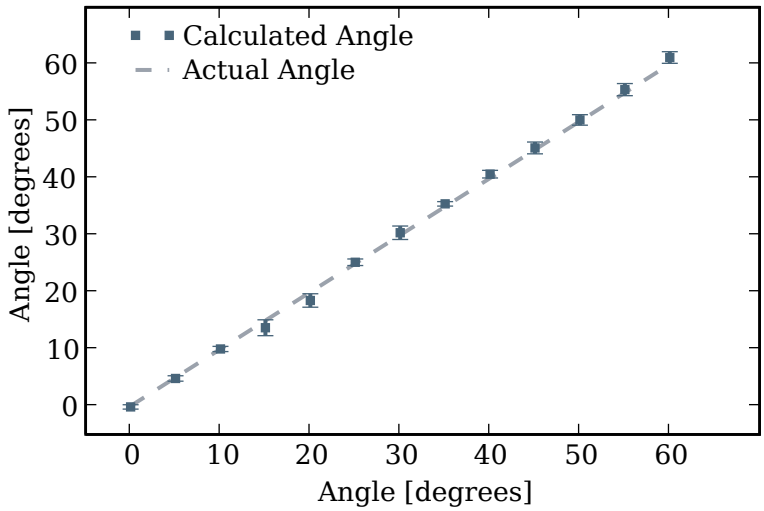


Figure 3.7: The plot shows the mean angle ϕ_{inst} calculated by the tracking algorithm for all 5 position. Errors bar indicate standard deviation.

For the anchor driver, angular accuracy is shown in Figure 3.7, the angular accuracy of the method is shown for different orientations of the instrument with respect to the ultrasound probe. For angles from 0 to 60 degrees of ϕ_{instr}, the instrument tracking algorithm accurately determined its orientation. Across all trials, the RMS difference of the angle calculated by the tracking algorithm and the angle measured by the testing setup was 1.07 degrees. There was no dependence on the accuracy of the algorithm and the orientation of the instrument.

The tip position was also tracked for the surgical anchor driver. The tip position was compared to the actual tip position recorded by the three-axis positioning stage. The results are shown in Figure 3.8 for each orientation angle, where each data point

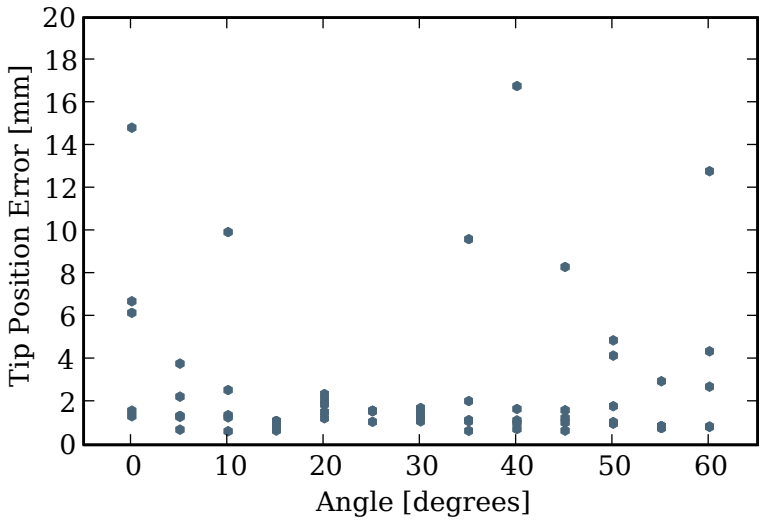


Figure 3.8: Tip position error from tank trials for each angle ϕ_{inst}.

corresponds to one of the five positions of the instrument and orientation angle. It was found that the tip position accuracy was extremely dependent on the accuracy of the passive marker identification algorithm. When the algorithm correctly identified the positions of the three passive markers, the tip position had an RMS error of 1.8 mm. However for eight of the trials, the algorithm mis-identified the marker location. As a result, the tip position error was greater than 5 mm.

Figure 3.9 graphically shows the results of the instrument tracking in a beating porcine heart. The figure shows an overlay of the tracked instrument position on the 3D ultrasound volume over 5 seconds. For each of the images, the overlay correctly matches the position of the instrument as the heart beats around it.

In our experimental setup, the instrument tracking technique required 1.7 s to initially detect the instrument in the entire ultrasound volume. For subsequent tracking, the algorithm required 32 ms per volume. This speed is well within the 40 ms required for the algorithm to keep pace with the 25 volumes per second generated by the ultrasound machine. This performance is a significant improvement over implementation on a CPU. When running the same algorithm on a CPU (Pentium 4, 3 GHz), the algorithm took 11.7 s for the initial detection and 0.54 s for subsequent tracking of the instrument. As a result, the GPU based tracking approach is 14 times faster than a CPU based approach.

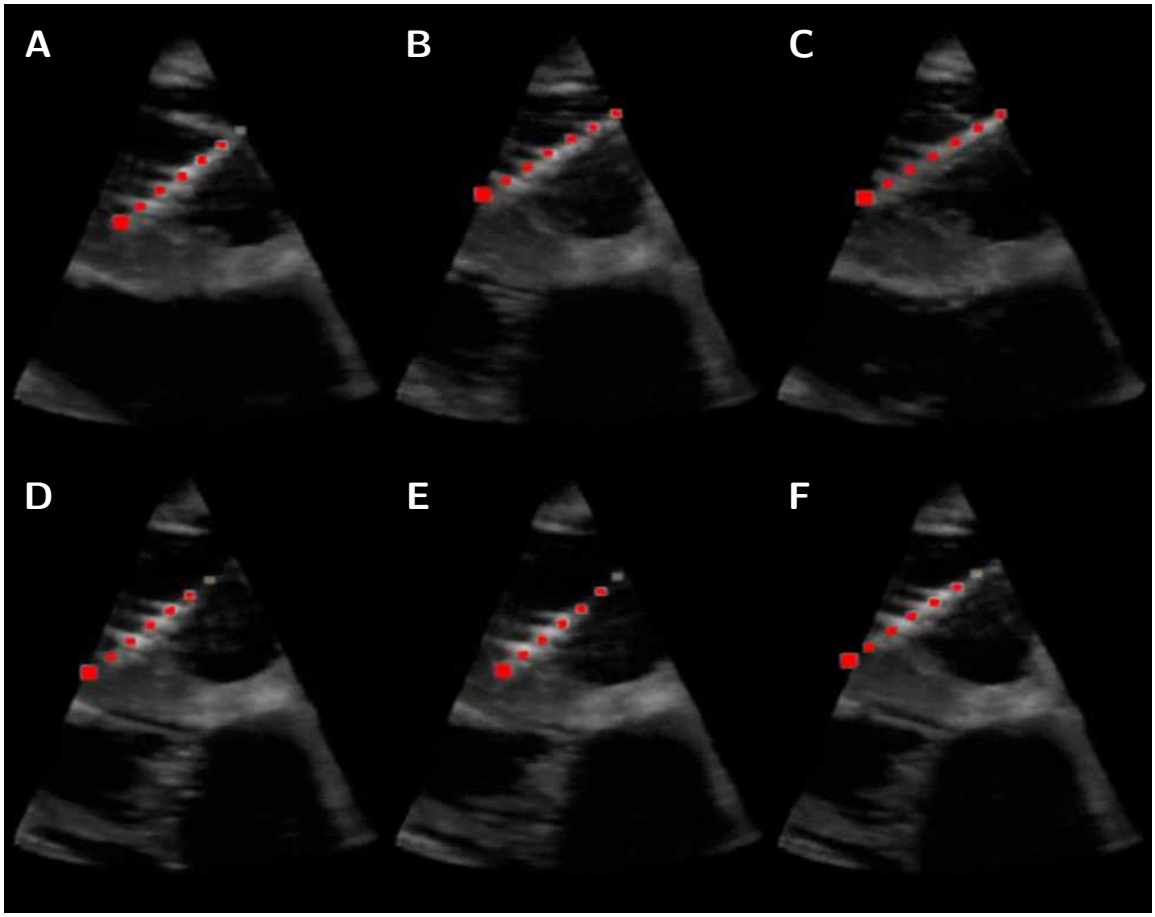


Figure 3.9: Ultrasound images of the instrument inside a beating porcine heart. The red dots indicate the instrument position calculated by the tracking algorithm. Each image (A-F) are images taken each second for 6 seconds.

3.3 Discussion

This work demonstrates for the first time real-time tracking of surgical instruments in intracardiac procedures with 3DUS. The algorithm was both capable of distinguishing instruments from fast moving cardiac structures and efficient enough to work in real-time. The generalized Radon transform is effective here because it integrates over the length of the instrument shaft to minimize the effects of noise and spatial distortion in ultrasound images. By taking advantage of the unique shape of surgical instruments, we were able to correctly distinguish instruments from cardiac structures of similar intensity. The Radon transform was also selected because of its amenability to parallel implementation, which was exploited by using GPUs to achieve real-time performance. As a result, it was possible to detect an instrument

in the ultrasound volume in 32 ms, which is sufficient to handle the 25 volumes per second produced by the ultrasound machine.

The approach demonstrated good performance in both tank trials and an *in vivo* study. These studies accomplished two separate and complimentary goals. The tank trials permitted rigorous exploration of the accuracy of the technique in a controlled environment with precision that is not possible in a surgical setting. The *in vivo* trials, on the other hand, tested the technique's ability to correctly distinguish an instrument from tissue in the presence of highly dynamic aberrations and clutter typical of cardiac ultrasound imaging. Rigorously characterizing accuracy is difficult in this setting, due to the limited range of instrument motion and access to the intracardiac space.

