Application of Augmented Reality to Laparoscopic Surgery

by

Jeremy D. Ackerman

A Dissertation submitted to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering.

Chapel Hill

2002

Approved by:

Henry Fuchs, Ph. D., Advisor

Guido Gerig, Ph. D., Reader

Anthony M. Meyer, M. D., Ph. D., Reader

Benjamin Tsui, Ph. D., Reader

Stephen M. Pizer, Ph. D., Reader

JEREMY D. ACKERMAN
Application of Augmented Reality to Laparoscopic Surgery.
(Under the direction of Henry Fuchs, Ph. D..)

ABSTRACT

The usefulness and feasibility of an augmented reality visualization system for laparoscopic surgery is examined. This technology could enable physicians to see data, such as medical images, as well as the patient's 3-D anatomy, in the context of the patient rather than viewing a 2-D video monitor. Three new results will be presented:

1. A study whose chief findings were that augmented reality improves performance for medical students and that exposure to augmented reality may improve later performance under laparoscopic visualization for both medical students and practicing surgeons.

2. An image processing method that can acquire three-dimensional structure inside the highly specular environment of the abdomen will be shown.

3. A design for a high-speed depth-extracting laparoscope that can be miniaturized and be made operating room compatible will be shown.

Laparoscopic surgery is a minimally invasive abdominal surgery that is performed through small incisions. Surgeons view the procedure on a video monitor that shows a camera's view from inside the abdomen through a small telescope inserted through an incision. Laparoscopy is associated with a loss of natural hand-eye coordination because the work is viewed indirectly on a video screen. Augmented reality may improve performance by restoring natural hand-eye coordination during this type of procedure.

A study of 40 medical students and 8 surgeons was conducted to compare performance of a surgically relevant task under laparoscopic and simulated augmented reality visualization. The study uses *simulated augmented reality*, a technique in which an idealized augmented reality system is created using physical mock-ups, to explore the utility of future augmented reality systems.

A proof-of-concept real-time range acquiring laparoscope has been constructed. The system uses a high-speed implementation of structured light depth extraction. It uses specialized hardware and algorithms. Results from preliminary studies on biological tissues show the effectiveness of this approach.

To the surgeons who hold lives in their hands and the engineers who shape those hands

Acknowledgements

I would like to thank everyone who assisted me in this endeavour.

Amber, Robroy, and Jalapeño for their help, support, encouragement, and most of all love.

Henry Fuchs and members of my committee for their direction and advice.

My parents for their support, love, and proof-reading.

Caoline Green, Gentaro Hirota, Kurtis Keller, Mark Livingston, Andrei State, Greg Welch, Mary Whitton, and many other members of the Department of Computer Science, past and present, for advice and comraderie.

John Morton, A Lucktong, Tim Farrell, and other members of the Department of Surgery for their time, advice, and use of the Laparoscopic Institute of North Carolina's facillities and equipment.

Chris Wiesen, Michael Jiruotek, and Keith Muller for assistance with statistical analysis.

Charles Simms and Ethicon Endosurgery for supplies.

Preface

Minimally invasive surgery and biomedical engineering have traditionally been and continue to be inherently multidisiplinary fields of study[Litynski96]. It is my hope that this dissertation provides both the history, as well as the prospects for even greater advances in the future in this field.

Towards the end of the 19^{th} century the role of surgeons dramatically increased as they were able to do increasingly more with less mortality and pain. The introduction of the medical x-ray in 1896 created the field of radiology wereby physicians were able to get more information about their patient with no apparent harm or discomfort.

In 1966 the television show *Star Trek* aired with its widely disseminated vision of the future. The medical technology in this series has remained relatively constant: a miraculous handheld instrument that measures many different types of bodily function and quickly allows the physician to locate abnormalities, medications for nearly every conceivable condition from radiation poisoning to alien viruses, and finally, amazing surgical instruments that cut, probe, dissect, repair, and regenerate, using small hand-held instruments that leave no scar, no bleeding and allow for a rapid recovery. The Star Fleet doctor's approach is based on the assumption that science will progress.

While modern laparoscopic surgery is a dramatic improvement over previous methods, it is not the vision of the future. Although patient outcomes are generally better for procedures that can be done laparoscopically, there are many limits on what can be done that way. Improved technology will broaden the applicability of laparoscopic and other minimally invasive procedures. These technologies will make the procedures easier to perform and to learn, and improve patient outcomes.

This work will contribute to this vision.

In the course of the research and writing that that make up this work, I have, on several occasions considered whether or not to use quotations to introduce each chapter. My first inclination was not to use any quotations, but later, as I became more

focused on the writing, it became more appealing. For a few weeks I was determined to use quotations from Dr. Seuss – he was the first "doctor" who I identified by that title after all, and his works still provide a great deal of comfort to me. Unfortunately, Dr. Seuss was neither a physician nor a technologist and the relevance of his work to my own was difficult to establish.

At that point I was facing a peculiar problem in writing the section on the history of laparoscopy in Chapter 2. I found that several of the sources disagreed, and not just on minor details. In my quest for closure I happened upon a translation from German of the first known scientific article, written in 1806, to introduce the principle of, as well as a device for, endoscopy. The author, Philip Bozzini, may have published two additional papers earlier, but the only record of them is a reference to the earlier publication in letters. As I read Bozzini's paper, several things struck me at once. The first was that many of the problems Bozzini sought to solve with his device remain problems, although of a different degree, nearly two hundred years later. Just as the problems he sought to solve are similar to those that I address, his motivation, goals, and methods, with the benefit of two hundred years of technical innovation, were strikingly similar to mine.

The following day I placed a quotation from that paper at the beginning of each chapter. I found that this helped to make smoother transitions between the ideas but, more importantly, it helped me focus my writing. I conclude the preface with the quotation used to introduce Chapter 4, because it best sums up my hopes for this entire work:

> *"Surgery will not only develope new and previously impossible procedures, but all uncertain operations which depended on luck and approximation will become safe under the influence of direct vision, since the surgeon's hand will now be guided by his eyes"*
>
> *– Philip Bozzini, 1806*

Contents

List of Tables

List of Figures

Chapter 1

Overview

"An Invention to Visualize the Internal Parts of the Body and their Diseases, with Illustrations"

– Philip Bozzini, 1806

In this dissertation I examine the usefulness and buildability of an augmented reality visualization system for laparoscopic surgery.

Laparoscopy is the endoscopic examination of the abdominal, or, more precisely, peritoneal cavity. A modern laparoscope is a rigid device with an optical channel which allows a view into the peritoneal cavity. A camera is attached to the exterior end of the laparoscope to allow recording and more comfortable viewing of the procedure. The optical channel is usually surrounded by a second channel which transmits the light that illuminates the peritoneal cavity. The light is supplied by fiber optic cable from a high intensity light source.

The laparoscope and other surgical instruments are introduced into the peritoneal cavity through sleeves called trocars. These sleeves are 2-14mm in diameter. Smaller instruments, those less than 4mm, may be introduced without a trocar. An insufflator is attached to a trocar, often the one holding the laparoscope, and it creates and sustains a pneumoperitonium, a gas filled peritonium. This inflation creates enough room for the surgeon to work. When the intervention is complete, the pneumoperitonium is released and the instruments and trocars are removed. The larger incisions may only need a few sutures.

Laparoscopic surgery is part of the growing field of minimally invasive surgery. The goal of minimally invasive surgery is to perform the surgical intervention with as little damage to other structures as possible. These procedures typically result in faster recovery for the patient but present greater difficulty for the surgeon. Chapter 2 contains a detailed description of laparoscopic surgery, its history, its instrumentation, its risks, and its benefits.

Early results from computer graphics demonstrations showed the potential of augmented reality in medical applications. Augmented reality technology allows users to see computer generated imagery merged with the real world.

Augmented reality systems are built around some type of display. Some displays include video and see-through head-mounted displays, flat panel displays, CRT monitors, and projective displays. Hybrid systems using multiple types of these displays are possible and may offer some advantages to using a single type of display.

These displays in general, and head-mounted displays in particular, can be broadly divided into two types: optical see-through and video see-through. Optical see-through displays allow the user to view the real world either directly or through some optical device. Augmentation is achieved by optically mixing the computer generated video with the view of real objects using either a prism or half-silvered mirror system or by directly projecting onto real objects. In video see-through systems, the user sees a view of the real world acquired via one or more cameras. Synthetic imagery is combined with the camera's view on the host computer or with a separate video overlay system.

Video see-through systems offer the ability of synthetic imagery to fully occlude real objects, while optical see-through systems allow higher resolution viewing of the real world. Optical see-through systems present greater challenges for callibration and registration.

There have been a large number of proposed uses for augmented reality. In some applications the user would like to see how a virtual object, perhaps a proposed building, would look in its natural setting. In other applications, the user's view is supplemented with labels or technical information, perhaps guide repairs of a complex system. Another application is to allow the user to see real things that normally cannot be seen, perhaps allowing a view through a wall or inside of a surgical patient.

Augmented reality can provide a more intuitive way to understand and observe data that can be shown in other ways. In the first example, architects might build models, and draw or render pictures. In the second example the technician might be given a paper manual or a computer with technical documents. In the final examples, images could be displayed on monitors.

While augmented reality has excellent intuitive appeal, there is little evidence demonstrating improved performance using these systems compared with more conventional methods of conveying information. An overview of current augmented reality displays and systems is given in Chapter 3.

For medical applications augmented reality allows physicians to see data, such as medical images, in the context of the patient rather than viewing films or a video monitor. Several demonstration systems have shown that augmented reality can look compelling. A strong anecdotal case can be made that this type of visualization improves performance by restoring natural hand-eye coordination. A summary of current developments in medical augmented reality is given in Chapter 4. These results motivate further application of augmented reality to other medical problems but also raise two important questions: is augmented reality for laparoscopy useful, and is it achievable?

The question of utility is addressed using *simulated augmented reality*, a technique in which an idealized augmented reality system is created using physical mock-ups. This environment provides a modality to test hypotheses about the utility of the augmented reality visualization. Methods to realistically simulate augmented reality using a physical mock-up are discussed in Chapter 5.

A study of 40 medical students and 8 surgeons using simulated augmented reality was conducted to compare performance of a surgically relevant task under laparoscopic and simulated augmented reality visualization. Based on a review of tasks used to test performance in laparoscopy, a modified task was developed to test the effect of the simulated augmented reality visualization on surgical performance. This task is different from other tasks in that accuracy rather than time-to-completion is the primary outcome measure. Development of this task is discussed in Chapter 6.

The chief finding from the study was that simulated augmented reality improves performance in medical students, but not in trained surgeons. Important secondary findings were that exposure to augmented reality may improve later performance under laparoscopic visualization and that degrading the augmented reality view significantly reduces the performance enhancement. These results are presented in Chapters 6 and 7.

The major barrier to building an augmented reality system to assist laparoscopic surgery is having a way to acquire the internal geometry of the surfaces of the patient's organs intraoperatively in real-time. Background about three-dimensional acquisition methods including detail of structured-light based triangulation methods are discussed in Chapter 8. Structured light methods appear promising because they can be implemented in real-time and they can be scaled to the small size needed for laparoscopic surgery. The real-time laparoscopic environment poses a number of difficult technical challenges for image processing and hardware implementation.

A proof-of-concept real-time range acquiring laparoscope was constructed. This system utilizes high speed structured light and uses specialized hardware and algorithms. Results from preliminary studies on biological tissues show the effectiveness of this approach. These image-processing methods are discussed in Chapter 10.

The constructed prototype has a number of significant failings as a result of being naively constructed from available components. This system has helped to clarify the design issues in building the hardware components of a real-time laparoscopic depth-extraction system. These problems and proposed solutions are discussed in Chapter 9.

The results of this research demonstrate the potential usefulness of augmented reality for laparoscopy. The long term potential of this technology is to allow surgeons an intuitive way to combine multiple data sources intraoperatively. This may give rise to new and improved surgeries in the future. Future developments, including those in simulated augmented reality as a test-bed for future augmented reality systems, laparoscopic augmented reality systems, and hardware components are discussed in Chapter 11.

Chapter 2

Laparoscopy

"Until now, we were unable to look into the internal cavities and spaces of the living animal body"

– Philip Bozzini, 1806

This chapter will provide a brief description of laparoscopic surgery, its history, laparoscopic technique and instrumentation, and a brief review of the risks and benefits associated with laparoscopic surgery. Developing directions within laparoscopic surgery will also be discussed.

2.1 Overview

Laparoscopic surgery is a type of minimally invasive surgery or minimal access surgery. Minimally invasive surgery is an approach to surgical procedures in which healthy tissues are spared to facilitate rapid patient recovery. While many minimally invasive procedures, including laparoscopy, are at times associated with increased risk of complications as compared to their traditional counterparts, most patients experience faster recovery, shorter hospital stays, and less pain. These procedures place greater demands on the surgeons and require specialized training because the same procedure is completed with less working space, less information, and awkward visualization.

Laparoscopy is broadly defined as the use of a rigid endoscope to perform an examination of or intervention on abdominal organs. Laparoscopic surgery is also commonly known as key-hole or "band-aid" surgeryi because of the small size of the incisions used. These incisions are closed using with a few sutures or even just a band-aid. The laparoscope, an endoscope designed for these procedures, is inserted into the abdomen through a small incision. Most laparoscope procedures are now performed by watching a video monitor displaying the view of a small camera mounted on the laparoscope rather than by directly looking through the laparoscope.

During a laparoscopic procedure the abdomen is usually inflated (insufflated) producing a gas-filled space in the abdomen called pneumoperitoneum. This space is needed to move the laparoscope and the instruments during the procedure. Gasless laparoscopic methods have also been introduced to create a working space without inflating the abdomen[1].

The laparoscope and other instruments are usually inserted into the abdomen through a sleeve-like device called a trocar. Trocars usually contain a system of seals so that only a small amount of gas leakage from the abdomen can occur and pneumoperitoneum is maintained.

Instruments designed for laparoscopy usually feature a tiny tool on the tip of a long, thin shaft. The length of the shaft allows the surgeon to reach a large area inside the abdomen through a single, small incision. A wide variety of laparoscopic instruments including, grasping, cutting, stapling, suturing, and imaging devices have been created to facilitate these procedures.

A more detailed description of steps common to all laparoscopic procedures appears in Section 2.4. Descriptions and images of laparoscopic instruments appear in Section 2.3.

2.2 History[2]

Laparoscopic surgery, as it is currently performed, is a fairly recent development. However, the basic principles and methods used in modern laparoscopic surgery date back over one hundred years, and the general notion of examination of abdominal organs without making a large incision dates back over a thousand years.

Despite the relative newness of the field, there is a great deal of confusion, if not controversy, about who, when, and where key developments were made. Much of this confusion is related to parallel and perhaps independent development of similar methods and equipment.

[1] "Gasless" laparoscopy is not widely used because it can be difficult to use and gives inferior exposure of the operative site[Lukban00, Curet00], but it is an available technique. These methods use mechanical retraction to create a neutral-pressure space inside the peritoneum rather than CO_2 insufflation and may be useful in patients with high risk of cardiovascular complications from pneumoperitoneum[Galizia01].

[2] The material in this section is derived from many sources that are not explicitly referenced within this text. The chief among these are [Litynski96] which is by far the most complete and well researched. Other general references I consulted include [Spaner97] and [Semm95].

The earliest historical evidence of endoscopy comes from excavation of Pompeii (buried in an eruption in 79 A.D.). Archeologists found "specula" thought to have been used for vaginal, rectal, nasal, and otic examinations. A reference that is likely to have been contemporary occurs in the Talmud of Babylon (collated and recorded during the 4^{th} century, but recorded largely from 1^{st} century scholars) and refers to a lead pipe used to visualize the cervix via the vagina.

The Arabian physician Albukasim (reported to have lived between either 936-1013 A.D. or 980-1037 A.D.) is credited as being first to examine an internal organ (the cervix) using reflected light. Giulio Cesare Aranzi (1530-1589) followed later using mirrors and a camera obscura.

In 1806 Philip Bozzini (1773-1809) published an extensive description of the design and use of a candle-lit tubular instrument designed to examine inside the human body. He named this device *liechlieter* or "light conductor". While it is not clear that he used this device himself, experiments in both live subjects and cadavers were conducted in 1807 and included the visualization of stones and tumors in the urinary bladder. The Vienna Medical Faculty discounted the invention as "a mere toy" and experiments with the device were discontinued[Bush74]. The 1806 paper begins with a vision of a wide range of uses Bozzini saw for his invention, including many of the current uses for endoscopes[Rathert74].

Desormeaux introduced an instrument in 1853, similar to devices previously introduced by others. This was an open tube cystoscope with no optical system for magnifying the view. The devices burned alcohol and turpentine as a light source and used a lens to help focus the light. In 1868 Adolf Kussmaul (1822-1902) used a device of this type to attempt, unsuccessfully, an esophagoscopy in a professional sword swallower (who presumably could tolerate having a long tube thrust down his esophagus).

In 1876 Maximilian Nitze (1848-1906) developed an endoscope which closely resembled the modern laparoscope. This endoscope used a water-cooled electric light source. Most importantly, it contained an optical system to magnify the view and improve optical efficency. This device was extensively modified for various types of examinations including those of the stomach, esophagus, urethra, and bladder.

In 1901 George Kelling(1866-1945) used a Nitze-type cystoscope that he built to perform the first modern laparoscopy. He had been experimenting with the use of "Lufttamponade" – using high pressure insufflation of gas to stop bleeding in the abdominal cavity. Kelling wanted to directly observe the effect of Lufttamponade

on internal organs and introduced a Nitze-type cystoscope through the abdominal wall. This method of creating a gas-filled space in the abdomen proved not to be very successful for controlling bleeding because of the high mortality at the pressures needed to stop bleeding. Using this technique he was able to observe abdominal organs in dogs. Kelling called his technique "coelioscopy", but did not attempt it in humans. Kelling's Lufttamponade is now called pneumoperitoneum and significantly lower insufflation pressures are used.

Hans Christian Jakobaeus (1879-1937) is usually credited with popularizing laparoscopy in humans and coining the term "laparoscopy". He published an article "Über Laparo und Thorakoskopie" ("On Laparo and Thoracoscopy") describing 109 procedures in 69 patients in 1912. He described the appearance of a variety of conditions including tuberculous peritonitis, cirrhosis, and metastatic cancer. Jakobaeus also described difficulties and complications of laparoscopic examination, and preferred to perform the procedure in patients with accumulation of fluid in their abdomen (ascites), which he replaced with air to create pneumoperitoneum, because of the risk of bowel injury. In addition to examining the abdomen, Jakobaeus also examined the pleural cavity (the cavity in the chest containing the lungs) during induced pneumothorax. Both pneumothorax and pneumoperitoneum were widespread treatments for disseminated tuberculosis in the chest and abdomen respectively. Pneumoperitoneum and pneumothorax were induced with air, oxygen, or nitrous-oxide.

In 1924 Zollikofer proposed the use of carbon dioxide for generating pneumoperitoneum. Carbon dioxide is more rapidly reabsorbed from the peritoneal cavity thereby reducing both post-operative discomfort from residual gas in the peritoneum and risk of gas embolism. Carbon dioxide also reduced thermal complications due to heating from the light source, and compared to use of room air, nitrous oxide, or oxygen, reduced the risk of fire.

In 1918 Otto Goetze invented an automatic needle for increasing the safety of creating pneumoperitoneum in 1918. His contribution was overshadowed by that of Janos Veress (also referred to in the literature as Jacob, Veres, Verres or Verré in part due to variations in spelling he used for his own publications), a Hungarian physician, who published a description of his own needle (first used in 1932) in 1938[Veress38, Bridgewater99]. Veress's primary application of the needle was drainage of abdominal fluid (ascites) and creation of therapeutic pneumothorax to treat tuberculosis. This device, with only a number of small improvements, became and remains, an important part of methods to create pneumoperitoneum. It is dis-

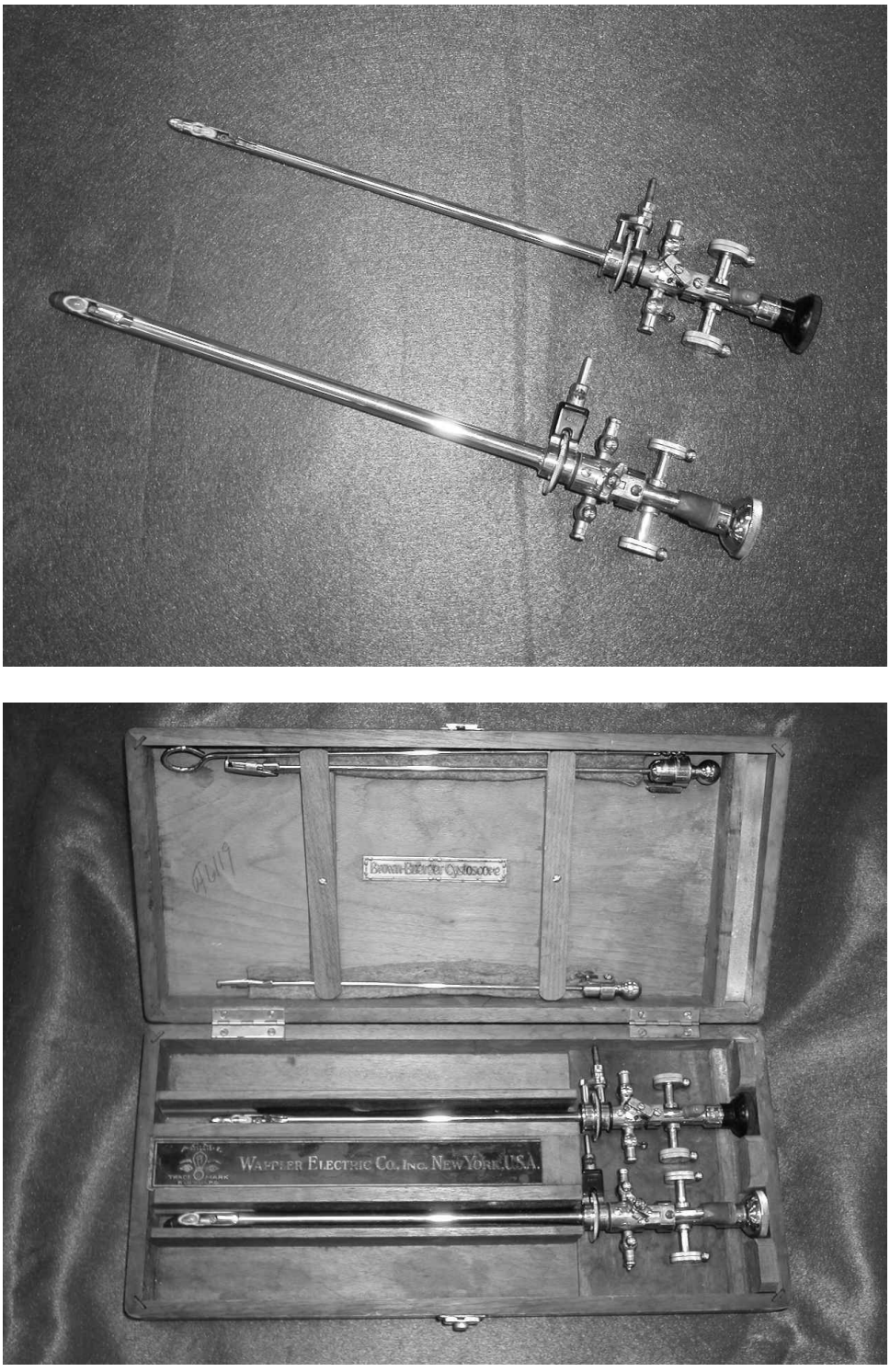


Figure 2.1: An image of two cystoscopes (circa 1910) from the UNC historical instruments collection. Instruments similar to these ones were used for early laparoscopies. Courtesy of Health Sciences Library, University of North Carolina at Chapel Hill.

cussed in more detail in Sections 2.3.2 and 2.4.1.1, and in Figure 2.5.

It should be noted that much of the early work on laparoscopy, particularly in Europe, was done by internists rather than gynecologists or surgeons. Early applications of laparoscopy by internists include treatment of tuberculous peritonitis via pneumoperitoneum, and laparotomy for distinguishing mechanisms of liver disease. Laparoscopy was widely employed by gastroenterologists from approximately 1950 to the 1970s when it was largely supplanted by x-ray and ultrasound imaging technologies. Later flexible endoscopy became the diagnostic procedure of choice.

Laparoscopic techniques were used for gynecological procedures, most notably tubal sterilization described by Power and Barnes in 1941 and later by Raoul Palmer in 1962. Demand for this procedure led to widespread education of gynecologists to perform laparoscopic examinations and interventions. For this reason gynecologists performed the largest portion of interventional laparoscopy between 1960 and 1980 and much of the instrumentation and technique for modern laparoscopy was developed by gynecologists. Kurt Semm developed an automatic carbon dioxide insufflator in 1967 which led to wider acceptance of laparoscopic pelvic examination due to the decreased risk of excessive pressure during pneumoperitoneum. Kurt Semm also invented a large array of devices that now form the basis of the tools used for laparoscopic surgery.

Another significant improvement in laparoscope design was the introduction of glass and crystal rod optics by Harold Hopkins in 1960. The first devices using this optical system were introduced by Karl Storz in 1966. This dramatically improved the optical efficiency of laparoscopes.

Video cameras were occasionally used to record laparoscopies, but they were used infrequently because of the difficulty associated with positioning a laparoscope with a large and bulky video camera attached. In the early 1980s CCDs[3] allowed for the miniaturization of video cameras. Introduction of the video laparoscope allowed multiple people to view the procedure simultaneously (in contrast to only using an optical device).

The first laparoscopic appendectomy was performed by Semm, a gynecologist, in 1980. Gynecologists began doing elective appendectomies regularly during other laparoscopic procedures. Mouret, a surgeon, is generally credited with performing

[3]CCD – a charge-coupled device, invented in 1969, is an array of photo dectectors printed onto a silicon chip. These sensors can be more sensitive than photograph film, can be integrated directly with other circuits on the silicon chip, and are used for many applications including satelitte imaging, video, and digital cameras.

the first fully laparoscopic cholecystectomy (gall bladder removal) in 1987. His effort was however actually predated by Erich Mühe in 1985. This has become the most widely performed laparoscopic surgery. Semm, Mühe, and to a lesser degree Mouret received much more criticism than acclaim for their initial attempts.

Today laparoscopy is performed by gynecologists, urologists, and a wide array of surgical specialists. Minimally invasive procedures making use of similar techniques are now employed in most parts of the body.

2.3 Instrumentation

A large variety of instruments have been developed for use during laparoscopic surgery.

2.3.1 Laparoscopes

The laparoscope is the fundamental instrument in laparoscopic surgery. A laparoscope is a long tubular telescope. Modern laparoscopes are very similar in form to early cystoscopes (telescopes used to examine the inside of the urinary bladder) which were used as the first laparoscopes. An example of an early cystoscope is shown in Figure 2.1. Most laparoscopes feature two optical channels, a viewing channel, and an illumination channel. The illumination channel usually surrounds the viewing channel and conducts light using unstructured fiber optic filaments. Light is introduced into this channel by a port which usually extends perpendicularly to the shaft of the scope near the head. The light is transmitted through this channel and onto the operative site via a ring or crescent shaped opening on the scope's tip. The viewing channel extends from the head of the scope to the tip and is the channel through which a surgeon can either directly view the operative site or, more typically, attach a video camera for display on a video monitor. There are several significant variations. Some laparoscopes do not have an optical viewing channel but instead have an imaging sensor (usually CCD) at the tip of the laparoscope. The output from this sensor is attached directly to display devices.

Stereo laparoscopes, usually called "three-dimensional laparoscopes", feature either two sensors at the instrument's tip or two optical viewing channels, Image splitters and stereo adapters are also available that make use of a single lens laparoscope through a single optical channel. The output from these scopes or splitters can be shown on special stereo displays (usually requiring special glasses) to

give an effect similar to "3-D" movies[Durrani95]. Despite a significant number of research articles on the subject it is not clear what, if any benefit there is to these devices. While some studies report some performance improvement, others do not, and many report adverse effects, such as headaches from use of the stereo devices[Mueller99, Breedveld00, Herron99].

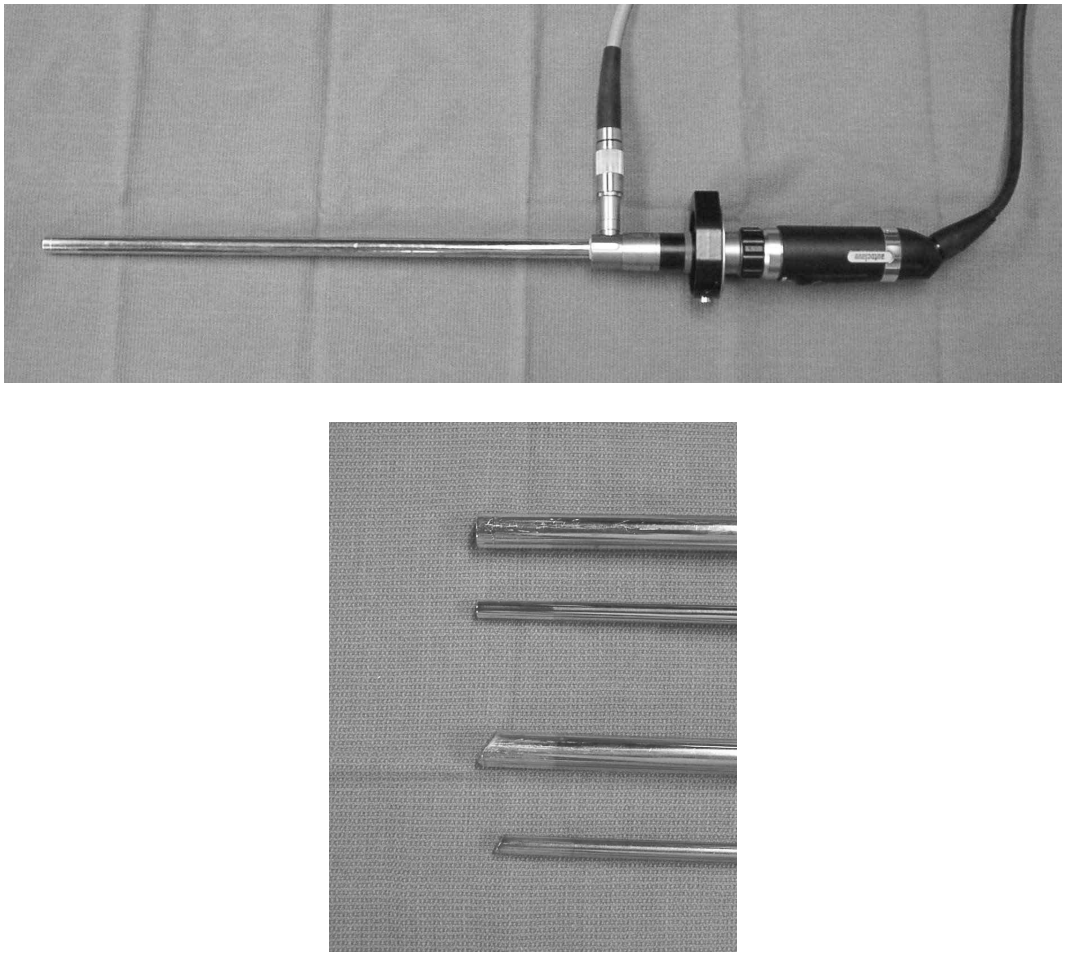


Figure 2.2: Top: a typical laparoscope (10mm) with a camera and light guide cable attached as they would be used in surgery. Bottom: the tips of an assortment of laparoscopes with different diameters and viewing angles.