The Radon based instrument tracking highlights the importance of coatings and material selection. Coatings had a large effect on misidentification and accuracy (Figure 3.6). The tape covering on stainless steel produced a near uniform angular scattering of the acoustic energy and therefore was tracked accurately for all orientation angles. For large angles, the other materials did not scatter the acoustic energy and as a result were not tracked. Huang *et al.* [23] confirmed that uncoated metal instruments that were more than 20 degrees from perpendicular to the ultrasound probe reflected almost no ultrasound energy back to the ultrasound probe. They also noted that adding surface coating produces a more diffusive interaction with the ultrasound pulse.

The tank experiments also showed that the tracking algorithm can accurately track a surgical instrument. Across all trials, it correctly identified the orientation of the instrument to within 1 degree. By adding passive markers, the full six degree of freedom tracking was possible. This is of importance to surgical applications because it enables instrument tip tracking. For most cases, the accuracy of the technique was within 1.8 mm. However, as seen in Figure 3.8, the marker detection algorithm sometimes incorrectly identifies the position of the passive markers. As a result, the tip position is incorrectly calculated along the shaft axis. While not the focus of this work, promising improvements to the passive marker detection have been proposed by Stoll *et al.* [45].

With real-time instrument tracking techniques, it is now possible to introduce guidance enhancements to aid in our target procedures, intracardiac surgery. Real-time tracking can now be used for instrument overlays and navigational aids to help the surgeon deal with the distorted appearance of tissue and instruments in 3DUS images.

Chapter 4

Robot Control with 3D Ultrasound

The goal of performing tasks on a beating heart introduces new challenges because of the dynamic nature of the environment. Mitral valve surgery is especially complicated due to rapid movements of the valve during ventricular contraction. Extra caution is necessary because a misstep could result in tearing fragile valve leaflets. To work in this environment, the movement of the mitral valve must be addressed. A surgical robot could be of great assistance under these conditions. For example, a common task in surgical procedures is anchoring an annuloplasty ring to the mitral valve annulus to stop regurgitation. With ultrasound guidance, a surgical robot could perform this task by autonomously moving to a target on the annuloplasty ring and deploying an anchor. The robot could also synchronize its movements to the beating heart and the valve would appear stationary to a surgeon.

While using robotics for beating mitral valve repair is new, there is a history of using robotics in other beating heart procedures. Nakamura *et al.* [33] have shown that it is possible to synchronize a robot with the external heart wall for use in performing a coronary artery bypass graft. By using two cameras, they tracked the outside of the heart and subsequently controlled a robot to follow these movements. Motions made by an operator on a master robot were mapped on top of this synchronization. As a result, it was possible to provide a stationary view of the heart wall for the surgeon.

Other researchers have reported success using visual servoing techniques to control a robot under ultrasound guidance. Early work done by Stoll *et al.* [46] used 2D ultrasound to visually servo a robot end-effector to positions within the ultrasound image. Electromagnetic trackers registered the ultrasound image coordinates to the robot coordinates. A phantom target was identified in the ultrasound by image thresholding and the robot successfully moved to the target position.

Hong *et al.* [21] developed an ultrasound guided robotic system for percutaneous cholecystostomy where a needle is inserted into the gallbladder. The robot was a two degree of freedom robot that was physically constrained to move within the plane of the 2D ultrasound image. A 2D Hough transform identified the needle in the ultrasound image and a deformable contour was used for gallbladder tracking. With

Table 4.1: Description of the coordinate transforms.

$\mathbf{T}_{Instr}^{US}$, $\mathbf{T}_{US}^{Instr}$	Transformation determined by the six degree of freedom instrument tracking in the ultrasound volume.
$\mathbf{T}_{Wrist}^{Instr}$, $\mathbf{T}_{Instr}^{Wrist}$	Constant transformation relating the positional offset of the instrument tip and the robot wrist.
$\mathbf{T}_{Robot}^{Wrist}$, $\mathbf{T}_{Wrist}^{Robot}$	Robot kinematics that determine the position and orientation of the wrist with respect to the robot base frame.
$\mathbf{T}_{US}^{Robot}$, $\mathbf{T}_{Robot}^{US}$	Total transformation from ultrasound coordinates to the robot base coordinates.

known needle and target position, the robot was able to direct the needle to the gallbladder.

More recently, Ortmaier *et al.* [53] tracked the fingers of a endoscopic grasper with 2D ultrasound [53]. They developed a controller that does not require the instrument to lie in the ultrasound plane. In addition, accurate registration of the the robot coordinate frame to the ultrasound coordinate frame is not needed. However, a target location was not tracked in the ultrasound, instead the operator specified the target location in the ultrasound image. All of these techniques rely on 2D ultrasound as feedback for the robot guidance. Registration of the ultrasound frame to the robot frame is obtained through physical constraints [21], external tracking systems [46], or online estimation of the image Jacobian [53].

Recently it has been shown that 3D ultrasound can be used for visual servoing. By registering a three axis robot to a 3D ultrasound probe *a priori*, Pua *et al.* [39] was able to command a robot to move to a manually detected target location in the ultrasound image. Stoll *et al.* [47] performed robot control without *a priori* registration of the ultrasound and robot coordinate system. This work used the 3D ultrasound for full six degree of freedom tracking of the robotic end-effector, and as a result the full registration of the robot and ultrasound coordinates was found. The method was verified by moving the end-effector to a prescribed point in the ultrasound volume. However, tracking speeds were limited to 2 Hz, and the target location was not automatically identified in the ultrasound image.

The work presented here uses 3D ultrasound for visual servoing of a robot for surgical procedures. It builds upon the previous work in 3D ultrasound based visual servoing and addresses its shortcomings. First, the speed of the instrument tracking is addressed by using the fast tracking algorithms presented in Chapter 3. In addition to tracking the instrument as others have done, we describe a technique to use the ultrasound volume to track a target in parallel with the instrument tracking. With both a tracked surgical instrument and a target site, we verify the utility of these tracking techniques by visually servoing a small robot in a water tank.

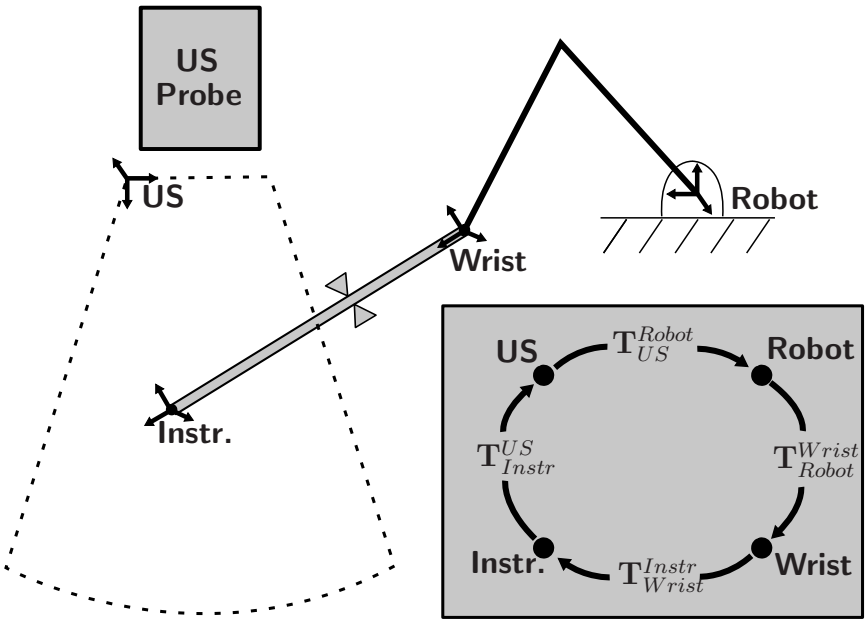


Figure 4.1: Schematic of the ultrasound guided robotic system. The ultrasound probe tracks the position of the surgical instrument that is controlled by a robot. The relationships between the ultrasound image (US), instrument (Instr.), robot wrist (Wrist), and robot base (Robot) coordinate systems are shown.

4.1 System Overview

At the most basic level, controlling a robot requires finding a desired end-effector location or trajectory. When using ultrasound, the desired trajectory or position of the end-effector is often defined with respect to the ultrasound volume. This is especially the case when target locations or anatomy are tracked using the ultrasound images. For a robot to move to, or avoid these locations, their positions in the robot coordinates are needed. This corresponds to a basic coordinate transformation from ultrasound image coordinates to robot coordinates

$$p_{Robot} = \mathbf{T}_{US}^{Robot} p_{US} \tag{4.1}$$

where p is a target location in the ultrasound coordinate frame (p_{US}) and the robot coordinate frame (p_{Robot}). As seen in Figure 4.1, this transformation is broken down into components

$$\mathbf{T}_{US}^{Robot} = \left(\mathbf{T}_{Instr}^{US}\mathbf{T}_{Wrist}^{Instr}\mathbf{T}_{Robot}^{Wrist}\right)^{-1}. \tag{4.2}$$

$\mathbf{T}_{Instr}^{US}$ is calculated using the six degree of freedom instrument tracker described in Chapter 3. The hybrid projection and passive marker technique provides the location and orientation of the instrument in the ultrasound volume. $\mathbf{T}_{Wrist}^{Instr}$ is a

static translational transformation from the tip of the instrument to the connection of the instrument to the robotic wrist. This transformation is determined *a priori* by measuring the instrument length. The final transformation, $\mathbf{T}_{Robot}^{Wrist}$, is found from the robot joint encoders. This transformation describes the transformation from robot base coordinates to the robot wrist coordinate frame.