Laparoscopes sometimes have additional features. Self-cleaning and defogging scopes feature a small spray nozzle directed at the viewing lens on the tip. This allows the surgeon to clear fog or debris from the scope's tip without removing it from the patient. Some laparoscopes feature irrigation and suction channels for clearing the operative site. Some laparoscopes feature an instrument channel which allows the introduction of an instrument into the abdomen through the scope without creating

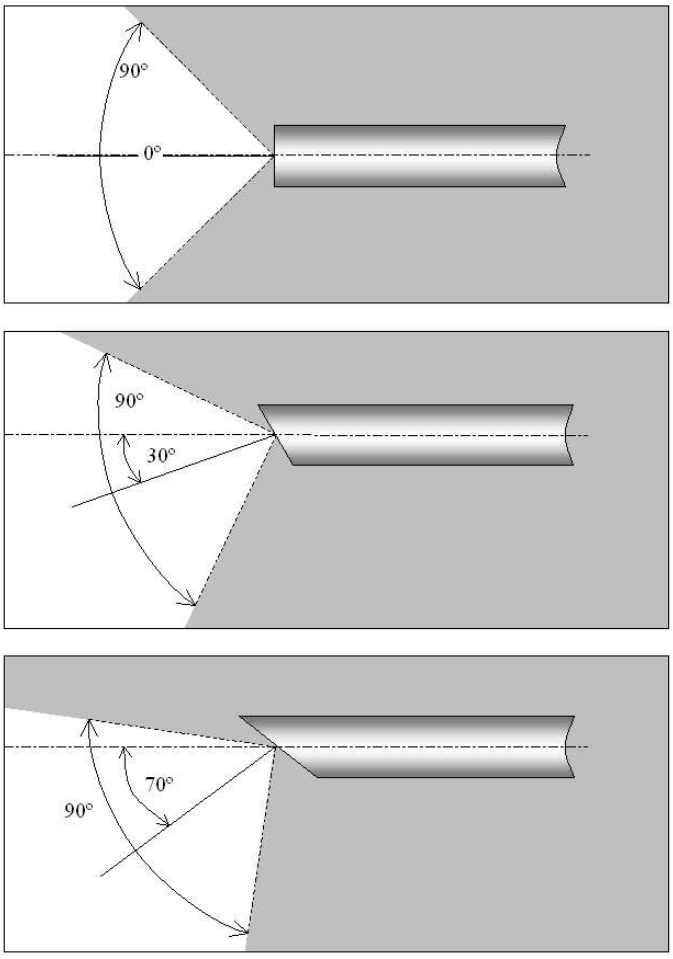


Figure 2.3: Laparoscopes with 0°, 30°, and 70° angled views.

an additional port. These channels are most frequently used for biopsy instruments or for laser ablation in procedures requiring little movement of tissues and organs.

Laparoscopes are available in several diameters (see Figure 2.2). Larger diameter scopes generally provide clearer and brighter visualization, but may be harder to navigate into tight locations and require a larger incision.

Laparoscopes are also available with different viewing angles and fields of view (see Figure 2.3). Angled scopes have a direction of view off axis from the body of the scope. Angled scopes have the advantage of being able to view places that straight scopes cannot. This is because straight scopes can only view in the direction that the tip is pointing which limits them to places that the tip can actually reach. Angled scopes are more difficult to learn to use. In addition to varying angle of view, scopes also vary in how wide a field of view they have. Scopes with narrower field of view create more magnification of the operative site, while wide field of view scopes provide a better overview of the surgical site.

Laparoscopes are usually used with video cameras to magnify the view, to document procedures, to improve ergonomics, and to allow more than one person to observe the procedure. Laparoscopes need a strong light source (typically a 400 or 800-Watt xenon source), a video camera, and a video monitor. Light is conducted from the source to the laparoscope via flexible fiber-optic cables. A small video camera is attached directly to the scope (or sometime built into the scope). A typical video tower containing a light source, camera controller, and video monitor is shown in Figure 2.4.

2.3.2 Insufflation

The insufflator is a device used to create and maintain pneumoperituneum. These devices contain automated control systems that regulate the gas pressure in the abdomen.

The Veress needle consists of a spring-loaded blunt-tipped tube inside a sharp outer needle. The needle is inserted with the inner tube retracted. Pressure from dense tissue (e.g., skin and muscle) prevents the blunt inner portion from advancing while the needle is in dense tissue. On reaching a space or less dense tissue, the inner tube advances thereby creating a guard that prevents the sharp out needle from puncturing deeper structures.

The shield prevents cutting injury from the needle to abdominal organs. Once the Veress needle has penetrated the peritoneum, its placement is verified using the

Figure 2.4: A typical laparoscopic "tower" (manufactured by Olympus). It contains a video monitor, a camera controller (first shelf), a light source (second shelf), and an insufflator (third shelf). Laparoscopic towers frequently also contain a video recorder or video printer for documentation during procedures.

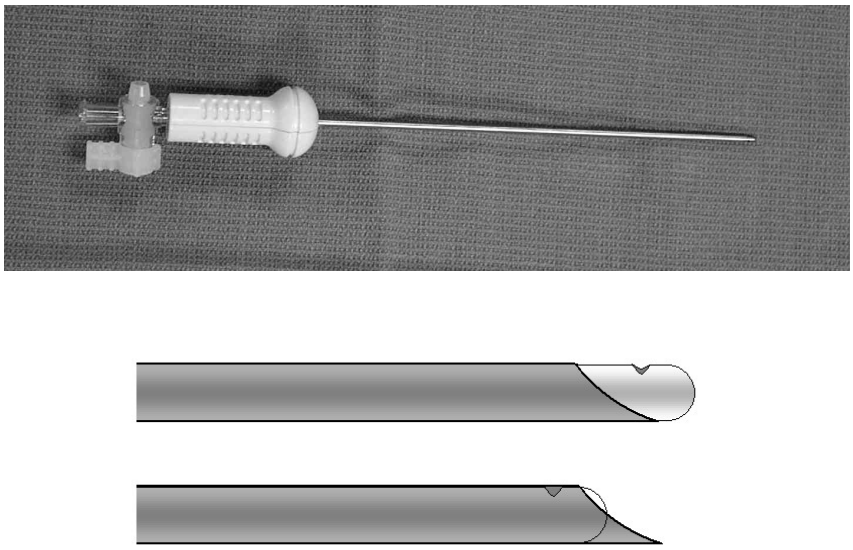


Figure 2.5: Top: a modern Veress needle (Ethicon Endoscurgery). Bottom: the tip of the Veress needle appears (bottom) as the needle is advanced with the blunt tip retracted and the cutting edge of the needle fully exposed, (upper) the blunt inner tube extended. A small area of the blunt tip is exposed while advancing to trigger the release mechanism. The inner tip is spring-loaded and extends to prevent accidentally puncturing internal organs after the needle is advanced through the peritonium. An open air channel is available through the needle with the tip in either position allowing suction to verify correct needle placement (detecting bleeding or fluid indicates incorrect placement) as well as pressurized air for insufflation.

technique described in Section 2.4.1.1 Footnote 4, and the needle can be attached to the insufflator. For an image of a typical Veress needle and detail drawings of the tip, see Figure 2.5.

2.3.3 Access

A trocar or "port" is a metal or plastic sleeve inserted through the abdominal wall through which laparoscopic instruments gain access to the abdomen. Trocars are used to enter the operative site within the body.

Many trocars have a self-retaining blade (similar to the mechanism in the Veress needle) to assist insertion and to prevent injury to abdominal organs. The self retaining blade is exposed while the trocar is inserted, but then is shielded automatically on entering the peritoneum. Others feature a non-bladed screw-threaded tip, which enables a twisting motion to advance the trocar through a small incision. Many trocars have a system of seals which prevents leakage of gas through the trocar both when there is an instrument being passed through the trocar and when the trocar is in its passive mode.

Many trocars also can be connected to the insufflator to monitor and to maintain the intraabdominal pressure. When initial access and pneumoperitioneum is established using the direct visualization method, a trocar is often used for beginning insufflation rather than using a Veress needle.

Trocars are available in a variety of diameters to allow use of a variety of instruments (see Figure 2.6. Small ports are used whenever possible since they require the smallest incisions for insertion and are least likely to leak.

2.3.4 Interventional instruments

A wide variety of cutting, dissecting, and grasping instruments have been created for use during laparoscopic procedures. Most of these devices are similar in function to those developed for open procedures, but feature an instrument tip (typically miniaturized) on the end of an elongated shaft (Figure 2.7). Several different styles of handles have been developed for laparoscopic instruments as well. Several styles of handle are illustrated in Figure 2.8. Some instruments are single use (disposable), while others are designed for sterilization and re-use. There are some partially reusable instruments, also called "reposable", which use single use instrument tips on a re-usable handle. Some laparoscopic instruments also feature an articulated head

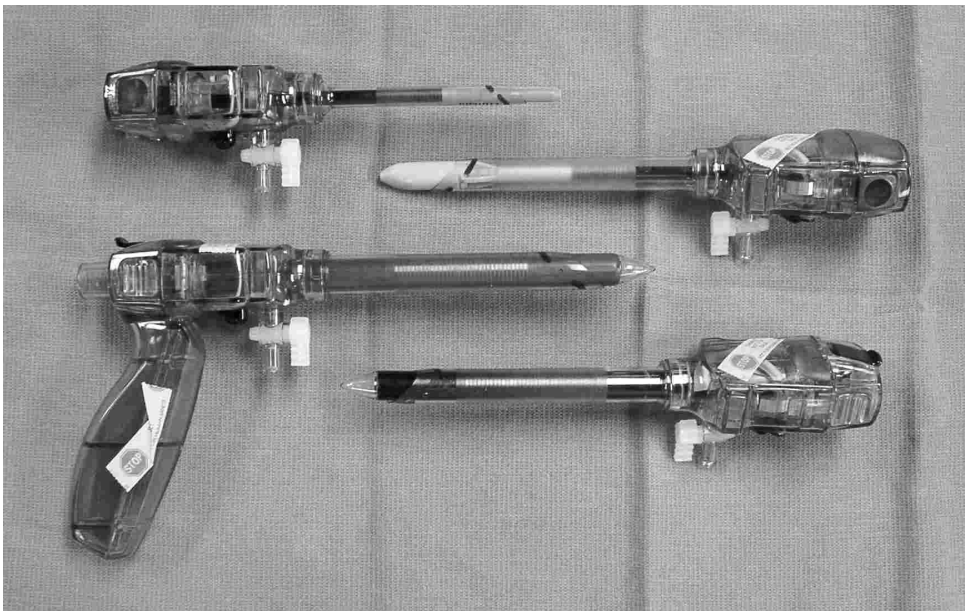


Figure 2.6: An assortment of laparoscopic trocars manufactured by Ethicon Endo-surgery. Trocars come in a variety of diameters with many options to assist gaining access. The top trocar is a 5mm port and the second is a 10mm port. These both feature self-retaining cutting blades. The bottom two are 10mm ports designed for optical access. Ridges on the clear point may be used to dissect through abdominal wall tissue under visualization through a scope placed in the trocar. In all of these devices the cutting portion can be removed to insert other instruments.

which gives additional degrees of freedom for the instrument tip. Most articulated instruments add a single direction of flexion a short distance from the tip. Many articulated instruments have a small number of preset locking positions for the flexing joint rather than allowing continuous adjustments during use.

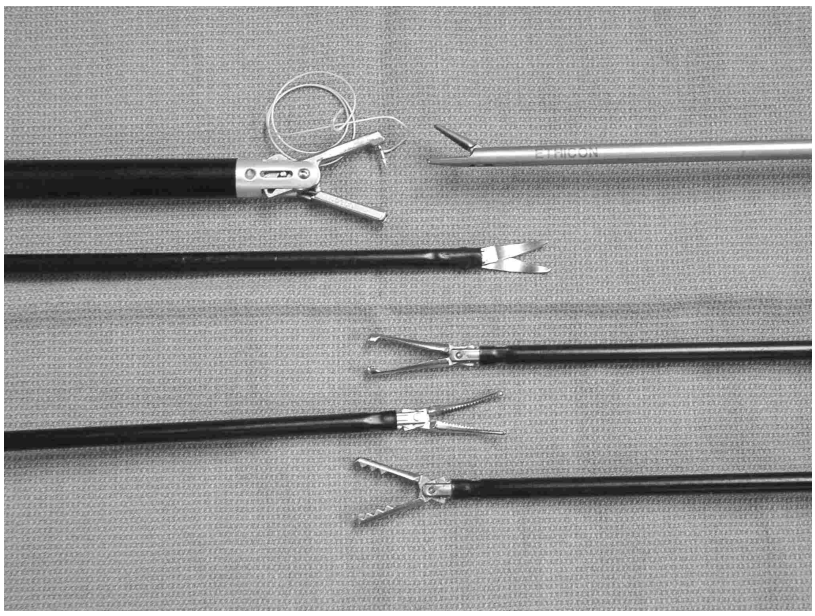


Figure 2.7: The working tips of some of the variety of laparoscopic instruments available. From top to bottom, starting on the right, these are: a needle driver, an Endo-Stitch suturing device, a scissors, an atraumatic grasper, curved dissector, and a tissue forceps.

2.3.4.1 Cutting and cautery

Cutting and dissecting instruments can be augmented with cautery features. These features are useful for controling bleeding (hemostasis), avoiding bleeding while cutting tissue. They also enhance the cutting ability of an instrument. Many instruments are designed with contacts for use of monopolar electrocautery.

More specialized instruments have contacts for bipolar electrocautery. Monopolar electrocautery cauterizes and cuts through tissue by applying current at the electrical tip. Risk of electrical injury away from the intended point of contact is reduced by placing a large grounding pad under the patient and insuring that there is good conduction with the patient. Bipolar electrocautery works primarily through the heating effect caused by rapidly osscillating current passing through the instrument tip or through tissue placed between the two electrically isolated tips of a single

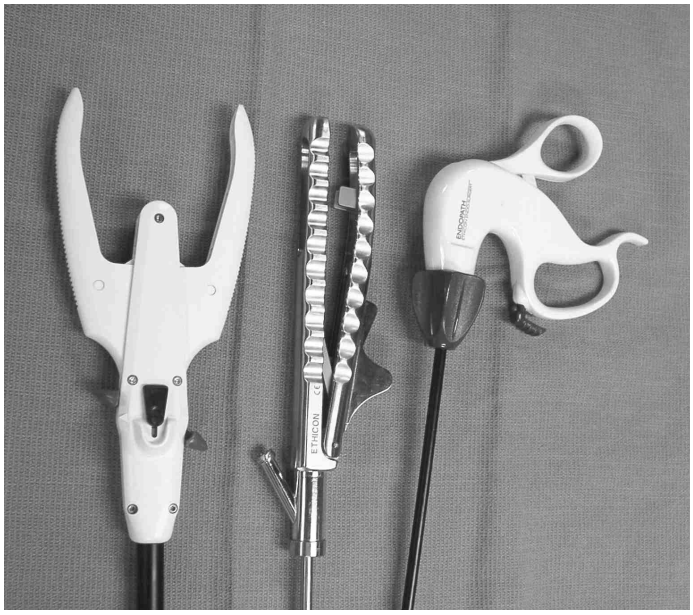


Figure 2.8: Laparoscopic instrument handles. Several additional varieties have been developed that are not shown here. Pictured here are (from left to right) an Endo-Stitch suture device, a needle driver, and a grasper.

instrument. Electrocautery must be used with caution to avoid unintended burns and electrical injury. Electrocautery can sometimes cause unseen injury away from the site of application due to arcing between the conductive surfaces of internal organs.

Harmonic cutting and cautery devices are also widely used. Harmonic instruments cut by a localized heating effect produced by high speed vibration of the instrument tip. These devices cause rapid cautery of small blood vessels with less heat generation than other cautery devices. Harmonic devices carry no risk of electrical injury but are substantially more expensive.

Lasers are also used for cutting and cautery and were quite popular in the early 1990s. Laser cautery devices use fiber-optic fibers to transmit high energy laser energy inside the body. The fibers are designed to focus the energy very near the tip of the fiber so that it burns and cuts on contact with tissues. In some instances the energy from the laser can be redirected and cause unintended and difficult to detect burns away from the instrument tip. Several studies have shown that electrocautery is more efficent and less expensive to use than laser cautery and it is now not widely used [CR90, Corbitt91].

2.3.4.2 Suturing, stapling, and ligation

Closing wounds, ligating structures, and attaching tissue surfaces is more difficult and time consuming when done laparoscopically as compared to open surgery. For this reason a substantial number of devices have been created for these tasks. These include rods designed to push externally tied knots through the trocars and into position, pretied loops on the end of a positioning device, specialized knot tying instruments, stapling, and clipping devices.