4.1.1 Estimating Robot and Ultrasound Registration

As was seen in Chapter 3, the instrument tracking is susceptible to mis-identification of the passive markers that caused errors in tracking. However, in the ultrasound guided robotic system (Figure 4.1) we can assume the ultrasound probe is held static or in constant relation to the robot. As a result, while $\mathbf{T}_{Instr}^{US}$ and $\mathbf{T}_{Robot}^{Wrist}$ are varying as the instrument moves, their product $\mathbf{T}_{US}^{Robot}$ remains constant. During surgery, this could be ensured by mechanically fixing the robot and ultrasound probe to the operating table. This assumption allows for filtering erroneous instrument tracking measurements. With the filtered transformation ($\bar{\mathbf{T}}_{US}^{Robot}$), the target location is found

$$p_{Robot} = \bar{\mathbf{T}}_{US}^{Robot} p_{US}. \tag{4.3}$$

To filter the transformation $\mathbf{T}_{US}^{Robot}$, we first convert the homogeneous transformation matrix into a displacement component, and a rotational component represented by a quaternion. Quaternions are a compact representation of a rotation matrix as unit vector [14]

$$\mathbf{q} = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^T \tag{4.4}$$

where $[q_1, q_2, q_3]$ is the axis of rotation, and $q_0 = \cos(\frac{\phi}{2})$, where ϕ is the angle of rotation. Quaternions are an attractive representation of rotations in 3D dimensions because of their compactness and efficiency. For this application, the simplicity of finding an average rotation and the angle between rotations is of principal importance.

The homogeneous transformation matrix $\mathbf{T}_{US}^{Robot}$ is converted to a displacement-quaternion pair, $(\mathbf{t}, \mathbf{q})$, using the following method. Starting with the structure of a homogeneous transformation matrix

$$\mathbf{T} = \begin{bmatrix} & \mathbf{R} & & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.5}$$

the displacement $\mathbf{t}$ is directly extracted from the last column of the matrix such that

$$\mathbf{t} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T \tag{4.6}$$

with t_x, t_y, and t_z representing the x, y, and z translation components of the ho-

mogeneous transformation. The quaternion equivalent to the three by three rotation matrix ($\mathbf{R}$) is found using

$$\begin{array}{rcl} q_0 & = & \pm\frac{1}{2}\sqrt{1+R_{11}+R_{22}+R_{33}} \\ q_1 & = & \frac{1}{4q_0}(R_{23}-R_{32}) \\ q_2 & = & \frac{1}{4q_0}(R_{31}-R_{13}) \\ q_3 & = & \frac{1}{4q_0}(R_{12}-R_{21}) \end{array} \tag{4.7}$$

where R_{ij} corresponds to the i^{th} row and j^{th} column of the rotation matrix. Equivalent forms of Equation 4.7 exist to avoid computation inaccuracies when q_0 is close to zero [14].

As seen in Chapter 3, the instrument tracking algorithm can produce inaccurate outliers of the actual instrument position. It is necessary to develop a technique to first remove outliers before estimating $\mathbf{T}_{US}^{Robot}$. A leader-follower clustering technique [11] is used to group estimates of $\mathbf{T}_{US}^{Robot}$. The algorithm takes each new sample of the transformation $(\mathbf{t}_i, \mathbf{q}_i)$ and compares it to each cluster C found from previous measurements. To determine similarity between the current measurement and a cluster, we use the difference between the current translation $\mathbf{t}_i$ and the average translation of the cluster $\bar{\mathbf{t}}_C$

$$\Delta\mathbf{t}_C(\mathbf{t}_i) = |\mathbf{t}_i - \bar{\mathbf{t}}_C| \tag{4.8}$$

and the angle between the rotation quaternion of the current transformation $\mathbf{q}_i$ and the average quaternion of the cluster $\bar{\mathbf{q}}_C$

$$\Delta\theta_C(\mathbf{q}_i) = 2\arccos\left(\frac{\mathbf{q}_i \cdot \bar{\mathbf{q}}_C}{|\mathbf{q}_i|\,|\bar{\mathbf{q}}_C|}\right). \tag{4.9}$$

The accuracy of the instrument tracking technique was found in Chapter 3. The standard deviation of the positional (σ_t) and orientation (σ_θ) error was 1.8 mm and 1.1 degrees, respectively. If $\Delta\mathbf{t}_C(\mathbf{t}_i) < 3\sigma_t$ and $\Delta\theta_C(\mathbf{q}_i) < 3\sigma_\theta$ for one of the preexisting clusters, the new sample is added to the cluster. If no clusters match, the sample becomes the first member of a new cluster.

For estimating $\mathbf{T}_{US}^{Robot}$, the mean transformation of the cluster with the largest number of elements is used. Finding the average translation component of transformations is straightforward, at is is just he mean of the three components of the translation

$$\bar{\mathbf{t}}_C = \frac{1}{N}\sum_{i\in C}\mathbf{t}_i. \tag{4.10}$$

By using quaternions, the average rotation is easily calculated. Gramkow [16] showed quaternions provide a convenient means of finding an average rotation. In fact the

normalized barycenter

$$\bar{\mathbf{q}}_C = \frac{\sum_{i \in C} \mathbf{q}_i}{\left|\sum_{i \in C} \mathbf{q}_i\right|} \tag{4.11}$$

is a good approximation of the average rotation for rotations within 40 degrees. This criteria is true in this case because the clusters contain measurements that are within $3\sigma_\theta$, about 3 degrees.

The filtered transformation, $\bar{\mathbf{T}}_{US}^{Robot}$, is then calculated from $\bar{\mathbf{t}}_C$ and $\bar{\mathbf{q}}_C$ of the dominant cluster using

$$\bar{\mathbf{T}}_{US}^{Robot} = \begin{bmatrix} 1-2(\bar{q}_2^2+\bar{q}_3^2) & 2(\bar{q}_1\bar{q}_2-\bar{q}_0\bar{q}_3) & 2(\bar{q}_0\bar{q}_2+\bar{q}_1\bar{q}_3) & \bar{t}_x \\ 2(\bar{q}_1\bar{q}_2+\bar{q}_0\bar{q}_3) & 1-2(\bar{q}_1^2+\bar{q}_3^2) & 2(\bar{q}_2\bar{q}_3-\bar{q}_0\bar{q}_1) & \bar{t}_y \\ 2(\bar{q}_1\bar{q}_3-\bar{q}_0\bar{q}_2) & 2(\bar{q}_0\bar{q}_1+\bar{q}_2\bar{q}_3) & 1-2(\bar{q}_1^2+\bar{q}_2^2) & \bar{t}_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.12}$$

This filtered transformation, $\bar{\mathbf{T}}_{US}^{Robot}$, provides a stable estimate of the actual relationship between the ultrasound volume and the robot. In addition, it incorporates a mechanism for discarding erroneous measurements that are introduced by the instrument tracking algorithm.

4.1.2 Target Tracking

Using ultrasound allows for tracking of not only the robot end-effector, but common surgical targets. Tracking targets and the robot end-effector with the same device, ultrasound, has many advantages. The biggest advantage is that no additional sensors are necessary, thereby reducing system complexity and eliminating the need for registration of the systems. Tracking anatomical features in ultrasound is possible, but it is still an open area of research. Instead, the problem is simplified by tracking fiducials on surgical targets. For mitral valve repair, a useful target is the annuloplasty ring which is attached to the annulus during surgery. Tracking this particular target is simplified because the device can be specifically designed for ultrasound based tracking. In this case, we investigated tracking an X, a structure consisting of two crossed strings that is easily added to an annuloplasty ring (Figure 4.2A).

In addition, this fiducial was specifically chosen because detecting two intersecting lines is ideally suited for the modified Radon transform already implemented for instrument axis detection. To identify this marker, it is first necessary to distinguish it from the surgical instrument that is also in the ultrasound image. This is accomplished by identifying the two lines in the ultrasound image that are both intersecting and perpendicular. These conditions ensure that the instrument shaft and target marker are not mis-identified by the tracking algorithm. The intersection of the lines defines the position of the target, and the cross product defines its orientation (Figure 4.2A).

The target tracking algorithm uses the same Radon based algorithm as the in-

strument tracking. As a result, we are again able to take advantage of the parallel architecture found on consumer graphics cards to reach real-time performance.

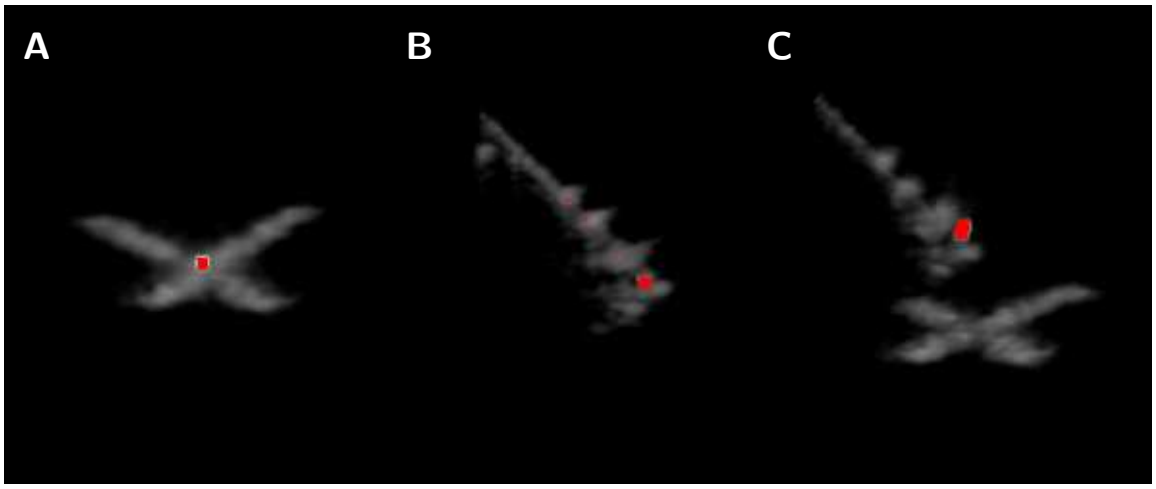


Figure 4.2: 3D Ultrasound images of the target (A), instrument (B), and instrument with target (C). The red dot in (A) indicates the tracked target position. In (B), the red dot indicates the tracked instrument tip position.

4.1.3 Mitral Valve Motion

The robotic system must be tested under conditions mimicking mitral valve surgery. The biggest challenge is handling the fast motions intrinsic to the cardiac cycle. To quantify these conditions, a transthoracic 3DUS image of human mitral valve was obtained using a SONOS 7500 ultrasound machine (Philips Medical, Andover, MA). The motion of the mitral valve throughout the cardiac cycle was captured at 25 volumes per second. For analysis, the mitral valve annulus was manually segmented from the ultrasound volumes. In total, fifty data points were identified around the annulus for each of 20 ultrasound volumes, consisting of one heart beat. The location of these data points was verified by an echocardiologist for accuracy.

For each time point, the 50 data points were spatially averaged to find the centroid of the annulus. Principal component analysis of the centroid's three dimensional position identified the major directions of mitral valve movement. The mitral valve movement is projected onto the first, second, and third principal components in Figure 4.3. The eigenvalues of the covariance matrix (21.75, 5.49, 2.97) shows that 70% of the explained sample variance is captured in the first principal component. This results in only an RMS error of 1.4 mm and max error of 3.1 mm between the one dimensional mitral valve trajectory and the complete three dimensional trajectory.

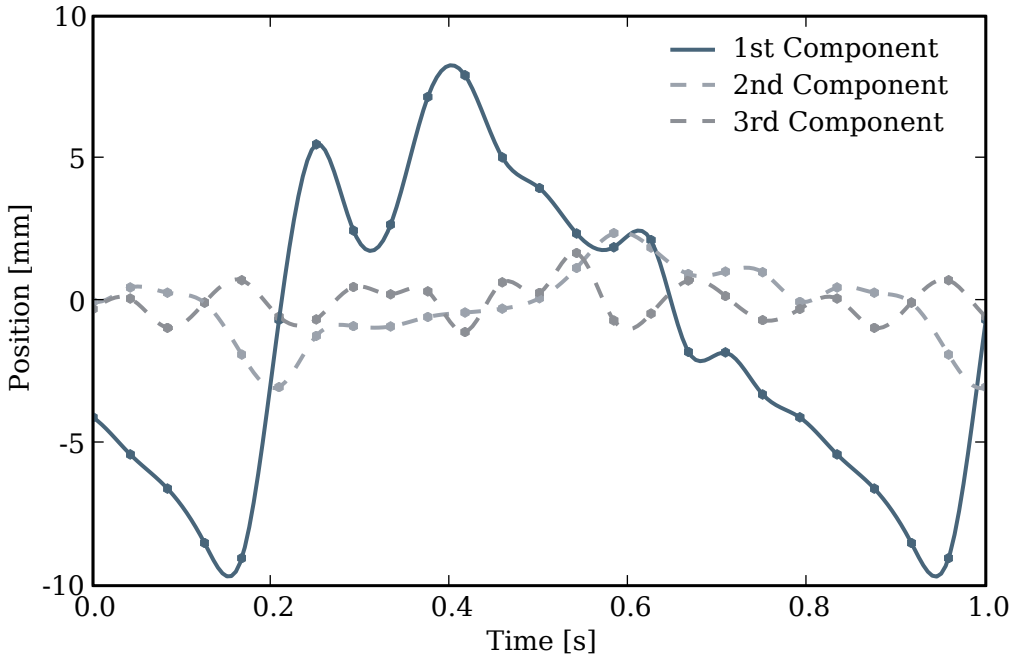


Figure 4.3: Motion of the the mitral valve annulus in a beating human heart. The three curves are the first, second, and third principal components of the mitral valve motion for one cardiac cycle. The first component captures 70% of the movement's sample variance.

To approximate mitral valve motion, only the major one dimensional component is used because it captures the translational motion. Rotational movements are not considered and have been found to be negligible [25]. A one degree of freedom mitral valve annulus was built using a cam and a cam follower. The cam profile produces cyclic one dimensional translation motion of the follower. The follower is connected to the target shown in Figure 4.5. As a result, the target location movement mimics the major component of the mitral valve's translation. Since the cam is connected to a electric motor its speed, or approximate heart rate, can be varied.

4.1.4 Robot Control

With the tracked instrument tip and target location, the robot is controlled to maintain a constant distance from the target. By synchronizing the movements of the robot to the target position, the robot serves a useful function in canceling out the dynamic mitral valve movements. The robot controller uses the current target

position in the robot coordinate frame

$$p_{target} = \bar{\mathbf{T}}_{US}^{Robot} p_{target,US}. \tag{4.13}$$

$p_{target,US}$ is calculated using the target tracking algorithm described in Section 4.1.2. $\bar{\mathbf{T}}_{US}^{Robot}$ is the filtered transformation obtained from the method described in Section 4.1.1.

For robot control, the trajectory is defined in terms of the robot wrist position w_{target}. Due to the kinematics of the port, if the port position is located at p_{port}, the target wrist position is

$$w_{target} = p_{target} + L\left(\frac{p_{port} - p_{target}}{|p_{port} - p_{target}|}\right) \tag{4.14}$$

where L is the length of the instrument shaft.

A PD controller minimizes the error between the target wrist position w_{target} and the robot wrist position in robot coordinates p_{wrist}. The current wrist position p_{wrist} is found using the robot kinematics embedded in the $\mathbf{T}_{Wrist}^{Robot}$ transformation. Given that the position of the wrist is defined as $\begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}^T$ in the robot wrist coordinate frame,

$$p_{wrist} = \mathbf{T}_{Wrist}^{Robot} \begin{bmatrix}0\\0\\0\\1\end{bmatrix}. \tag{4.15}$$

The robot controller calculates the wrist force necessary to move the wrist to the desired location using

$$F = k_p\left(w_{target} - p_{wrist}\right) + k_d\left(0 - \dot{p}_{wrist}\right). \tag{4.16}$$

k_p and k_d are the proportional and derivative gain of the PD controller and $\dot{p}_{wrist}$ is measured by the robot joint encoders. Combining equations 4.13, 4.14, 4.15, and 4.16 yields

$$F = k_p\left(\bar{\mathbf{T}}_{US}^{Robot} p_{target,US} + L\left(\frac{p_{port} - \bar{\mathbf{T}}_{US}^{Robot} p_{target,US}}{\left|p_{port} - \bar{\mathbf{T}}_{US}^{Robot} p_{target,US}\right|}\right) - \mathbf{T}_{Wrist}^{Robot}\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right) - k_d\dot{p}_{wrist}. \tag{4.17}$$

Equation 4.17 calculates the wrist force F in terms of: known constants measured *a priori*, instrument length L and port location p_{port}; the target position tracked with ultrasound $p_{target,US}$; the filtered ultrasound to robot coordinate transformation $\bar{\mathbf{T}}_{US}^{Robot}$; and the robot kinematics $\mathbf{T}_{Wrist}^{Robot}$. k_p and k_d were hand tuned to 0.22 $\frac{N}{mm}$ and

0.005 $\frac{N}{mm/s}$ for high response times without losing stability. The closed-loop frequency response of the underdamped system is shown in Figure 4.4.

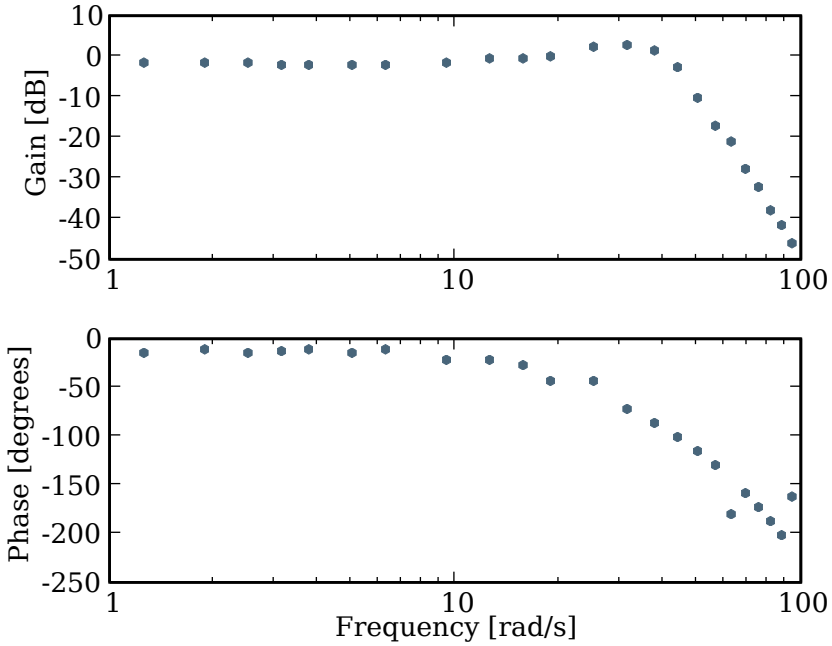


Figure 4.4: Closed-loop frequency response of the robot with PD controller. (-3 dB point is 7 Hz)

4.2 Experimental Evaluation

Visual servoing using 3D ultrasound was tested in a water tank. A surgical instrument was attached to the end of a small robot (Phantom, Sensible Technologies, Woburn, MA). The robot was used to maintain a constant offset from the instrument tip to the target position. The target fiducial was constructed by crossing two lengths of 0.6 mm diameter kevlar string (Figure 4.5B).