2.4 Techniques

All laparoscopic procedures have several steps in common. The first step is obtaining access to the abdominal cavity. Once an instrument (either a trocar or Veress needle) has been placed in the abdomen, the process of insufflation can be begun.

2.4.1 Obtaining access

Initial access, that is creating pneumoperitoneum and placing the first trocar, can be done using a Veress needle, using an open technique, or using an optical access technique. In each of these methods, the goal is to introduce an instrument into the peritoneum with the least chance of causing damage to any abdominal organs so that the peritoneum can be inflated to create pneumoperitoneum. Usually the outer layer of skin, is cut with a scalpel before introducing a trocar, even one with a cutting blade, to reduce the amount of force needed to penetrate into the peritoneum.

2.4.1.1 Veress needle technique

The Veress needle technique can be faster than the open technique and requires a much smaller incision.

A Veress needle (described in Section 2.3.2 and shown in Figure 2.5) is inserted (percuntaneously) into the abdomen, usually with mechanical retraction to pull the abdominal wall away from abdominal organs. This retraction has not been proven to be benefical[Briel00], but it is widely recommended. During insertion, suction from a syringe is applied through the needle to detect entry into large blood vessels, fluid filled spaces, or viseral organs. After the needle is placed and verified[4], insufflation is

[4]The following five procedures can be used. A 10ml syringe containing 5ml of saline is attached

started. After the appropriate amount of pressure in pneumoperitoneum is reached, the first trocar is blindly inserted through the abdominal wall (using a self-retaining blade) and into the pneumoperitoneum. Subsequent trocars are placed by watching their insertion into the pneumoperitoneum from inside using the laparoscope.

The Veress technique is less likely to succeed in obese patients. Studies have shown that multiple insertions are frequently required in these patients and that there is a higher incidence of other complications. This may be due to the increased transit distance of the needle and because there is less effective retraction. Preperitoneal insufflation, inflation of a space under the skin but outside of the peritoneum, must also be guarded against[Curet00]. Alternative approaches using the Veress needle include trans-uterine insertion which is safe and effective in obese patients[Santala99]. The Veress technique is also not recommended in patients who have had prior abdominal surgery. These patients have a higher probability of adhesions (attachment of abdominal organs to each other and/or the abdominal wall) which puts them at significantly higher risk of injury to abdominal organs during both needle insertion and trocar insertion[Lecuru01].

2.4.1.2 Open technique

The open technique is preferred under some conditions because insertion is done under direct visualization minimizing the risk of injury to abdominal organs. There are a large numbber of variations on the open technique, first published by Hasson in 1978[Hasson74], so a typical approach is described here.

In the open technique an incision is made just below the umbilicus and careful disssection is made through each of the layers of the abdominal wall. The peritoneum is either incised with a scalpel or with the blade of a trocar with a self-retracting blade. The incision site is grasped and lifted creating an organ-free area to advance the trocar into. This trocar is usually sutured in place to prevent movement and leakage. The insufflator is then attached to an air inlet on the trocar to create pneumoperitoneum. As in the Veress technique, subsequent insertions are observed laparoscopically to avoid injury.

---

to the Veress needle. First the needle is aspirated and any recovered fluid is identified (ideally none). Next some of the saline is injected and there should be no resistance. The needle is again aspirated and no fluid should be collected. With the syringe not connected to the needle, a drop of saline can be placed in the hub of the needle and it should be pulled into the needle ("drop") with retraction of the abdominal wall (the "drop test"). Finally, the Veress needle should be able to be advanced another 1-2cm without resistance.

While the open technique can usually be shown to be slightly safer than other techniques for gaining access there are several disadvantages. The first is that it can be more difficult to get and maintain a good seal around the trocar. Without a good seal pneumoperitoneum will be lost. The second major disadvantage is that the open technique requires making a larger initial incision than other methods thereby increasing the risk of incision site complications and visible scarring.

2.4.1.3 Optical Access

Optical access utilizes a special trocar on which the cutting blades (often plastic ridges or screw threads) are located on the outer surface of the trocar and a clear viewing window is available at the tip. A 0°-angle laparoscope (viewing straight ahead) is placed in the trocar so the tissues being penetrated can be observed as the trocar penetrates them[Melzer95]. This allows more rapid recognition of possible injuries (as compared to blind insertion in the Veress needle technique), and a better seal and less scarring (as compared to the open technique).

2.4.2 Insufflation

After a Veress needle or trocar has been placed, insufflation is begun. Insufflation is the inflation of the abdomen to create a working space. Insufflation also lifts the abdominal wall away from abdominal organs thereby decreasing the risk of accidental injury to internal organs when inserting additional trocars.

Modern insufflators have automated regulation controls and monitors to insure that proper insufflation is maintained. Overly high pressure can cause injuries such as gas embolism or impede proper blood flow. Low pressure leads to loss of pneumoperitoneum and the loss of the working space. High pressure (over 14mmHg in a healthy patient) is associated with an increased risk of complications including cardiovascular effects, air embolism, postopertative pain, decreased rate of healing, increased rate of tumor seeding.

2.5 Benefits and Risks of Laparoscopy

There is a general consensus that laparoscopy is sometimes beneficial, that there is an increased risk for some complications, and that laparoscopic procedures are generally

more expensive but may be less expensive if extended care costs are considered. Beyond these broad statements there is little agreement.

Since laparoscopy became widely used, the surgical community has accumulated a large body of literature on its risks and benefits. Widespread use of laparoscopy, particularly in the setting of general surgery, is a relatively recent development so almost all of the literature before 1990 is anecdotal in nature. Until recently large clinical trials, primarily done retrospectively, were rare. Recently published retrospective clinical trials exclude often patients operated on before 1996 citing the rapid changes in instrumentation and technique that would confound meaningful analysis. Many of the studies examining the costs and benefits of laparoscopic surgery are done retrospectively and may use historical cost and complication rates as the basis for comparision. This type of comparision of laparoscopic procedures with open procedures is problematic because the managment of patients during and after open procedures is more variable. A small number of randomized, prospective studies have been performed which may give more reliable information.

The benefits of laparoscopy are fairly widely known and accepted outside of the medical community. While faster recovery, smaller scars, and less postoperative pain make these procedures attractive to patients, less lost productivity, and shorter hospital stays leading to presumed lower overall costs have been more important forces driving the rapid acceptance of laparoscopic procedures by health care providers and funders. However, in many instances the procedure itself is more expensive when performed laparoscopically due to the extra cost of specialized equipment and longer operative times.

What has largely escaped notice outside the surgical community is the increased risk of morbitity and, in some instances, mortality associated with many laparoscopic procedures.

2.5.1 Costs

Almost all studies have shown increased "upfront" costs of laparoscopy. This cost is mainly due to increased equipment costs and increase in operating room time. Cost savings from laparoscopy have also been demonstrated. These savings are derived from the cost of postoperative care, and indirect savings through increased productivity. The wide range of studies that have attempted to document the associated costs and savings of laparoscopy are difficult to compare because of differences in methods of calculating costs and reimbursement structures[Greenberg01, Jonsson00].

The true cost savings of laparoscopy is a contentious issue[Hopkins00], and depends highly on the procedure in question (even the type of instruments used can be important), and costs associated with a particular hospital or healthcare system. For some procedures the difference in average hospital stay is very small[Nicholson01]. In some cases this difference is narrowing due to improvements in the open techniques. In general, laparoscopic cholecystectomy (gallbladder removal) is cost effective[Calvert00], laparoscopic appendectomy is marginally cost effective[Long01], and laparoscopic hernia repair is not cost effective[O'Boyle01].

2.5.2 Benefits

The publicized benefits of laparoscopic surgery are generally upheld by the published literature. In particular, decreased postoperative hospitalization, less postoperative pain, faster improvement in quality of life, and decreased hospital stays are most frequently sited. Another important factor for some patients is the better cosmetic result, smaller scars, of laparoscopy.

Common post-operative complications including wound infections, incisional hernia, urinary tract infection, and pneumonia are decreased with laparoscopy[Poulin01]. A related finding is that patients who have laparoscopic procedures have less postoperative stress[Nishiguchi01]. Blood loss during laparoscopy has been shown to be lower for some laparoscopic procedures as compared to open procedures[Nguyen01]. Formation of adhesions, attachments of organs to themselves, each other, and the abdominal wall, that create an increased risk for future surgery and bowel obstruction are less common with a laparoscopic approach than with an open approach[Polymeneas01].

Among the elderly, laparoscopic surgery has important documented benefits such as reducing pain and length of hospital stay (by approximately 40%)[Tuech00]. Significantly, elderly patients undergoing laparoscopic surgery were more likely to maintain their level of independence than those having open surgery (94% laparoscopically versus 77% open)[Stocchi00].

Laparoscopic surgery is being applied to children with increasing frequency. Open surgery can be very difficult in children because of the small size of structures and working space. Laparoscopic procedures have been shown to be better or comparable in safety and efficacy for some difficult procedures such as splenectomy[Liu00, Minkes00].

2.5.3 Risks

There are currently few reasons why a patient absolutely should not have laparoscopic surgery. Initially, many reasons were advocated including, but not limited to current pregnancy, prior abdominal surgery, acute conditions (such as appendicitis and cholecystitis), obesity, and cirrhosis. As techniques have advanced, these limitations on laparoscopy have become relative rather than absolute contraindications. Likewise, the risks of laparoscopic surgery for all patients has dramatically decreased over the past ten years. There remain however, many risks of laparoscopic surgery, both proven and suspected.

The first group of complications of laparoscopy are injuries related to gaining access to the operative site. The most common of these are damage to blood vessels and abdominal organs while inserting the Veress needle or trocars. The risk of these injuries is estimated to be between 0.09% and 0.27% with the open methods having the lowest incidence and optical and Veress methods having higher incidence of this type of injury [Mayol97, Schafer01, Catarci01]. Vascular injuries usually involve the arteries or veins in omentum or mesyntery, and damage to abdominal organs is most frequently bowel perforation. Up to 5% of these injuries have been reported to be fatal, and approximately 25% require reoperation to repair [Catarci01]. These injuries are usually minor if they are immediately recognized but pose a greater risk if they remain undetected.

Pneumoperitoneum itself presents potential risks. Inflation of the abdomen gives greater resistance to breathing. For spontaneously breathing patients this increases the work of breathing, and for ventilated patients this requires that greater pressure be applied to adequately ventilate them. The carbon dioxide gas used for insufflation diffuses through tissue and may be absorbed. At extremely high pressure there is some risk of gas embolism (gas bubbles forming in blood vessels), but under normal conditions the carbon dioxide is rapidly converted to carbonic acid. This additional acid is rapidly eliminated, but may cause acidosis in some patients, particularly those with known respiratory, renal, or cardiovascular disorders. There is also some indication that increased abdominal pressure may also increase intracranial pressure. Post-operatively many patients feel varying degrees of discomfort due to trapped gas from pneumoperitoneum. All of these effects are more likely to cause problems as the pressure of pneumoperitoneum is increased.

In some instances the laparoscopic approach is associated with other risks (specific to the procedure). In the case of laparoscopic cholecystectomy, injury to the bile

duct or hepatic artery occurs in approximately 1.4% of patients[Calvete00]. Only about 40% of these injuries are recognized during the procedure and all need surgical repair[Bingham00]. In gastric bypass, patients undergoing the laparoscopic procedure have a higher incidence of anastomotic stricture (post-operative narrowing of the bowel where it was reattached)[Nguyen01], and some studies reveal a higher rate of abcess formation following laparoscopic appendectomy.

There is some evidence that suggests that laparoscopic resection of malignant tumor is not as effective as open resection. In particular, metasteses (often to the trocar sites) and recurrance appear to be more common. A molecular mechanism for this effect has been proposed[Hajri00].

Common complications of all types of surgery including adhesion formation, systemic or localized infections, and incisional (port site) hernia are also risks of laparoscopic surgery. These risks are generally lower than those for open surgery but they are not infrequent. Because laparoscopic procedures often take longer than open procedures, patients may be at an increased risk of anesthesia related complications. Conversion to open surgery is often reported as a complication of laparoscopic surgery. This is misleading in that conversion can be used as a safety net for the surgeon in the event of injury or other complicating factor. Surgeons choose to convert to open surgery when that is the best option for successful patient management.

2.6 Novel approaches to laparoscopy

Beyond incremental improvements in techniques and instrumentation for laparoscopy there are several significant new advances in laparoscopy. In some instances approaches initially used in laparoscopy are now being introduced into open procedures to improve visualization, reduce the size of incisions, enable telecollaboration capabilities, and improve ergonomics. The hybridization of open and laparoscopic procedures also includes hand-assisted laparoscopic surgery in which surgeons use their hands under laparoscopic visualization. Finally, the introduction of robotic assistance for laparoscopy has introduced many new possibilities.

2.6.1 Hand assisted laparoscopy

Hand assisted laparoscopic surgery is a technique whose benefit has now been shown for a variety of procedures. Such surgery is performed nearly identically to conven-

tional laparoscopy except that a single larger incision, approximately 7cm but varying depending on the size of the surgeon's hand, is made. A special trocar (e.g., Hand-Port by Smith and Nephew or LapDisk by Ethicon Endosurgery) is placed in this incision to prevent loss of pneumoperitoneum. When needed the surgeon can insert a hand through this port allowing direct manipulation of the surgical field with visual guidance through the laproscope (see Figure 2.9).

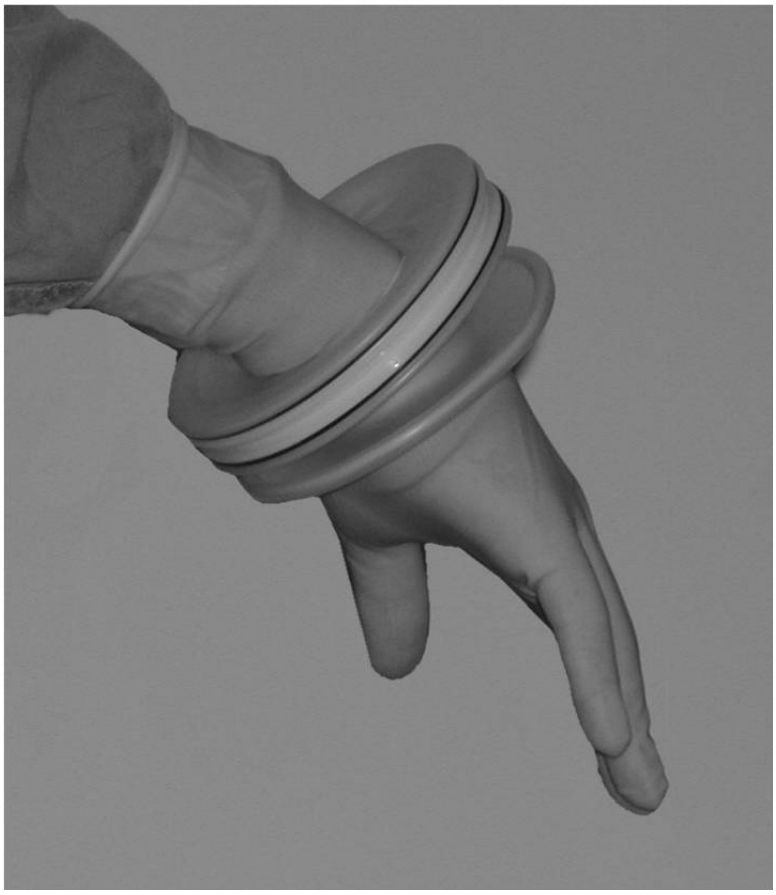


Figure 2.9: A LapDisk (distributed by Ethicon Endosurgery) used for hand assisted laparoscopic surgery. The upper ring contains a sealing mechanism to create an air-tight seal around the surgeon's hand. The lower ring is made of a springy material that can be deformed and inserted through a small incision. When the material springs back to its normal shape the rubber membrane around both rings forms an airtight seal with the incision.

This technique has been shown to be beneficial in several procedures, particularly in those where a larger incision is needed anyway (e.g., live-donor kidney harvesting, large solid tumor resection). The benefit appears to come from two sources: increased tactile feedback from direct manipulation, and increased dexterity over using

laparoscopic instruments.

2.6.2 Hybrid techniques

The widespread acceptance of laparoscopy has meant that the instrumentation for laparoscopic procedures is now available in many operating rooms. This has lead to the development of hybridized techniques in open surgery where laparoscopic instruments and techniques are applied to open surgeries.

One example of this type of hybridization include use of laparoscopes to magnify and better visualize manipulations in small incisions. Another example is to use a laparoscope and robotic camera driver to get good video of open procedures for teleproctoring applications. Suture and staple devices developed for laparoscopy can be used in open procedures.

2.6.3 Surgical robotics

Several robotic[5] devices are now available and, in some instances, are approved for use during laparoscopic surgery. The three most widely used robotic assistance devices are the *Zeus* and *AESOP* devices built by Computer Motion [CompMo] and the *da Vinci* device by Intuitive Surgical[Intuitive]. The *AESOP* is a remote-controlled camera holder used to aim and hold the laparoscope. The *Zeus* and *da Vinci* systems are master-slave devices which control laparoscopic instruments.

2.6.3.1 Camera driving robotics

AESOP is now widely used in the United States. The first version of this robot was introduced in 1994 and was controlled by foot pedals or by a miniaturized joystick. Later versions added additional degrees of freedom of motion as well as voice control. The alternative to using a device like *AESOP* is to have an assistant hold the laparoscope. This job is often given to an assisting surgeon, resident, or medical student. In contrast to an assistant, the robot always holds the laparoscope steady and is consistant in how it moves the camera. Computer Motion estimates that a single *AESOP* can be used for an average of 240 operations per year. While a useful

[5]It is important to point out that these devices, while commonly called robots, perform few, if any, automatous actions. They are more accurately described as remote control or master-slave devices because they are directly guided by the user. In keeping with the medical literature, I will refer to these devices as robots.

tool, *AESOP*'s primary benefit is reducing the number of personnel needed during laparoscopic procedures.

2.6.3.2 Robotics for manipulation

The second type of surgical robot has had a more dramatic effect in the operating room. While there are substantial differences between the two currently available devices, they operate on the same general principles and therefore will be considered together.

Current systems feature three robotic arms which control the position of a laparoscopic camera and two laparoscopic instruments. A wide variety of instruments can be placed in the robotic arms as needed for specific procedures. These instruments frequently have additional joints not typically seen in traditional laparoscopic instruments. The surgeon sits at a console and grasps a control device with each hand. The control devices can be moved with six degrees of freedom with an additional open-close control. The instrument tips inside the patient mimic the motion of the control device except that these motions are scaled down and tremor can be eliminated. See Figure 2.10 for a view through one of these systems. The extra joints in the laparoscopic instruments act like a wrist and give the instrument tips two additional degrees of freedom of motion (for a total of six).

The surgeon may be provided with a stereo or monocular view of the surgical site. The camera is positioned so that the relationship between the camera and the instruments is similar to that between the surgeon's eyes and hands. These systems have the theoretical ability to provide enhanced force reflection, magnification of the forces acting on the robotic tool tips back to the surgeon, but this has not yet been used clinically. A schematic overview of such a system deployed in an operating room is shown in Figure 2.11.

These systems provide a primary benefit of increased dexterity, particularly in difficult to access locations. They also provide an intuitive way to use laparoscopic instrumentation. A final potential benefit of these systems is for teleoperation. These systems are in fact tele-operative devices in that the surgeon is actually some distance from the patient.

While the surgeon is usually in the same operating room as the patient, several demonstrations have been performed where the surgeon has been at a considerable distance (up to 7000km away, from New York to Strasbourg)[Marescaux01]. While several trans-Atlantic telesurgeries have been successfully performed, surgeons are

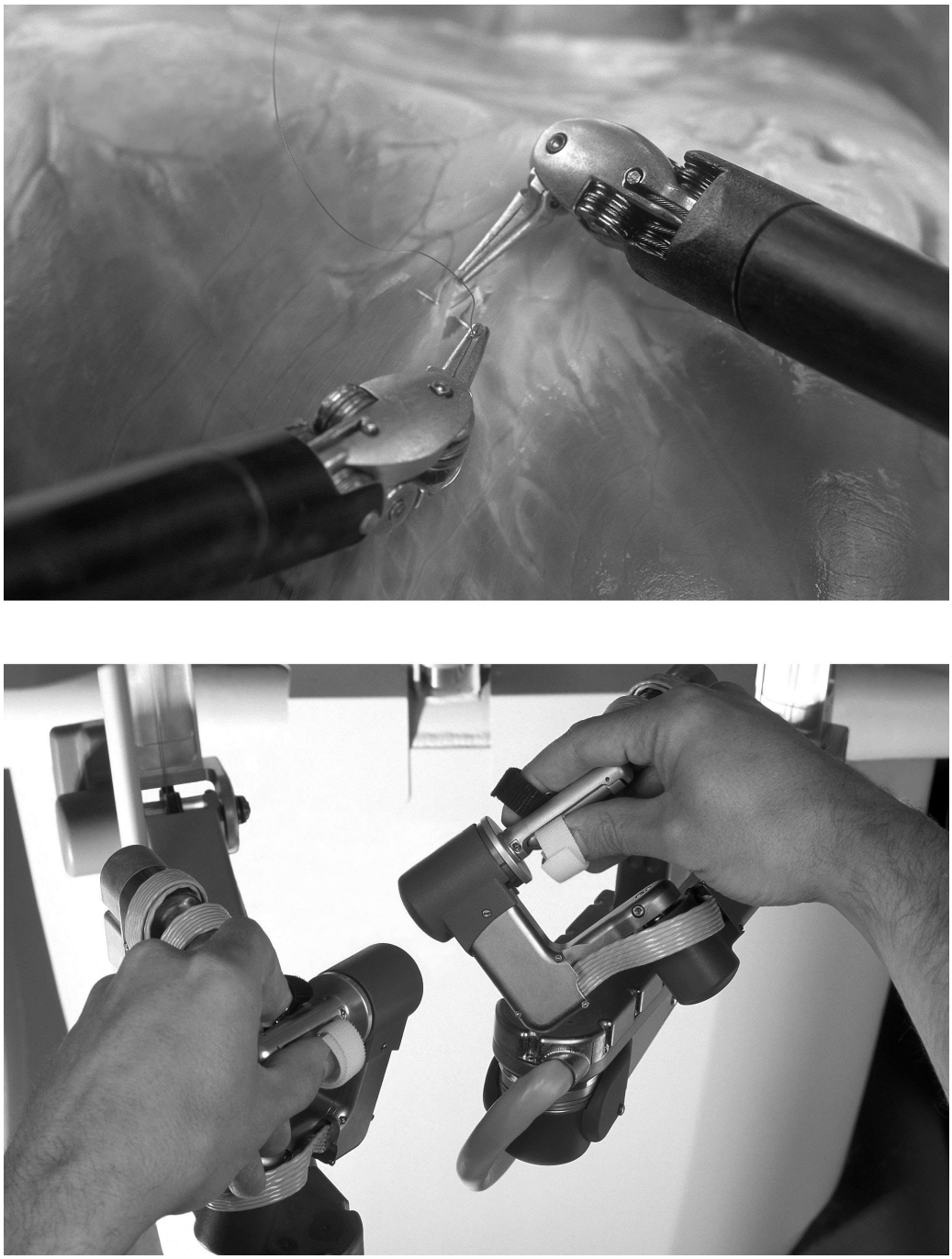


Figure 2.10: Top: the surgeon's view of the operative site while using the *da Vinci* robotic system. Bottom: the surgeon's hands in the controllers (image courtesy Intuitive Surgical, Inc.).

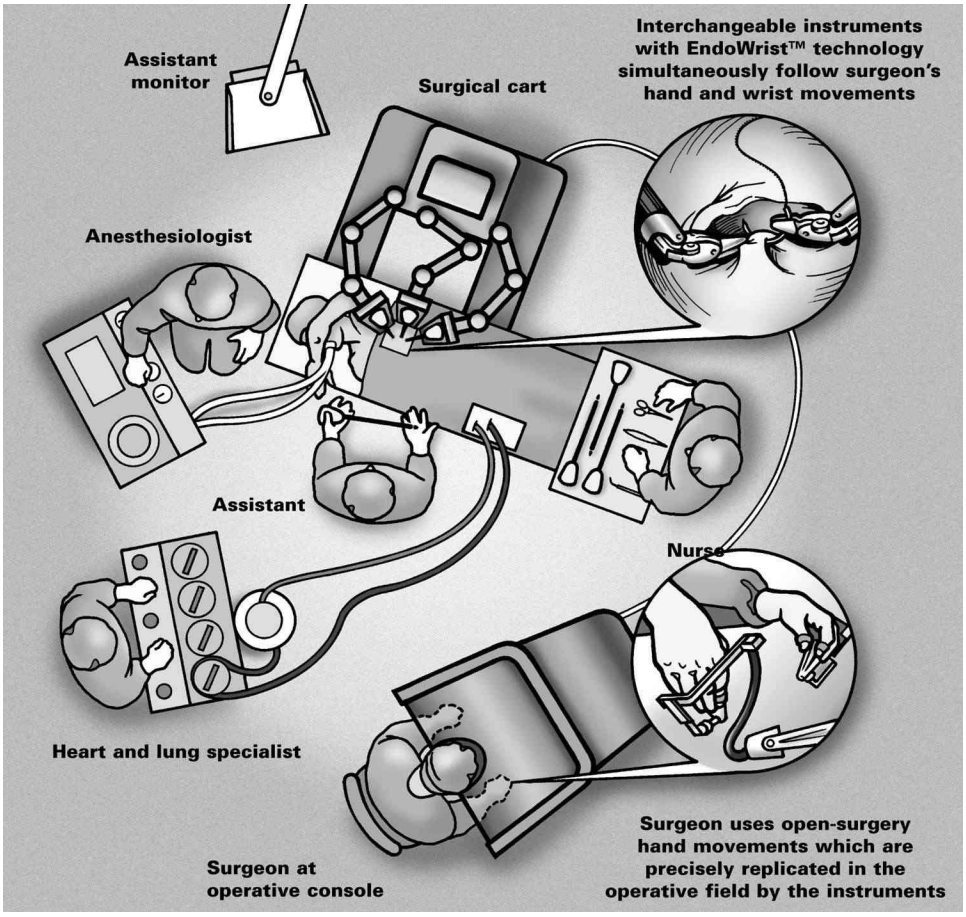


Figure 2.11: A schematic overview of the *da Vinci* robotic system deployed in an operating room (image courtesy Intuitive Surgical).

still needed at the remote site (where the patient is) in order to initially place the trocars, control electrocautery, and to take over the procedure in the event of system failure.

These systems still have several inherent problems. While local navigation (e.g., the surgeon wants to follow an anatomical structure currently partially in view) is relatively easy, disorientation to the larger scale (e.g., trying to find the gallblader while looking at the pancreas) is not uncommon. Surgeons also lose track of which way true "up" and "down" are, and are sometimes suprised when a suture falls "up" on their screen when they release it. A final problem with these systems is cost. With the exception of a small number of procedures that have only been performed with robotic assistance, no substantial benefit has been shown using these devices.

Although today only a relatively small number of surgical robots are in use, they are one of the most promising technologies in surgery. The most impressive clinical contribution of these devices is the ability to do coronary artery bypass procedures ("heart bypass") without the standard thoracotomy (opening the chest by cutting the sternum and spreading the ribs). The robotics give the surgeon the dexterity needed to work using instruments inserted between ribs[Shennib98]. Not suprisingly, patients require significantly less pain medication and recover from surgery faster[Kappert00].

In addition to the improvement in dexterity that may be achievable by scaling motion and the additional degrees of freedom, these robotic systems represent a significant increase in the amount of computational power available in the operating room. Their acceptance opens the way for other systems and devices that may share resources for improving coordination and visualization. Most importantly, this technology has made, and will likely continue to make, new procedures possible.

Chapter 3

About augmented reality[1]

"The eye guides the other senses and verifies their impressions."

– Philip Bozzini, 1806

An augmented or "mixed" reality system is one in which computer graphics objects are added into the user's surrounding scene in a meaningful way in near real-time. Augmented reality systems can represent a wide continuum from heads-up displays, where very simple graphics are added to the user's view in very small quanitites, to near virtual reality applications, where only small parts of the real world, such as the user's hands, are visible to the user. Augmented reality includes applications ranging from "head-up" displays for fighter pilots to complex real-time visualizations of multiple data sets in medical applications. Other applications include previewing architectural and engineering design.

Augmented reality is similar to the computer graphics that are frequently used in movies and in television where special effects, even major characters, are added to a scene after the original filming is completed. While many of the computer graphics principles and techniques used to generate these video and film sequences are similar to augmented reality systems, they are different in that the producers generally add the imagery off-line. The time gap, often many months after a film is shot, allows the artists and animators the luxury of adjusting and correcting the computer graphics to make them perfect. It also allows them to employ rendering and registration techniques that would not be possible to use in a real-time system.

Augmented reality systems have a variety of technical requirements. The fundamental technical requirements are a display, a way to align the real and the virtual

[1]This section is significantly derived from H Fuchs and J Ackerman. Henry Fuchs and Jeremy Ackerman. Displays for augmented reality: Historical remarks and future prospects. In *Mixed Reality – Merging Real and Virtual Worlds*, pages 31–40, 1999.

(tracking), and a source of data for augmentation. Most built augmented reality systems include graphics hardware, object and head tracking, hardware and software mechanisms for blending real and virtual objects, and specialized display devices. Displays will be emphasized because of the significance the choice of display mechanism has on other design considerations. With the exception of tracking, these considerations will be discussed in the context of display mechanisms.

3.1 Summary of Augmented Reality Display Technologies and Examples of Their Use

Typically optical see-through or video see-through displays are used in augmented reality. Head-mounted projector displays are also used. It is important to note that the conception of augmented reality systems should not be limited to applications using head-mounted displays. Two alternative display technologies, small area and large area displays, are described later in this section.

3.1.1 Head-Mounted Displays

Head-mounted displays have been used for augmented reality applications since the 1960s [Sutherland65], and have been the mainstay of many augmented reality systems to date. Two kinds of head-mounted displays are used: optical see-through and video see-through. As their name implies, optical see-through systems combine the real and synthetic imagery via some optical merging mechanism, most commonly a "half-silvered" mirror. Video see-through systems combine synthetic images with "real" images of the user's surroundings by combining two video streams, one typically generated by computer, the other coming from a video camera mounted on, or near, the user's head.

3.1.1.1 Optical See-Through Head-Mounted Displays

The first augmented reality system was developed in the late 1960s by Ivan Sutherland [Sutherland68], whose work was done first at Harvard University and then at the University of Utah. This display is shown in Figure 3.1.

Sutherland's display demonstrates the key components and principles needed for optical see-through displays. There is a distinct display for each eye, since each eye's view of the world is from a slightly different vantage point. The see-through

capability allows each eye to view simultaneously the user's surroundings and the computer-generated imagery. A half-silvered prism sitting directly in front of each eye allows some of the light of the user's surroundings to pass through to the user's eyes. Also passing through to the user's eyes is some of the light that originates at the face of the CRTs mounted on the sides of the head. Each eye observes the combination of two images superimposed one on top of the other. If light levels in the environment are carefully adjusted, and both the environment and the synthetic imagery are sufficiently simple, both can often be comprehended simultaneously.