The surgical instrument was a custom anchor driver that consists of a 1.5 mm stainless steel tube. Inside the tube is a push rod that inserts the anchors into the target site. The coatings and instrument markers were mounted to the exterior of the instrument. To produce distinct elements, wire was wrapped around the instrument shaft. Uncoated metals such as the stainless steel used for surgical instruments are highly reflective in ultrasound. To ensure that the instrument is visible in ultrasound, a more diffusive interaction with the ultrasound pulse is desired. As a result, electrical heat-shrink tubing (0.30 mm thickness) was applied to the instrument in order to improve the reflectance.

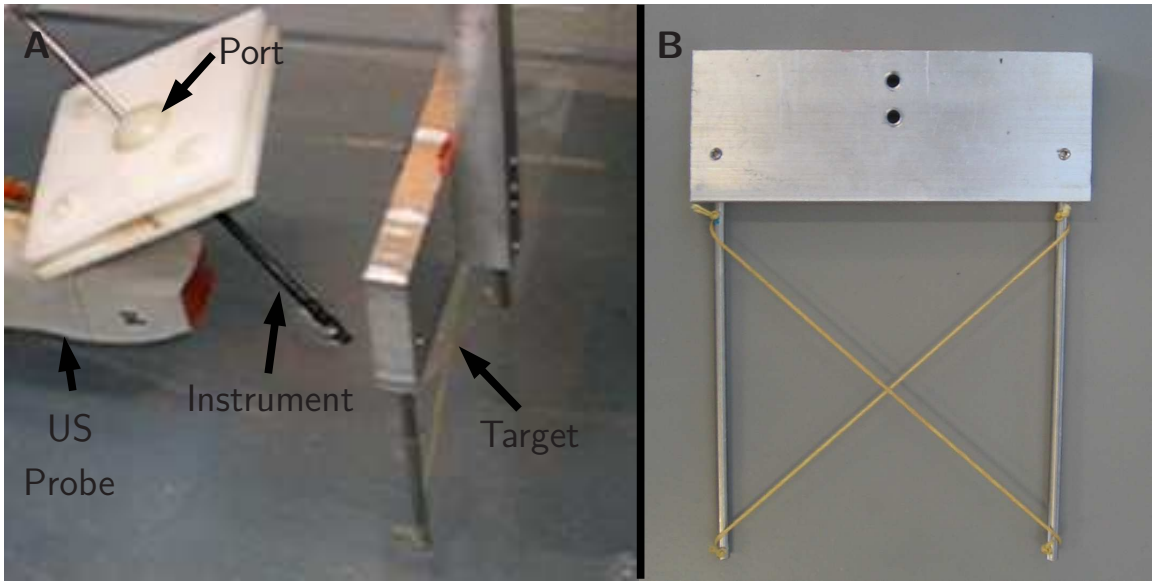


Figure 4.5: (A) Image of the tank trial experiments. The ultrasound probe images a surgical instrument and target submerged in a water tank. The surgical instrument is attached to a robot and passes through a port. During the trials, the target (B) moves left to right, simulating the one-dimensional movements of the mitral valve.

A sliding ball joint approximates an instrument passing through a port in the heart wall. The target and instrument are both imaged by a real-time 3D ultrasound probe in a water tank. (Figure 4.5).

In this setup, the ultrasound data is streamed from the ultrasound machine (SONOS 7500, Philips Medical, Andover, MA) to a personal computer over a gigibit LAN using TCP/IP. The data stream is captured on the target PC and passed to both the instrument tracking and target tracking algorithm running in separate asynchronous threads. Both the tracking algorithms calculate the modified Radon transform on the graphics processing unit (8800GTS, nVidia Corp, Santa Clara, CA). The instrument and target positions are then passed to a third thread that contains the 1 kHz control loop for the robot. The entire system runs on a dual CPU AMD Opteron 285 2.6 GHz personal computer with 4 GB of RAM.

The target position was measured with an angular potentiometer to provide the actual target position. In addition, the robot tip position was recorded using the joint encoders of the robot. Both these measurements were made at 1 kHz and stored to a data file for analysis.

The system's ability to maintain a constant offset from the moving target was tested at different speeds. The target was attached to a cam that mimics mitral valve movements. The cam was run at speeds of 1 Hz, 0.7 Hz and 0.3 Hz, corresponding to

60, 40, and 20 beats per minute. While 40 and 20 beats per minute are not realistic heart beats, they did provide a means of characterizing the system performance.

4.3 Results

Figure 4.6 shows the results of the target tracking algorithm. The tracked target position shows good correspondence to the actual target position. Maximum error of the tracking algorithm was 1.7 mm and the RMS error was 1.1 mm. The computation time was 30 ms for detecting the target in each ultrasound volume. In addition, each ultrasound volume required 30 ms to produce and transfer to the tracking computer. As a result, the target tracking algorithm had a delay of 60 ms from the actual target position.

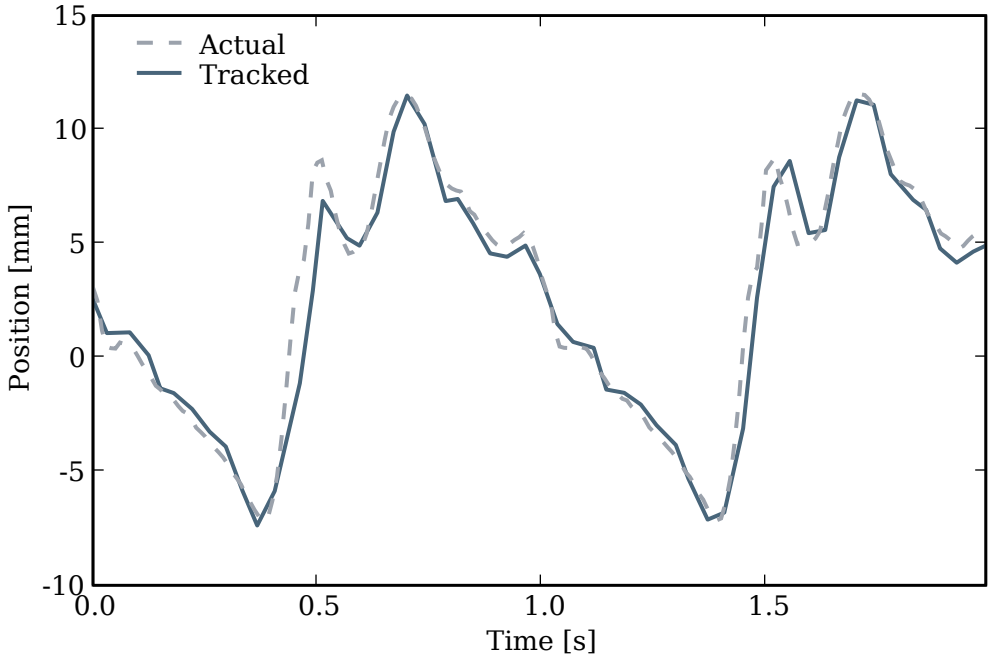
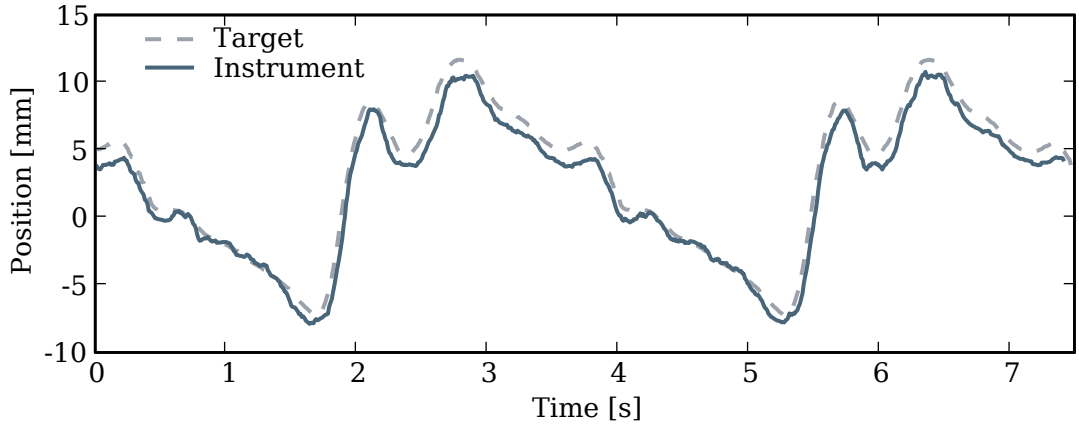
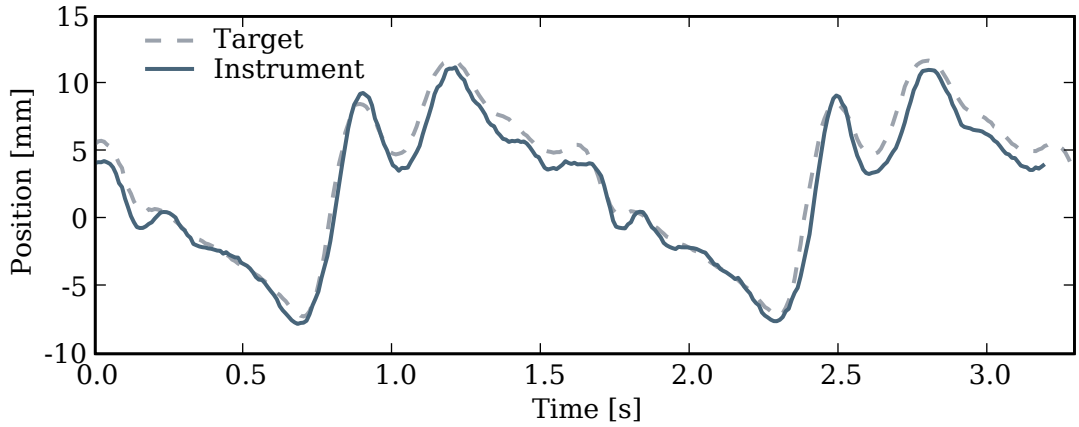