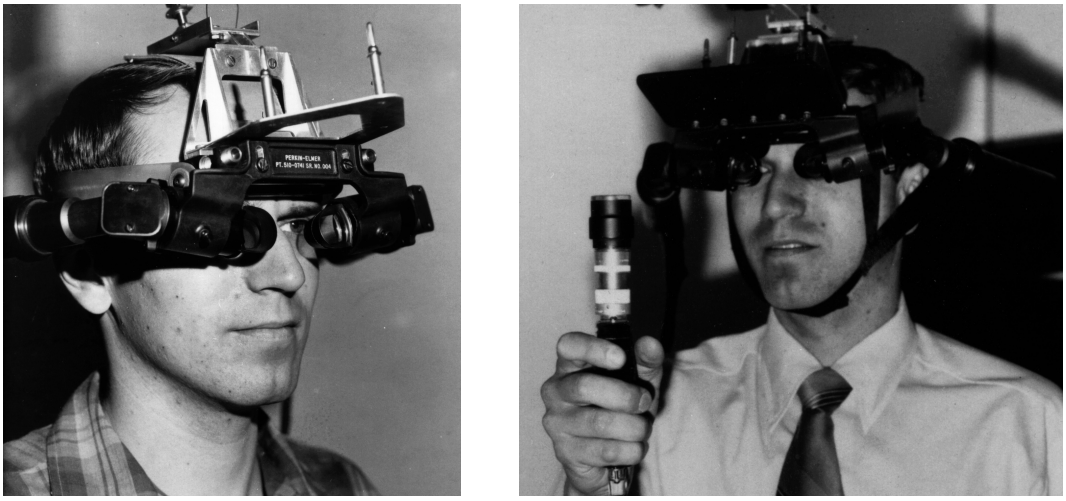


Figure 3.1: Ivan Sutherland's head-mounted display (circa 1970). Note the CRTs on either side of the head and the optics in front of the CRTs to put the image in front of the eyes. Images are viewed on half-silvered prisms, giving the user a view of both the synthetic imagery and the real world. (Courtesy of Department of Computer Science, University of Utah).

More recent implementations of optical see-through displays have utilized newer display technologies and novel optical systems to improve the performance of this type of head-mounted display. Other display devices, such as small LCD panels, can be placed directly between the viewer's eyes and the environment, and significantly reduce the complexity of the optics and the weight and cost of the device. Several displays based on this principle are commercially available. Many displays of this type were designed with other applications in mind.

Optical see-through displays have the very difficult problem of accurate registration of real and synthetic imagery in space and time. Changes in head position or orientation can occur very rapidly during a person's normal interaction with the surrounding environment. The combined processes of tracking (orientation and po-

sition), and rendering and display of synthetic imagery can easily take close to 100 milliseconds. By the time an image corresponding to the tracking data gathered at the beginning of the rendering cycle can be displayed, a person's head has already moved considerably. The image that is displayed is misregistered with the real world, which may lead to confusion, misinterpretation of the scene, and disorientation.

A second difficulty with optical see-through displays is light intensity. If the synthetic imagery is too bright relative to the ambient light, the user's environment will not be visible. If the reverse is true, the synthetic imagery will be either invisible or visible only as a vague shadow in the environment.

A final difficulty with optical see-through displays is that of occlusions. If a synthetic object should appear in front of a real world object, it will, in general, appear to be a semitransparent ghost floating in front of the object. If parts of the real world object are brightly lit, those portions of the synthetic object that are in front of the bright spots will appear more transparent or invisible.

An optical see-through display system manufactured by CAE is used in their simulators and trainers for military aviation (Figure 3.2). This head-mounted display is expensive and heavy, making it impractical for many other applications. It uses two light-valve type projectors for each eye. The central part of the user's visual field, the area with greatest acuity, is fed with one projector, while the larger surrounding, and lower acuity, region is fed by the second projector. An optical system is used to combine the light from each projector. This system provides greatest display resolution in the central area of vision where it is needed most.

In the CAE system, the use of a head-mounted display reduces the space required for a flight simulator to the size of the simulated aircraft's cockpit, without the additional space needed for other display techniques used by other manufacturers. A see-through head-mounted display is preferable to a virtual reality type display to enable the pilots to see the instruments and their own hands as they learn to operate an unfamiliar cockpit. This environment is ideal for the limitations of optical see-through devices discussed above. There is no reason why synthetic imagery of the environment that the aircraft is flying through should ever need to be overlaid on the real objects inside the cockpit. The simulator designers can design the lighting systems inside the cockpit to insure that both the real world and the synthetic imagery are always visible to the user.

Optical see-through displays were also a part of the demonstration of augmented reality air hockey given by Mixed Reality Labs at SIGGRAPH'98. In this demon-

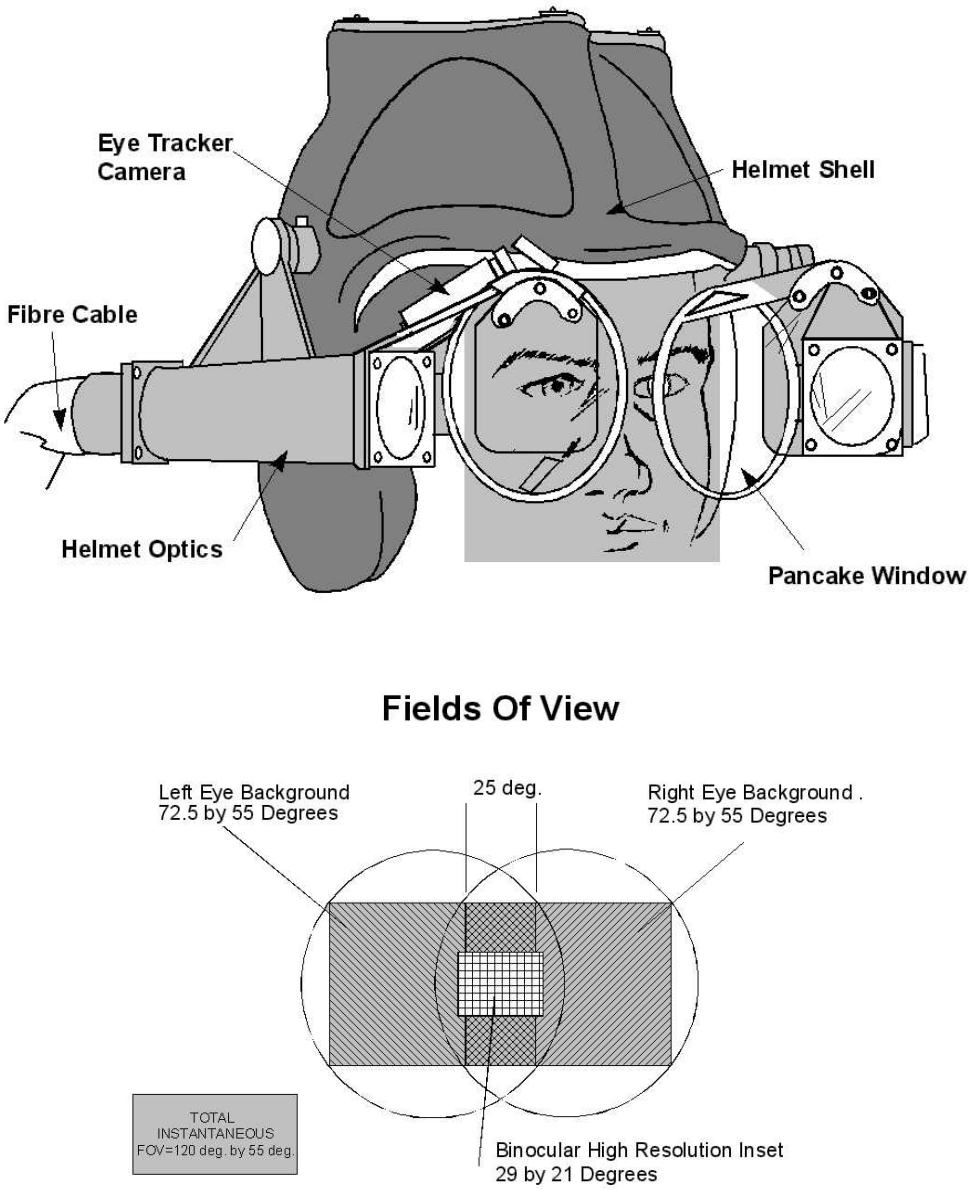


Figure 3.2: CAE HMD featuring high resolution foveal displays and low resolution peripheral displays. Top: schmatic, bottom: layout of high and low resolution displays in the visual fields. (images courtesy CAE)

stration, both optical and video see-through displays were available to conference attendees to play air hockey. The puck and paddles were rendered (registered to a real table) in the head-mounted displays of two players. The displays used were relatively lightweight and comfortable [Ohshima98].

Optical see-through displays can also be used for non-head-mounted displays. For example, "head-up displays" for automobile drivers, while currently limited to display of dashboard and infrared imagery for night driving, presumably could be provided with more sophisticated augmentation. Augmentation more similar to that done in optical see-through head-mounted displays can be achieved in almost any instrument where the user observes the real world through an optical assembly (e.g., microscope or binoculars). Augmentation through optical assemblies has similar problems of registration and alignment as in the case of a head-mounted system.

3.1.1.2 Video See-Through Head-mounted displays

A video see-through display gives the user a view of the world through one or more video cameras mounted on the display. Synthetic imagery is combined with the image captured through the video cameras by the computer, and the combined signal is sent to the display.

No video see-through displays are commercially available at this time. It is, however, relatively simple to mount one or two small video cameras on a standard head-mounted display designed for virtual reality (Figure 3.3) however, Additional hardware and software are needed to combine real and synthetic imagery. One approach is to employ a frame grabber for each camera. Alternatively, techniques like chroma-keying could be employed to add the video signal into the data stream after the synthetic imagery is rendered.

Video see-through also has a number of underlying difficulties. Among these are camera offset, limited field of view, temporal lag, fixed focus, and limited depth of field. When only a single camera is used, stereopsis is also lost. Difficulties such as camera offset and limited field of view have been overcome in an experimental display built collaboratively by the University of North Carolina and the University of Utah (Figure 3.4).

Without carefully designed optics, it is difficult to align the camera's view with the normal viewing axis of the eye. This means that the image sent to each of the user's eyes is taken from a perspective other than that of the observer's eyes. A significant change in horizontal spacing between the cameras distorts the user's sense

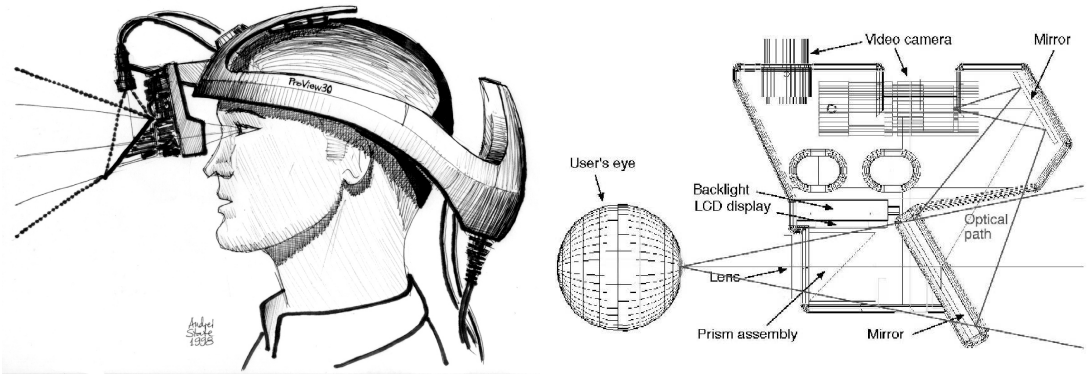


Figure 3.3: Two examples of ways to mount cameras on a video see-through so that the image of the world is reflected up to the camera so that the optical path length between world objects and the camera is very close to the path length to the user's eyes. The mirrors are moved with the displays when the displays are adjusted for optimal viewing. Right: modification of an commercially available display, Left: schematic of a custom-built video see-through display. (©1998, Department of Computer Science, UNC-Chapel Hill)

of depth because the stereo pairs have an effective interpupillary distance different from that to which the user is accustomed. Typically it is easiest to mount the camera or cameras vertically above the user's eye. This displacement gives the user a false sense of height, particularly for relatively close objects.

Another problem with video see-through is that of field of view. Care must be taken in the selection of cameras and lenses so that the field of view of the camera and lens combination matches the field of view of the displays. Without this precaution, the user's view of the world will appear noticeably distorted.

Like optical see-through displays, video see-through displays suffer from temporal lag. However, it is possible to compensate for lag, such that the effects of lag are not noticeable in a given image being displayed by the head-mounted display. This can be achieved by synchronizing video image capture with the report from the tracker. The synthetic imagery generated from that tracker data must be displayed with the video image captured at the same point in time. This mechanism of dealing with the lag inherent in the system solves the misregistration problem, but may introduce noticeable delays between when movements in the real world occur and when they are seen.

Fixed focus and depth of field are problems inherent in using video cameras. Objects outside the focal distance of the cameras will be out of focus, and typical cameras have a relatively narrow depth of field. While it may be possible to use auto-

Figure 3.4: Photograph of a video see-through head-mounted display built in collaboration with the UNC-Chapel Hill and the University of Utah. This display features a relatively narrow field of view for the displays, but has an open design allowing unencumbered vision around the display. The displays and cameras are mounted on independently adjustable pods to optimize viewing performance. The camera view is aligned with the user's eyes by optics which fold the optical path. (©1998, Department of Computer Science, UNC-Chapel Hill)

focus type mechanisms to insure that objects seen by the camera are in focus, changing the focus may change the camera model significantly and lead to misregistration of synthetic imagery.

The problem of environmental lighting faced by optical see-through systems is not a concern in video see-through systems. Lighting must be considered when selecting cameras and camera settings, as cameras operate optimally on a much narrower range of illumination than can the human eye.

Occlusion poses a challenging problem in video see-through head-mounted displays. When a real object passes between the user and the location of a virtual object in space, the synthetic object will typically, and incorrectly, fully occlude the real object that it should appear to be behind. This confusion of depth cues can destroy the illusion that the virtual objects truly exist in the user's environment. In some applications moving objects may be tracked with position sensors, or another method may be employed to determine the three-dimensional geometry of the real

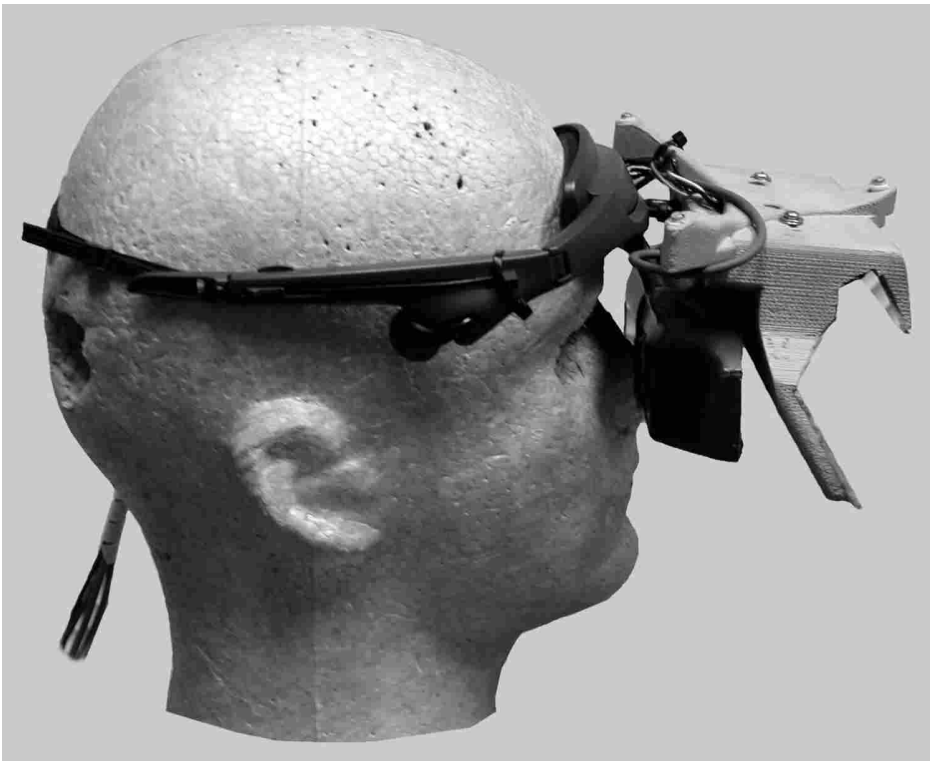


Figure 3.5: Photograph of a modified Sony Glasstron used in a medical augmented reality system. The camera view is approximately aligned with the user's eyes by optics which fold the optical path. The LEDs used for tracking are integrated into the design in a position designed to optimize their exposure to the tracking detector (©2002, Department of Computer Science, UNC-Chapel Hill)

world. Without this type of three-dimensional information about real objects, this type of occlusion error is difficult, if not impossible, to solve.

At the University of North Carolina, video see-through head-mounted displays have been used extensively for the medical applications being developed [State96a]. Video see-through displays were chosen because of the ability to synchronize synthetic imagery with the video image from the real world, thereby reducing the likelihood that any individual image shown to the user would misrepresent the spatial relationship between real and virtual objects. Our system has used many different displays, including those we were involved in designing (Figure 3.4), those that were built by re-engineering commercially available displays, and those made simply from commercially available displays (Figures 3.3 and 3.5).

Our system has used several displays originally designed for use in virtual reality applications. The wide field of view displays that these systems typically have, while

helpful for a sense of immersion in virtual environments, do not provide adequate resolution for the surgical site at the center of the user's view. Our medical collaborators have also been unhappy with the inability to use peripheral vision and to see around the display easily. Our own design (Figure 3.4) and our most recent redesign of a commercially available display (Figure 3.5) solve these two problems by leaving the display open laterally to the display for each eye and by providing relatively narrow field of view displays.

Video see-through augmentation may also be applied to other situations. Augmentation through microscopes, endoscopes, and other tools where the user views images of real objects via a video signal are possible.

3.1.2 Head-mounted projector displays

An alternative method of displaying the combination of real and virtual objects is to project the virtual object onto real objects. Application of this technique to large areas making use of fixed projectors is discussed in Section 3.1.3.2. An alternative approach is to have the user wear projectors on his or her head.

Rolland [Gao01, Hua00, Argotti02] uses this technique in some of their systems. Projectors mounted on the user's head are optically aligned with the wearer's eyes using a half-silvered mirror. Real surfaces to be augmented must be coated with retro-reflective material. When images from the head-worn projectors strike the retro-reflective material, the imagery is directed directly back at the projector and passes through the half-silvered mirror to the wearer's eyes. Retro-reflective surfaces are extremely efficient so that augmentation can have achieved in a lit room and the projectors need not be extremely high intensity. Other observers in the room will not see the projected imagery, and multiple users can see their own, perspectively correctly rendered stereoscopic imagery using the same retro-reflective display surface.

Under the assumption that virtual objects are "under" the retro-reflective display surface, occlusion conflicts between real and virtual objects are easily resolved. An object that is not coated with retro-reflective material placed between the user and the retro-reflective surface will always fully occlude the virtual object. Displays of this type have currently only been built for lower resolution viewing (640x480) but higher resolution displays are currently feasible, and current designs are equivalent in size and weight to many head-mounted displays.

The requirement of retro-reflective display surfaces has both advantages and disadvantages. As long as a very irregular display surface is totally coated with retro-

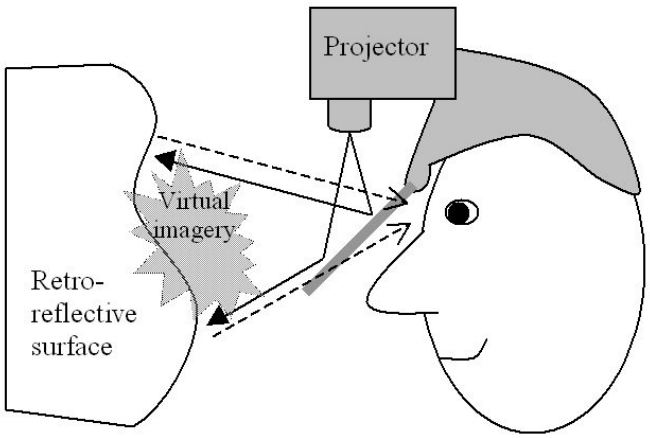


Figure 3.6: This figure shows a schematic of a head-mounted projector display. A pair of low-intensity video projectors are mounted on the user's head and reflected off a half-mirrored surface so that the center of projection is aligned with the user's eyes. When this light hits the retro-reflective surface the rays are coherently (that is the image is not lost) directed back to the user's eyes. Retro-reflective materials are extremely efficient so the projected images do not have to be visible to others, nor is a high-powered projector needed.

reflective material, the whole surface can be used as a display surface without any need for calibration. This can be employed usefully to hide portions of input devices (e.g., coat a Phantom haptic device with retro-reflective material so that portions of it are not visible while manipulating virtual objects) and display surfaces that would be less than ideal for other display modalities. The disadvantage of using retro-reflective material is that it constrains the locations of virtual objects in an augmented reality system to being "behind" coated surfaces. If large areas of retro-reflective material are available, virtual objects can be shown in front of the surface, but with a loss of the ability to correctly handle occlusion.

3.1.3 Non-head-mounted Displays

There are some environments and situations where wearing a head-mounted display is impractical or unfeasible. This may be true of situations requiring long term use, or extensive use of other resources in a non-augmented environment. The following subsections describe a few approaches to other types of displays.

Systems of this type provide the benefit of offering the higher resolution of a conventional monitor for augmented reality applications. The user also does not

experience the fatigue and discomfort of wearing a heavy and unwieldy head-mounted device. However, these devices are limited in their usefulness by their limited working volume and by potential space constraints.

3.1.3.1 Small Displays

Displays the size of a small monitor (approximately 30cm by 20cm) can be employed to build an augmented reality system. One approach was proposed in 1975 by Knowlton [Knowlton75] as a way of creating a programmable keyboard. In his system, a real but unlabeled keypad was to be augmented with button labels drawn by computer graphics and a half-silvered mirror was placed between the user and the keypad. The half-silvered mirror provides the user a view of the real objects behind the mirror, and also reflects the image from a conventional display (such as a CRT) to the user. This type of system can be thought of more generally as placing arbitrary real objects behind the mirror and adding appropriate synthetic objects.

Devices of this type can be further enhanced in several ways. If the user's head position is tracked, synthetic imagery can be drawn from the correct perspective. This device can be used to reflect alternating left and right stereo pairs rather than single perspective images. A stereo illusion of dimensionality can be created if the user wears shutter glasses when viewing the display [Knowlton77].

Displays of this type face a problem with illumination similar to the one encountered in optical see-through head-mounted devices. Real objects behind the mirror will wash out the synthetic image if they are more brightly lit than the illumination provided by the monitor. The optical see-through problem of not being able to fully occlude real surfaces with synthetic imagery is also present with these devices.

These displays have a working volume limited to the space behind the mirror. Generally the mirror needs to be maintained at a fixed geometry relative to the conventional display and is, therefore, not easily moved to a new position.

In situations where the user may want to move very close to the mirror, the display may fail in two ways. First, the user may obstruct the path of the light from the monitor to the mirror, thereby blocking the light path needed to view the synthetic imagery. Second, as the user approaches the display, each pixel subtends a greater portion of the user's field of view. As a result, the user can gain no better enhancement of their view of the synthetic objects by moving their head closer to where they perceive the object to be.

This type of display is currently being used in surgical simulators so that the

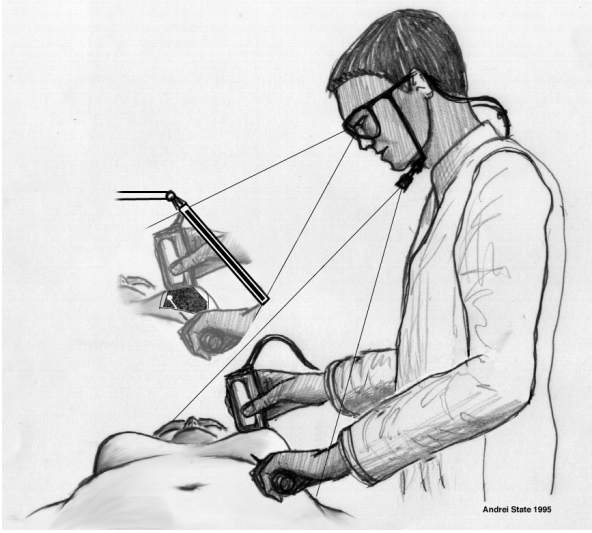


Figure 3.7: One method of how a relatively small display might be employed for use in an augmented reality medical system. In this example a strategically placed camera enables video see-through augmented reality. (by Andrei State ©1998, Department of Computer Science, UNC-Chapel Hill)

surgeon's hands are visible, and in some telesurgical devices under development. Figure 3.7 shows the proposed use of these types of displays in UNC's surgical augmented reality applications.

In the field of neurosurgery another novel approach to augmented reality is used. Preoperatively acquired images, including three-dimensional graphics created during surgical planning, may be overlayed onto the surgeon's view through the surgical microscope. A system of this type is now commercially available to assist stereotactic procedures in neurosurgery [COMPASS].

3.1.3.2 Large Area Displays

Large displays have perhaps received less recognition than they deserve for augmented reality applications. These displays may be employed in many configurations: front projection, back projection, and even conventional CRTs. Use of larger displays in augmented reality is perhaps best explained by example, because they are not usually thought of as augmented reality display devices.

A display like an ImersaDesk or Virtual Workbench is usually a large monitor, set on either a horizontal or tilted work surface. A user wears shutter glasses so that the display can show three-dimensional objects in stereo. A variety of input devices can be employed to interact with the objects being displayed. These displays are limited

to displaying objects near their surface. A virtual object can appear no "taller" than the plane from the far edge of the display to the user's eyes [Kalawsky93].

The large monitors and projector systems in civilian flight simulators may also be thought of as augmented reality devices. These displays augment the real environment of a large jet's cockpit with images of what the outside world should appear like.

Projects like UNC's Office of the Future [Raskar98] (Figure 3.8) take the previous two concepts a step further. In the Office of the Future, multiple projectors will be used to provide a virtual window into another Office. In addition to providing a compelling telepresent experience, workers in an Office will be able to collaborate on three-dimensional models both locally and with remote participants. Large area displays, unlike the head-mounted technologies, allow a user to easily switch back and forth between making use of the virtual objects and taking or reading notes, or using a conventional monitor.

Figure 3.8: A view into the Office of the Future. Large displays (provided by the projectors mounted in the ceiling), cover the walls and desk. The user participates in a discussion with his colleagues who appear to him to be sitting on the other side of his desk. He is able to use his personal computer to make notes while working collaboratively on a design for a new head-mounted display, which appears to be hovering over his desk. (by Andrei State ©1998, Department of Computer Science, UNC-Chapel Hill)

3.2 Tracking for augmented reality in small areas

Tracking is an important component of most augmented reality systems. Tracking is the collection of methods and equipment used to find the location and orientation of real objects and the user in an augmented reality environment. Many of the same techniques that are used for user interaction for virtual reality applications are used here.

Tracking is the key component that gives the system enough information about the real world to add virtual objects in the correct place. Tracking systems can take a wide variety of forms depending on the application. For instance the head-up targeting displays for fighter pilots might use radar to track the position of enemy planes and use the radar position to highlight the plane on the display. A system for displaying directions and labeling buildings for a truck driver might use a global positioning system (GPS) device to track the location of the truck and a computer vision-based system to track the driver's head[Welch01].

The most typically used head and instrument tracking methods for augmented reality in applications in a small working volume are mechanical, magnetic, infrared triangulation, and computer vision. Other methods used include accelerometer and ultrasonic tracking.

3.2.1 Mechanical tracking

Mechanical tracking systems typically work by attaching a segmented arm to the object to be tracked. Encoders, placed at each joint, report the current position of that joint. The position and orientation of the tracked object can then be calculated from the positions of the joints.

Mechanical tracking systems can be extremely accurate and offer good refresh rates. However, mechancial tracking devices can be difficult to use for some applications because of the mechanical tether. The arms can get in the way of some motions and they frequently have mechanical stops to prevent movement beyond the range of the encoders. For medical applications, mechanical trackers add challenges to sterilization if the tracked object is going to be in or near the sterile operative field.

3.2.2 Magnetic tracking

There are several different techniques that can be used for magnetic tracking devices. The general principle is that one or more emitters creates a magnetic field and a receiver attached to the object to be tracked detects characteristics of the field and uses these characteristics to calculate the receiver's position within the field. These tracking systems may not be accurate because the magnetic fields can be distorted by nearby ferromagnetic metals and electronic devices. In general, these devices are more accurate closer to the transmitter where the fields have the least distortion and are strongest.

Magnetic trackers are widely available and relatively inexpensive. While they need a wire connecting the receiver to a controller, the receivers can be sterilized. Unlike other tracking methods, magnetic methods do not require a clear line-of-sight between the emitter and the receiver. These devices are sometimes used for medical applications.

3.2.3 Computer vision based tracking

A large number of techniques are available within the framework of computer vision tracking. These methods identify features in images acquire from one or more cameras and use those features to gather information about the location of objects. Computer vision techniques can be used for identification of objects as well.

In video see-through augmented reality these techniques have been used on the video images to determine position and orientation of the user's head. These methods can be used as part of a closed feedback loop to avoid the "jitter" of synthetic imagery typically seen with other tracking methods because there is no reliance on an external tracking device. Typically several easily identifiable fiducial markers need to be scattered around the working area for these systems and the user needs to keep several markers in view at all times. If an insufficent number of markers are visible, tracking is no longer possible.

3.2.4 Infrared markers tracking

Use of infrared markers for tracking is attractive because infrared is not visible to the human eye, but is usually visible to CCD cameras and analog sensors. Infrared markers can be used as highly visible fiducial markers.

There are a variety of ways the infrared markers are used. There are active methods where infrared sources are sequentially lit and passive methods where retro-reflective markers send strong signals to cameras positioned near infrared sources. Another type of infrared-based tracking is outside-in systems where cameras look inwards for markers and inside-out systems where cameras on the tracked object look outwards for markers. All of these systems require line-of-sight between the cameras and the markers although redundant cameras and markers can be useful for lessening the chance that the system will be unable to detect markers.

Active systems illuminate infrared markers, typically LEDs, sequentially and use several detectors to calculate the position of each marker in turn. At least three markers are needed on any tracked object to determine its position and orientation. Passive systems find all of the visible markers from several different camera views and then attempt to determine which markers it found correspond to the markers seen from another camera's view. Both passive and active infrared systems do not necessarily require a wired tether.

Infrared tracking systems are also currently used for medical applications. They have very high accuracy, but the line-of-sight requirement can make them difficult to use. Spatial localization of markers in active systems can be done to approximately 0.1mm, with angular error depending on the distance between markers. These systems have a working area that is limited by the field of view of the cameras and the power of the infrared sources. Accuracy decreases as markers are moved away from the cameras.

3.2.5 Ultrasonic tracking

Ultrasonic trackers use an ultrasound source attached to the tracked object and several microphones to determine the position. These systems usually use time-of-flight to each of the microphones to triangulate position, but some systems use phase differences. These systems are often slower than other methods because they depend on the speed of sound. These systems have line-of-sight constraints similar to infrared-based systems because many solid objects reflect ultrasound energy. These systems may also be affected by noise in the environment and hard surfaces that might efficiently reflect ultrasound energy.

3.2.6 Accelerometer tracking

Accelerometers can be used to track the position of an object by integrating the accelerations the object undergoes over time to determine the velocity of the object over time and then integrate the velocity. The orientation of a static object may be determined by measuring the effect of gravity on several differently oriented accelerometers. Gravity makes tracking a moving object more difficult because it must be factored out of the signal from each sensor which can be very difficult for a moving object. Any errors in filtering the effects of gravity are compounded over time when using accelerometers so this method is most often used when other methods would be impractical or when updates in position are needed between samples from a more slower, but less error prone system.

Chapter 4

Medical augmented reality

"Surgery will not only develope new and previously impossible procedures, but all uncertain operations which depended on luck and approximation will become safe under the influence of direct vision, since the surgeon's hand will now be guided by his eyes"

– Philip Bozzini, 1806

The dream of many who apply technology to surgical problems is well summarized by Bozzini in 1806. New advances in imaging and computer technology raise the possibility of something that Bozzini likely did not imagine, that surgeons would some day have tools that let them see more than what can be seen by the eye or with any optical instrument. This chapter will explore how new technologies have enhanced what a surgeon can see, how augmented reality now allows surgeons to use this new vision to guide their hands, and will summarize relevant previous work in medical augmented reality and related technologies.

A wide range of new technologies and techniques exist that, despite use of computers, fit a more traditional mode of use. These include things like "virtual colonoscopy" and intervention planning software that make use of sophisticated image processing and modern computer graphics. The distinguishing characteristic between these systems and those discussed here is not the level of technology that is used but instead it is the degree of interaction at the time of intervention.