Figure 4.6: Tracking results of the ultrasound target at 60 beats per minute. The dashed line shows the actual target location and the solid line plots the tracked position of the target. Only one dimension of the 3D tracking is shown because the movement is one dimensional. (maximum error = 1.7 mm, RMS error = 1.1 mm)

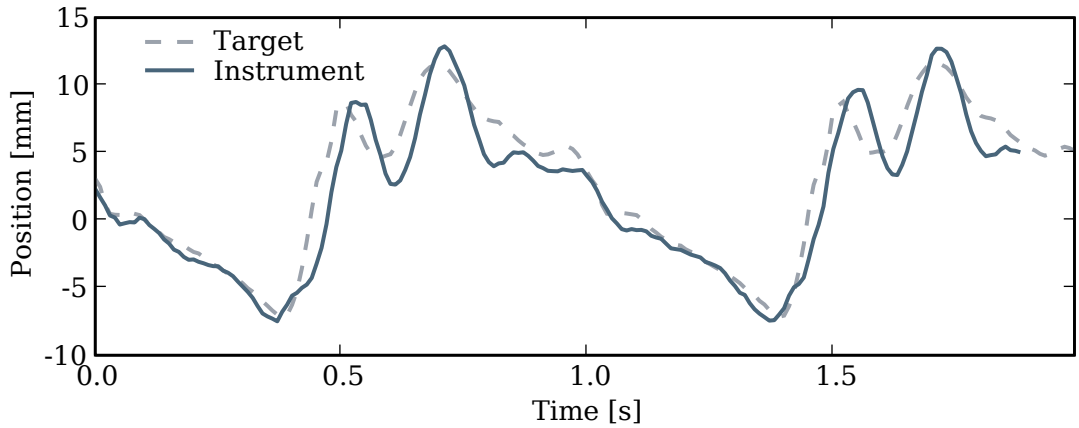
Figure 4.7 shows the results of the robot control task. The ability of the robot to synchronize its movements with the target was dependent on the speed of the target. For 20 beats per minute (Figure 4.7a), the robot was able to follow the



(a) 20 beats per minute (maximum error = 2.2 mm, RMS error = 1.0 mm, robot lag = 100 ms)



(b) 40 beats per minute (maximum error = 2.6 mm, RMS error = 1.0 mm, robot lag = 110 ms)



(c) 60 beats per minute (maximum error = 3.9 mm, RMS error = 1.5 mm, robot lag = 130 ms)

Figure 4.7: Results of the robot synchronization at 20(a), 40(b), and 60(c) beats per minute. Only one dimension of the 3D positions are shown because the movement is one dimensional.

target movements with minimal error. Throughout the cycle, the RMS error of the instrument tip was 1.0 mm. Similar to the target tracking, this error reached a maximum of 2.2 mm during the high speed movement. For the higher speed targets of 40 and 60 beats per minute, the errors increased accordingly. Figure 4.7b shows that for 40 beats per minute the RMS error was 1.0 mm and maximum error was 2.6 mm. 60 beats per minutes incurred a similar increase in error (Figure 4.7c). RMS error was 1.5 mm and maximum error was 3.9 mm. Due to the speed of the robot response and the 60 ms delay in the target tracking, the tip position lagged the actual target position considerably. For 20, 40, and 60 beats per minute the robot tip lagged the target by 100, 110, and 130 ms, respectively.

4.4 Discussion

A technique for using real-time 3D ultrasound for robot control is presented. With the ability to track an instrument and target at real-time speeds, controlling a robot with ultrasound guidance is now possible. We demonstrated that ultrasound based visual servoing can be used to position an anchor driver with respect to a target location. The anchor driver has able to maintain synchronization with a moving target that mimics high speed mitral valve motion. This ability is advantageous when working with dynamic tissue with limited visual and physical access, such as in intracardiac surgery.

Tank trials validated the robots ability to accurately follow cardiac motion for varying speeds. For cardiac motions analogous to a heart beating at 20, 40, and 60 beats per minute, the system accurately followed the desired trajectory. A RMS error of 1.5 mm is comparable to the instrument tracking accuracy of 1.8 mm found in Chapter 3. This error combined with errors introduced by the target tracking algorithm suggest that the inaccuracy of the robotic system is caused by the ultrasound based tracking algorithms.

The largest error in the system is the temporal error. While the robot properly followed the target trajectory, it lagged by up to 130 ms. From Figure 4.7, this lag is 0.10 ms, 0.11 ms, and 0.13 ms for 20, 40, and 60 beats per minute. 60 ms of this lag is attributed to the delay in the target tracking algorithm (Figure 4.6). This delay is comprised of 30 ms time to create the ultrasound volume and transfer the data from the ultrasound machine over a TCP/IP network. In addition, the tracking algorithm takes 30 ms. By the time the target is found in the ultrasound volume, it is 60 ms old. Most of the remaining lag in robot synchronization is caused by the response time of the robot. From Figure 4.4, the phase shift for frequencies between 0.5 Hz and 15 Hz is equivalent to a 30-50 ms response delay.

In order to address this delay, a model of the mitral valve is necessary to predict the target location. Modeling the motion of the external heart wall has been addressed by Nakamura *et al.* [33] and Bebek and Cavusoglu [2]. Their work takes advantage of the

cyclic nature of the heart motion for prediction. In addition Bebek and Cavusoglu [2] used the electrical activity of the heart through the EKG signal to help predict the heart's state. These techniques could be applied to tracking the mitral valve as well. Like the heart wall, the mitral valve's motion is cyclic. In addition, the EKG would be a great use in signaling the onset of left ventricular contraction. This contraction causes the fastest speeds of the mitral valve and results in the largest errors in the robotic system. The EKG would provide a clear indication of ventricular contraction, and the robot controller would act accordingly.

Additionally improvements to the robot are needed. As seen in Figure 4.4, the -3 dB point of the robot is 7 Hz. The 60 beats per minute mitral valve motion used in this study contained frequency components up to 10 Hz. From the frequency response of the robot, it is apparent that these high frequency components are not handled. For ultrasound guided robot control, a suitable robot must have the bandwidth greater than the Nyquist frequency of the ultrasound signal. This ensures that the robot can reproduce movements captured by the 3D ultrasound machine. For 24 Hz 3D ultrasound, a robot with 12 Hz bandwidth is necessary.

A major assumption of this work is that the robot and ultrasound probe do not move with respect to each other during the procedure. This assumption can easily be enforced by holding the ultrasound probe steady with a clamping device. However, this limits the mobility of the ultrasound probe and flexibility in moving the probe for different viewpoints. In addition, often times it is advantageous to press the probe up against the heart wall to obtain optimal intracardiac images. When pressed against a beating heart, the probe will be in constant motion. The current implementation relied upon an internal model where the probe does not move. As a result, the estimate of $\mathbf{T}_{US}^{Robot}$ is the mean of all measured $(\mathbf{T}_{US}^{Robot})_i$. When the probe is moving this is no longer appropriate, instead the system could be extended to include an internal model of the $\mathbf{T}_{US}^{Robot}$ that is updated by each new $(\mathbf{T}_{US}^{Robot})_i$.

This study did not encapsulate all the challenges that are inherent in performing a procedure inside a beating heart. It has been shown that the Radon based tracking algorithm is effective inside a beating heart in previous chapters. However, *in vivo* validation is still necessary for the total system and must be addressed. In actual surgical trials new challenges will undoubtedly arise, such as handling multiple fast moving structures, irregular heart beats, and the physical constraints imposed by the confined access to the heart. Additionally, issues such as biocompatibility must be addressed for both the coated anchor driver and target fiducial. While the target X used in this study was successfully tracked with the Radon based tracking algorithm, it is still an early design. Incorporating this fiducial in a clinically viable annuloplasty ring needs to be addressed.

Chapter 5

Conclusion and Future Work

This work provides tools and techniques to make ultrasound guided cardiac surgery a reality. Image guided interventions in general are an alternative to traditional procedures where surgeons have direct visual and physical contact with the surgical site. As surgery moves away from traditional procedures, interventions become more difficult. When surgeons no longer have physical contact with the tissue, they lose the sense of touch and haptic feedback that helps them perform dexterous manipulations. With the introduction of endoscopic cameras for minimally invasive procedures, surgeons no longer have direct visual contact. Imaging systems such as CT, MRI, fluoroscopy, and ultrasound, when used for guidance, further separate the surgeon from the task at hand. These imaging systems also have a disadvantage due to experience. Surgeons are not as accustomed to using medical imaging as they are natural vision. Compensating for shadows, specular reflections, and occlusions with natural vision is second nature as this is done in everyday life. Medical imaging technologies introduce an entirely new set of artifacts and distortions that are not as familiar as our natural vision systems.

The advantages of image guidance are well documented. Medical imaging reduces invasiveness of procedures considerably: catheter based procedures are supplanting open chest procedures by using angiography and fluoroscopy; MRI is making progress in guiding neurosurgery; ultrasound is being used to guide biopsies in the breast and the prostate. Each of these techniques permits the visualization of structures otherwise inaccessible or that are visualizable only with a greater degree of invasiveness.

The engineering challenge is to bring these systems up to par with optical and direct vision. A rich set of information is needed for the surgery to be done safely, quickly, and repeatably. These imaging technologies often provide a data set with a large amount of information useful for surgery. However, the proper way to process, interpret, and display this information remains an open area of research.

5.1 Ultrasound for Procedure Guidance

In this thesis, ultrasound was examined as a potential imaging technique for cardiac surgery. As with other systems, ultrasound has many disadvantages over direct visual contact. The images are difficult to interpret, low resolution, and have a small field of view. However, there are features that can be taken advantage of to improve performance. First, 3D ultrasound is real-time volumetric data. There is no other medical imaging modality that provides volumetric data at the temporal resolution of ultrasound. In addition, it is non-ionizing, compatible with a wide range of materials and anatomy, and comparatively inexpensive.

As a first step to improve the visualization of 3D ultrasound, stereo displays were investigated in Chapter 2. It was found in tank and *in vivo* studies that displaying the ultrasound volumes on a stereo display has a direct effect on surgical performance. Procedure times and errors decreased with the use of stereo displayed 3D ultrasound. Interpreting ultrasound images is a mentally difficult task and how the data is displayed is an important factor in reducing this mental workload. This thesis demonstrated that stereo displays are an important factor in ultrasound guided procedures.

As a precursor to further improvements to ultrasound visualization, an instrument tracking algorithm was described in Chapter 3. Surgical instruments produce artifacts that further degrade the ultrasound images. By detecting and tracking instruments, this major source of image distortion can be addressed. Knowing the location and orientation of the surgical instrument is arguably the most important piece of information for the surgeon. The main focus of a surgeon is navigating the instrument to a location within the target anatomy. Any misinterpretations of the instrument position lead to costly errors. The instrument tracking algorithm presented here, accurately measured the location and orientation of the instrument in tank studies and *in vivo*. Providing the surgeon with accurate instrument tracking will undoubtedly improve their ability to navigate surgical instruments

5.2 Real-Time Performance

Surgery necessitates real-time performance. To provide interactive information to the surgeon, data processing must occur on the order of minutes, seconds, or even milliseconds. Offline algorithms taking hours to calculate have there place in surgical planning, training, and assessment, but during surgery, algorithm speed is a critical design consideration. In this thesis, performance played a significant role in the design and implementation of the tracking algorithms. The instrument tracking algorithm had to meet a minimum update rate of 25 Hz.

To meet the specification of real-time performance, the software and hardware were considered up front. Currently, GPU performance is being driven by the com-

puter gaming industry. The current generation surpasses modern CPUs in transistor counts, 680 million to 300 million, and floating operations per second, 500 GFLOPS to 50 GFLOPS. However, these GPUs are highly specialized and much less flexible than CPUs. They excel at highly parallel operations but branching (if-then statements) decrease performance significantly. These factors weighed heavily on the design of the tracking algorithm of Chapter 3. The instrument tracking algorithm is posed as a maximization problem that can be solved by any number of optimization algorithms such as Nelder-Mead, simulated annealing, or genetic algorithms. In this thesis, the maximum was found with the traditionally slower 'brute force' method, where the parameter space is exhaustively searched to find the maximum. This was done precisely because of the capabilities of GPUs. Faster optimization algorithms require constant decision making to zero in on the maximum, something that is prohibitive on GPUs. An exhaustive search is decisionless and highly parallel, perfect for the GPU architecture.

Technology is not standing still, and it will not be long before the landscape changes and new hardware offers better performance. Other architectures will become more attractive, and the cost-flexibility-performance calculation will need to be reevaluated. Future real-time algorithms should go through a similar analysis as was done in this thesis to explore algorithmic options in tandem with hardware decisions. This upfront assessment was crucial to the real-time performance successfully achieved in the thesis.

5.3 Ultrasound and Robotics

Robots can play a large part in image guided procedures. While the imaging technology plays a role in visualizing inaccessible sites, robotics allows for greater flexibility in interacting with difficult to reach anatomy. When the invasiveness of the procedure is decreased the room to maneuver is correspondingly decreased. Robots provide a means to work in these cramped spaces. The use of robotics with ultrasound guidance was explored in two studies in this thesis. 3D ultrasound was used with a commercial surgical robot in Chapter 2. Subjects of all training levels were able to complete complex dexterous tasks using the robot and ultrasound. The need for high quality imaging is necessary for surgical robotics as they are used in minimally invasive procedures. Traditionally, endoscopic imaging is used, but incorporating different imaging techniques would extend the functionality of surgical robots. Ultrasound is a perfect tool, in cardiac, fetal, or surgeries within solid organs such as liver, where there is not a sufficient air cavity for endoscopic imaging. In these procedures, the surgical robot provides the flexibility to access these areas, but imaging is the missing piece. We have shown that ultrasound can be integrated easily with commercially available robotics systems for these procedures.

Medical imaging can also provide information to control robots in intelligent ways.

In Chapter 4, one such scenario was explored where the ultrasound provided instrument and target locations to control a robot. In this manner, the robot was able to maintain a constant offset to the target location. Other benefits could be surgical macros, where the robot would perform a small function such as placing an anchor in a target location, or prevent the robot from moving into delicate tissue.

5.4 Future Directions

With the work presented here, there are many promising new directions. While the stereo display was shown to be a significant improvement in Chapter 2, the volume rendering algorithm used in this study was purposely designed without any additional image processing. Additional enhancements to the ultrasound image were excluded in order to isolate the effects of the stereo display. However, a promising next step is to investigate various techniques that would improve a surgeon's ability to interpret ultrasound images. Techniques such as edge enhancement, intensity correction [55], and direct modification of the opacity transfer function used by the volume renderer [22] should be explored.

The instrument tracking algorithm was thoroughly tested in this thesis, although applications are just beginning to be explored. Questions about how best to display this information for surgery still remain. Instrument overlays is an obvious first step and testing its benefit needs to be explored. The benefits of augmented displays also are unverified. Many of these techniques are promising, but their benefit to surgery is unknown.

The robot servoing presented in Chapter 4 provides a significant first step for ultrasound guided robotic procedures. However, the system must be tested under *in vivo* conditions to fully address all the engineering challenges. Cardiac motion modeling and integration with secondary information sources such as EKG could also improve the robotic system's accuracy. However, the design of the robotic system presented here provides the framework to address these issues and make beating heart robotic procedures a reality.

Bibliography

[1] C. Bailey. The surgical treatment of mitral stenosis (mitral commissurotomy). *Diseases of the Chest*, 15:377–397, 1949.

[2] O. Bebek and M.C. Cavusoglu. Intelligent control algorithms for robotic assisted beating heart surgery. *IEEE Transactions on Robotics*, 2007.

[3] D.C. Bellinger, D. Wypij, K.C. Kuban, L.A. Rappaport, P.R. Hickey, G. Wernovsky, R.A. Jonas, and J.W. Newburger. Developmental and neurological status of children at 4 years of age after heart surgery with hypothermic circulatory arrest or low-flow cardiopulmonary bypass. *Circulation*, 100:526–532, 1999.

[4] E. Buffolo, C.S. de Andrade, J.N. Branco, C.A. Teles, L.F. Aguiar, and W.J. Gomes. Coronary artery bypass grafting without cardiopulmonary bypass. *The Annals of Thoracic Surgery*, 61:63–66, 1996.

[5] J.W. Cannon, J.A. Stoll, I.S. Salgo, H.B. Knowles, R.D. Howe, P.E. Dupont, G.R. Marx, and P.J. del Nido. Real-time three-dimensional ultrasound for guiding surgical tasks. *Computer Aided Surgery*, 8:82–90, 2003.

[6] S. Cotin, N. Stylopoulos, M. Ottensmeyer, P. Neumann, D. Rattner, S. Dawson, T. Dohi, and R. Kikinis. Metrics for laparoscopic skills trainers: the weakest link! In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, pages 35–43, 2002.

[7] A. Cuschieri. Minimal access surgery and the future of interventional laparoscopy. *American Journal of Surgery*, 161(3):404–407, 1991.

[8] A.M. Derossis, G.M. Fried, M. Abrahamowicz, H.H. Sigman, J.S. Barkun, and J.L. Meakins. Development of a model for training and evaluation of laparoscopic skills. *American Journal of Surgery*, 175:482–487, 1998.

[9] M. Ding and A. Fenster. A real-time biopsy needle segmentation technique using hough transform. *Medical Physics*, 30(8):2222–33, Aug 2003.

[10] K.J. Draper, C.C. Blake, L. Gowman, D.B. Downey, and A. Fenster. An algorithm for automatic needle localization in ultrasound-guided breast biopsies. *Medical Physics*, 27(8):1971–1979, 08 2000.

[11] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*. Wiley-Interscience, second edition, 2000.

[12] A. Durrani and G. Preminger. Three-dimensional video imaging for endoscopic surgery. *Computers in Biology and Medicine*, 25:237–247, 1995.

[13] V. Falk, D. Mintz, J. Grunenfelder, J.I. Fann, and T.A. Burdon. Influence of three-dimensional vision on surgical telemanipulator performance. *Surgical Endoscopy*, 15:1282–1288, 2001.

[14] J. Funda, R.H. Taylor, and R.P. Paul. On homogeneous transforms, quaternions, and computational efficiency. *IEEE Transactions on Robotics and Automation*, 6(3):382–388, June 1990.

[15] J. Gibbons. Application of a mechanical heart-lung apparatus to cardiac surgery. *Minnesota Medicine*, page 171, 1954.

[16] C. Gramkow. On averaging rotations. *International Journal of Computer Vision*, 42:7–16, April 2001.

[17] E. Grimson. Entering the brain: New tools for precision surgery. *The Bridge: Linking Engineering and Society*, pages 19–22, 2003.

[18] R. Gross, E.J. Watkins, M. Pomeranz, and E. Godsmith. Method for surgical closure of interauriclar septal defects. *Surgery, Gynecology and Obstetrics*, 96:1, 1953.

[19] G.B. Hanna, S.M. Shimi, and A. Cuschieri. Randomised study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy. *The Lancet*, 351(9098):248–251, 1998.

[20] D. Harken, L. Ellis, P. Ware, and L. Norman. The surgical treatment of mitral stenosis. i. valvuloplasty. *New England Journal of Medicine*, 239:801–809, 1948.

[21] J. Hong, T. Dohi, M. Hashizume, K. Konishi, and N. Hata. An ultrasound-driven needle-insertion robot for percutaneous cholesystectomy. *Physics in Medicine and Biology*, 49:441–45, 2004.

[22] D. Honigmann, J. Ruisz, C. Haider, G. Turk, J. van Wijk, and R. Moorhead. Adaptive design of a global opacity transfer function for direct volume rendering of ultrasound data. In *IEEE Visualization*, pages 489–496, 2003.

[23] J. Huang, P.E. Dupont, A. Undurti, J.K. Triedman, and R.O. Cleveland. Producing diffuse ultrasound reflections from medical instruments using a quadratic residue diffuser. *Ultrasound in Medicine and Biology*, 32(5):721–7, May 2006.

[24] J. Huber, N. Stringer, I. Davies, and D. Field. Only stereo information improves performance in surgical tasks. In *Proceedings of the SPIE - The International Society for Optical Engineering*, volume 5372, pages 463–470, 2004.

[25] J. Gorman III, K. Gupta, J. Streicher, R. Gorman, B. Jackson, M. Ratcliffe, D. Bogen, and L. Edmunds Jr. Dynamic three-dimensional imaging of the mitral valve and left ventricle by rapid sonomicrometry array localization. *The Journal of Thoracic and Cardiovascular Surgery*, 112(3):712–726, 1996.

[26] W.S. Kim, S.R. Ellis, M.E. Tyler, B. Hannaford, and L.W. Stark. Quantitive evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. *IEEE Transactions on Systems, Man and Cybernetic*, SMC-17(1):61–72, 1987.

[27] D. Kragic, P. Marayong, M. Li, A.M. Okamura, and G.D. Hager. Human-machine collaborative systems for microsurgical applications. *International Journal of Robotics Research*, 24(9):731–741, 2005.

[28] J. Kruger and R. Westermann. Acceleration techniques for GPU-based volume rendering. In *IEEE Visualization*, pages 287–292, 2003.

[29] D.F. Leotta. An efficient calibration method for freehand 3-d ultrasound imaging systems. *Ultrasound in Medicine and Biology*, 30(7):999–1008, Jul 2004.

[30] F. Lindseth, G.A. Tangen, T. Lango, and J. Bang. Probe calibration for freehand 3-d ultrasound. *Ultrasound in Medicine and Biology*, 29(11):1607–23, Nov 2003.

[31] C.L. Luengo Hendriks, M. Van Ginkel, P.W. Verbeek, and L.J. Van Vliet. The generalized radon transform: Sampling, accuracy and memory considerations. *Pattern Recognition*, 38:2494–2505, 2005.

[32] J.M. Murkin, W.D. Boyd, S. Ganapathy, S.J. Adams, and R.C. Peterson. Beating heart surgery: why expect less central nervous system morbidity? *The Annals of Thoracic Surgery*, 68:1498–1501, 1999.

[33] Y. Nakamura, K. Kishi, and H. Kawakami. Heartbeat synchronization for robotic cardiac surgery. In *Proc. IEEE International Conference on Robotics and Automation (ICRA'01)*, pages 2014–2019, May 2001.

[34] P.M. Novotny, J.W. Cannon, and R.D. Howe. Tool localization in 3d ultrasound images. *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, pages 969–970, 2003.

[35] K.S. Oh and K. Jung. GPU implementation of neural networks. *Pattern Recognition*, 37(6):1311–1314, 06 2004.

[36] S. Park, R.D. Howe, and D.F. Torchiana. Virtual fixtures for robotic cardiac surgery. In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, pages 1419–1420. Springer-Verla, 2001.

[37] K. Peitgen, M.V. Walz, G. Holtmann, and F.W. Eigler. A prospective randomized experimental evaluation of three-dimensional imaging in laparoscopy. *Gastrointestinal Endoscopy*, 44(3):262–267, 1996.

[38] R.W. Prager, R.N. Rohling, A.H. Gee, and L. Berman. Rapid calibration for 3-d freehand ultrasound. *Ultrasound in Medicine and Biology*, 24(6):855–69, Jul 1998.

[39] E.C. Pua, M.P. Fronheiser, J.R. Noble, E.D. Light, P.D. Wolf, D. von Allmen, and S.W. Smith. 3-d ultrasound guidance of surgical robotics: a feasibility study. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 53(11):1999–2008, 2006.

[40] J.D. Puskas, V.H. Thourani, J.J. Marshall, S.J. Dempsey, M.A. Steiner, B.H. Sammons, W.M. Brown 3rd, J.P Gott, W.S. Weintraub, and R.A. Guyton. Clinical outcomes, angiographic patency, and resource utilization in 200 consecutive off-pump coronary bypass patients. *The Annals of Thoracic Surgery*, 71:1477–1484, 2001.

[41] J. Radon. Uber die bestimmung von funktionen durch ihre integralwerte langs gewisser mannigfaltigkeiten. *Berichte Sachsische Akademie der Wissenschaften, Leipzig, Mathematisch-Physikalische Klasse*, 69:262–277, 1917.

[42] R.M. Satava. 3-d vision technology applied to advanced minimally invasive surgery systems. *Surgical Endoscopy*, 7(5):429–431, 1993.

[43] D.J. Scott, P.C. Bergen, R.V. Rege, R. Laycock, S.T. Tesfay, R.J. Valentine, D.M. Euhus, D.R. Jeyarajah, W.M. Thompson, and D.B. Jones DB. Laparoscopic training on bench models: better and more cost effective than operating room experience? *Journal of The American College of Surgeons*, 191:272–283, 2000.

[44] S. W. Smith, G.E. Trahey, and O.T. Von Ramm. Two-dimensional arrays for medical ultrasound. *Ultrasonic Imaging*, 14:213–233, 1992.

[45] J.A. Stoll and P.E. Dupont. Passive markers for ultrasound tracking of surgical instruments. *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'05)*, 8(Pt 2):41–8, 2005.

[46] J.A. Stoll, P.E. Dupont, and R. Howe. Ultrasound-based servoing of manipulators for telesurgery. In *Telemanipulator and Telepresence Technologies VIII Conference, Proceedings of the SPIE*, volume 4570, pages 78–85, Newton, MA, USA, October 2002.

[47] J.A. Stoll, P.M. Novotny, R.D. Howe, and P.E. Dupont. Real-time 3d ultrasound-based servoing of a surgical instrument. In *Proc. IEEE International Conference on Robotics and Automation (ICRA'06)*, Orlando, FL, USA, May 2006.

[48] D.A. Stump, N.A. Kon, A.T. Rogers, and J.W. Hammon. Emboli and neuropsychological outcome following cardiopulmonary bypass. *Echocardiography*, 13:555–558, 1996.

[49] Y. Suematsu, J.F. Martinez, B.K. Wolf, G.R. Marx, J.A. Stoll, P.E. DuPont, R.D. Howe, J.K Triedman, and P.J. del Nido. Three-dimensional echo-guided beating heart surgery without cardiopulmonary bypass: atrial septal defect closure in a swine model. *The Journal of Thoracic and Cardiovascular Surgery*, 130:1348–1357, 2005.

[50] Y. Suematsu, G.R. Marx, J.K. Triedman, T. Mihaljevic, B.N. Mora, S. Takamoto, and P.J. del Nido. Three-dimensional echocardiography-guided atrial septectomy: an experimental study. *The Journal of Thoracic and Cardiovascular Surgery*, 128:53–59, 2004.

[51] N. Taffinder, S.G. Smith, J. Huber, R.C. Russell, and A. Darzi. The effect of a second-generation 3d endoscope on the laparoscopic precision of novices and experienced surgeons. *Surgical Endoscopy*, 13:1087–1092, 1999.

[52] P. van Bergen, W. Kunert, and G.F. Buess. The effect of high-definition imaging on surgical task efficiency in minimally invasive surgery: an experimental comparison between three-dimensional imaging and direct vision through a stereoscopic tem rectoscope. *Surgical Endoscopy*, 14(1):71–74, January 2000.

[53] M.A. Vitrani, G. Morel, and T. Ortmaier. Automatic guidance of a surgical instrument with ultrasound based visual servoing. In *Proc. IEEE International Conference on Robotics and Automation (ICRA'05)*, pages 510–515, Barcelona, Spain, April 2005.

[54] E. Wu, Y. Liu, and X. Liu. An improved study of real-time fluid simulation on gpu. *Computer Animation and Virtual Worlds*, 15(3-4):139–146, 07 2004.

[55] G. Xiao, M. Brady, J.A. Noble, and Y. Zhang. Segmentation of ultrasound b-mode images with intensity inhomogeneity correction. *IEEE Transactions on Medical Imaging*, 21:48–57, 2002.

[56] J. Zeitlhofer, S. Asenbaum, C. Spiss, A. Wimmer, N. Mayr, E. Wolner, and L. Deecke. Central nervous system function after cardiopulmonary bypass. *European Heart Journal*, 14:885–890, 1993.

[57] W. Zhouping, L. Gardi, D.B. Downey, and A. Fenster. Oblique needle segmentation for 3d trus-guided robot-aided transperineal prostate brachytherapy. *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano*, 1:960, 2004.