Stereotactic surgery, that is localization of targets via a reproducable three-dimensional coordinate system, was introduced in 1906 for the study of structure and function in monkey brains. Variation between individuals meant that these systems were only reliable if target structures were usually near the landmarks. In the 1940s Spiegel and Wycis began using x-ray imaging to visualize the ventricals of the brain and location of the pineal body and used this information to more precisely locate structures via preoperative and intraoperative imagery [Kelly00].

The field of interventional radiology, the use of intraoperative radiological images to guide therapeutic interventions, is regarded to have begun in 1964. This represents a significant change in the view of radiological imaging in that the images were for the first time being used to directly guide the procedures and provide feedback to the procedure in progress. Interventional radiologists do a wide variety of procedures making use of real-time imaging, but they largely use traditional display mechanisms that show only the intraoperative image without the benefit of augmentation.

A variety of new technologies – display, automated segmentation, fast computing, and object tracking have given rise to a wide range of improved ways of displaying medical imagery.

Stereotactic surgery has developed in a number of significant ways. One advance has been the development of "frameless" stereotactic surgery. In these systems the position of the patients and instruments is electronically tracked so the procedures can be performed without a rigid frame to use as a reference. When instruments are tracked, this information can be displayed to the surgeon by showing the tracked position of the instrument on preoperative imagery. Further improvement can be made when the patient is imaged while the the tracked instrument is in place. This allows the surgeon to observe the changes between preoperative imagery and structure as it appears intraoperatively. In neurosurgery, craniotomy, opening the skull, can cause significant distortion of the brain making preoperative imagery difficult to use. New image processing techniques allow rapid segmentation and creation of three-dimensional models of anatomical structures from intraoperative scans. Other image processing techniques hold the promise of automatically distorting preoperative scans, perhaps by using another imaging modality, to match the structure imaged interoperatively. All of this augmenting information, including the source images, can be displayed on a computer monitor similar to what a physician might use to examine a CT or MRI scan.

Commercial and research products exist for assisting guidance of procedures using CT, MRI, and ultrasound. Once true three-dimensional information is available, it can be displayed in many other, potentially more useful ways.

4.1 Augmentation of the surgical microscope

In many procedures, particularly in neurosurgery, surgeons use optical microscopes to get a magnified view of the operative site. Three-dimensional information can be

Figure 4.1: An augmented surgical microscope in use in an operating room at Guys Hospital, London. (image courtesy David Hawkes)

added to these devices. Initial efforts, now used clinically, added manually created three-dimensional objects from preoperatively acquired scans to the surgeon's view through the microscope. These systems have now evolved into commercially available augmentation systems for the neurosurgical microscope[COMPASS, Rousu98, Alterman97]. Continuing research is examining the ability to add objects automatically segmented[1] from intraoperative scans [Edwards00, Edwards99, Edwards95, Jannin00].

These systems have been built as both optical and video see-through devices. The surgical microscope is tracked. Mechanical trackers are sometimes used for tracking in these systems because the surgical microscope, a large heavy device, is already mounted on a counter-balanced arm. Others use infrared tracking with widely spaced markers to achieve extremely high rotational accuracy. preoperatively acquired and segmented images are most frequently used for augmentation. Use of intraoperatively acquired and segmented images or preoperative images updated using intraoperative methods are being developed as additional sources for augmentation. An image of an augmented surgical microscope is shown in Figure 4.1.

Another area where augmentation of a surgical microscope is being done is in

[1]Segmentation is the process of identifying pixels or voxels in an image or volume as belonging to different object. In neuroimaging segmented images might be ones in which a tumor and key anatomical structures have been identified enabling a three-dimensional volumetric reconstruction of those features.

opthamology. These systems are used conventionally to examine and operate on a patient's retina. preoperative planning and images can be overlaid onto the live image of the patient's retina. The retina is a mobile, but essentially two-dimensional surface, so tracking is done by using computer vision methods to identify common landmarks and align the preoperative data with the current view[Berger01, Berger99].

Augmentation of a surgical microscope has a potential advantage over more conventional data displays in that the benefits of preoperative and intraoperative data do not require an additional display. The information is shown in the context of the patient without surgeons needing to look up from their work.

4.2 Augmentation of the operative field

One clinically used system allows the surgeon to see an augmented view of the patient on a display located near the operative field[Gering01, Grimson99]. This system is video see-through, but not head-mounted. A calibrated camera mounted at a known position provides the video image, and the patient's position (frequently situated in a scanner) is calibrated by a laser scanner. The view through this system is illustrated in Figure 4.2. Use of a fixed camera greatly reduces the demands on tracking because the surgeon's head is not tracked.

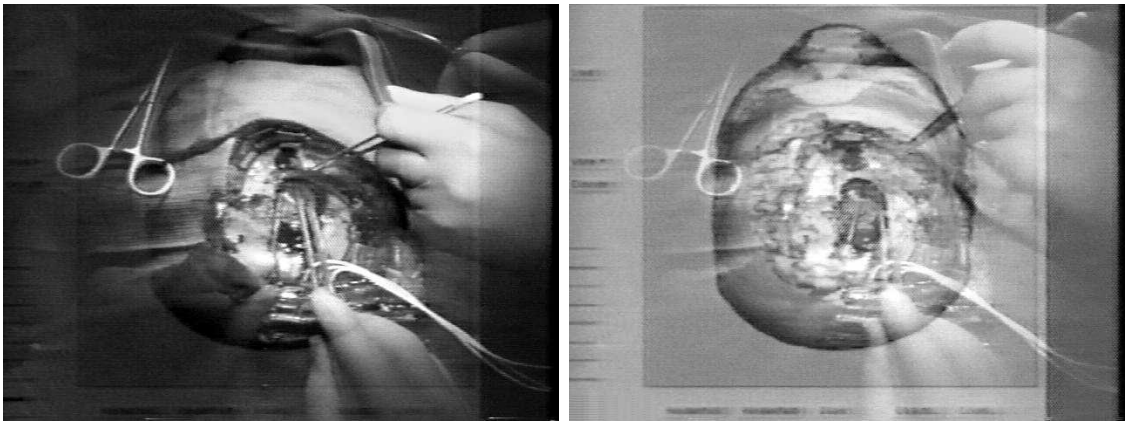


Figure 4.2: Right: raw video image from a camera placed near the operative field. Left: preoperative planning data, a segmented structure of interest, registered and overlayed with the intraoperative view. (Courtesy Surgical Planning Laboratory, Brigham and Women's Hospital)

At UNC a considerable volume of work has focused on ultrasound-guided needle breast biopsy. Similar research is also underway at Siemans Research. In these

systems, the image from a conventional ultrasound probe is shown correctly positioned in three-dimensional space on the patient. The UNC system, now in clinical trials, uses a video see-through head-mounted display. The physician's head, the ultrasound probe, and, in some instances, the biopsy needle are tracked using an active infrared tracking system. A model of the patient is acquired by mechanically sweeping with a tracked probe before the procedure is started. Using this system a radiologist trained to perform these procedures with augmented reality was able to place a biopsy needle 0.86mm more accurately [Rosenthal01].

A guidance system to guide the resection of non-palpable tumors has been tested in phantoms as well [Sato98]. This system uses a 3-D ultrasound probe (mechanically swept) to acquire a three-dimensional dataset. This data can then be overlayed on the patient allowing the surgeon to see non-palpable tumors before attempting resection.

A simpler device can be used to achieve similar effect, This device, developed by George Stetten is called the "Sonic Flashlight". It uses a small display and a half-silvered mirror. The display shows the video image from the ultrasound machine. The position of the mirror and display is selected so that the virtual (in this case reflected) image of the display is visually aligned with the objects generating the ultrasound image [Stetten00]. This system has the benefit over more conventional augmented reality systems in that there is no degradation of the user's view of the real world and there is no latency or tracker error. No tracking is necessary as long as the mirror is correctly positioned relative to the display.

An alternative method for displaying intraoperative augmentation is to project imagery directly onto the patient. Researchers at the University of Karlsruhe have developed systems to project preoperative planning data onto patients. These systems use a combination of a video projector and two video cameras to determine the surface geometry of the patient and correctly project the planning data. Deviations from the preoperative plan may be reported by such a system when comparision of the plan and the actual geometry of the patient are made. Accuracy of $\pm$1.5mm and refresh rates of 1.6Hz are reported for this system [Hoppe01a, Däuber02].

4.3 Augmentation for education

Medical educators are interested in augmented reality because it can be used to allow students to see internal structures relative to external landmarks. One such system is being developed by Rolland. In this system a "patient" wears a sleeve on their

Figure 4.3: Medical augmented reality research at the University of North Carolina. Top left: lab view of a cyst aspiration being performed using augmented reality guidance. Top right: view through the physician's head-mounted display. The ultrasound "slice" is drawn correctly registered to the patient inside a virtual hole drawn on the patient. Bottom: view through a physician's head-mounted display in a mock-up system of augmented reality for laparoscopic surgery. Live video textures are rendered onto pre-acquired, mechanically scanned geometery of the organs inside the anatomical model.

leg with multiple infrared tracking points. As the patient moves, the location of the tracked points are used to modify the position of the bones in a computational model of the knee. The user is then able to see bones "through" the patient's skin with proper articulation as the patient moves [Rolland97, Baillot00]. This system uses a bench-mounted optical see-through display so tracking is limited to the position of the patient's leg.

Rolland is also developing a system to demonstrate correct internal anatomy on mannequins used for training military medics in advanced life support skills. This system uses a head-mounted projecter display and the user's head is tracked. Detailed anatomical models are displayed, correctly rendered to the mannequin, to illustrate to students what they are trying to achieve during endotracheal intubation [Rolland].

Another educational application of augmented reality is to teach students how to conduct an ultrasonic cardiac examination. In this system the user is given a physical model of a patient's chest and a magnetically tracked ultrasound probe. A conventional video display is used to show the user both the ultrasound imagery corresponding to the current probe position and the position of the ultrasound "slice" relative to an anatomical model [Weidenbach00].

Chapter 5

Simulation of an idealized surgical augmented reality system

"Although our eyes may mislead us more than the other senses, the optical illusion is in most cases negligible when compared to the real image."

– Philip Bozzini, 1806

5.1 Motivation

Investigation of the usability and usefulness of future surgical augmented reality systems is difficult because of the known inadequacies of current systems. It is too easy to dismiss poor results, or even not to have an expectation of getting good results, because of the known failings of existing components. An evaluation of augmented reality which is independent of the limiting technologies is therefore desirable.

Current systems suffer from a large number of known technological issues. These problems include insufficient display and camera resolution, poor tracking accuracy, misalignment of the head-mounted display, and systemic latency[Holloway95].

The *Oxford English Dictionary* defines simulation as "The technique of imitating the behaviour of some situation or process (whether economic, military, mechanical, etc.) by means of a suitably analogous situation or apparatus, esp. for the purpose of study or personnel training."

In the context of research within computer science, the term "simulation" most frequently brings to mind numerical simulation or virtual environments. In this instance, simulation of an augmented reality environment with a virtual environment is not likely to be useful because the visual effect would ultimately be dependent on the same technologies needed to build an actual augmented reality environment. Numerical simulation is not a realistic option because the value of simulation, in this

case, is to measure how useful the augmented environment is to people, something which cannot be numerically simulated.

For some of the known technological inadequacies, head mounted displays for example, a much improved (and more highly specialized) system component could be designed and built based on technology available today. For others, tracking for example, currently available technologies do not appear to adequately meet the challenges of augmented reality. Still others, such as high speed depth extraction, appear to be promising but not fully developed technologies whose availability may be imminent.

The tasks of investigating surgical augmented reality systems are therefore twofold; to emphasize the potential benefits of the new technology while remaining vigilant against assuming components will be unreasonably good. A secondary task of this investigation is to identify realistic future roles for the technology. While the greatest benefit of augmented reality may well occur as the result of visualization of other data sources (data fusion, e.g., real-time imaging), only the role of visualization with respect to hand-eye coordination is addressed here. Later studies might examine the benefits of data fusion when techniques to examine that become available.

In keeping with these goals, two terms can generally be defined:

> *Simulated Augmented Reality*: An environment which simulates an idealized augmented reality system.
>
> *Degraded Simulated Augmented Reality*: A simulated augmented reality environment that has purposefully been degraded to simulate having one or more less than ideal components.

5.1.1 Simulated augmented reality environment

In concept, an augmented reality view of a minimally invasive surgery should appear nearly identical to an open procedure. The view of the operative site should be unobscured, and nearby structures and organs visible. Hidden structures can be found easily using external (superficial anatomical) landmarks as well as visible internal landmarks. Ideally, the main difference between minimally invasive surgery and open surgery would be the tools used to complete the task.

In reality, it is possible to predict a number of further limitations that even an *idealized* surgical augmented reality system would have. These limitations can be broadly divided into those that cannot be overcome based on currently available

technology and those whose limitations depend on the quality of technological components,

The following limitations appear to be inherent and fixed based on current technology:

- Limited field of view of the surgical site because of the need to rely on a laparoscope for that view.
- Only the front surface, relative to the laparoscope, of structures can be visible

The following limitations have an effect that will vary depending on the quality of available technology:

- Limited image quality of visible surfaces due to the resolution of the acquisition device.
- Limited image quality in the display of visible surfaces.
- Errors in range (depth extraction) information.
- Errors in tracking

Beginning with the idea of creating a physical object to create this simulated augmented reality environment, one method would be to modify a laparoscopic trainer. A laparoscopic trainer may be a carefully constructed device or it may be as simple as a cardboard box. In either case, objects to be worked with are placed inside an opaque construction. Two commercially available laparoscopic trainers are pictured in Figure 5.1. Video cameras (sometime actual laparoscopes with cameras attached) are placed inside this "box" so the trainee can see it on a video monitor. Small openings in the box allow the introduction of laparoscopic surgical instruments so that the trainee can practice basic techniques.

The simplest simulated augmented reality environment could be created by simply creating a large hole in a laparoscopic trainer, or removing the top, through which the trainee can directly observe their work. A more reasonable, yet still idealized, simulated augmented reality environment should simulate the effect of as many of the inherent limitations as possible. With few exceptions each of the technological limitations can potentially be simulated by physical means in a modified trainer. Simulation of these additional effects form the basis for the creation of a degraded simulated augmented reality environment. These are discussed in greater detail in

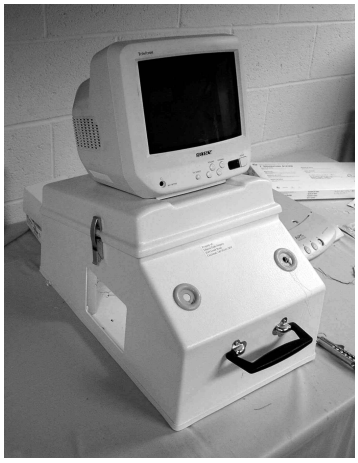


Figure 5.1: Two examples of commercially available laparoscopic trainers. The image on the right is of a transportable video trainer which features a fixed video camera mounted inside the trainer and a monitor integrated into a samll, portable package. The image on the left is a Pelvi-Trainer^{TM} which must be used in conjunction with a laparoscope and a video tower.

Sections 5.1.1.1 through 5.1.1.5. Figures 5.2 through 5.4 demonstrate some of these effects as they appear to a user of the trainer.

5.1.1.1 Limited field of view of the surgical site

In a buildable augmented reality system, there are two reasons why the field of view of the operative site is limited. The first occurs because of the limited field of view of the instrument acquiring the view of the operative site, the laparoscope. The second reason is the possibility of the synthetic opening into the patient, the "pit", not being sufficently large to allow a full view of everything that the laparoscope can see.

A simple way to simulate the effect of the limited field of view of the is laparoscope to place a piece of half-silvered glass between the user and the objects being examined. Under normal use, the illumination channel of the laparoscope delivers high-intensity lighting. In a lit room, this illuminated area is visible through the half-silvered glass while areas with less illumination are not visible. Because the illumination channel is designed to give diffuse light, it lights a much larger area than the field of view of the laparoscope with a gradual fall-off in intensity. The optical channel, normally used to see inside the body, can be used to project light to exactly the field of view of the scope and with a sharp cut-off. Alternatively, a hood can be attached to the tip of the laparoscope to give a sharper cut-off and to limit the field of view of the simulated depth extracting laparoscope. The effects of these techniques are shown in

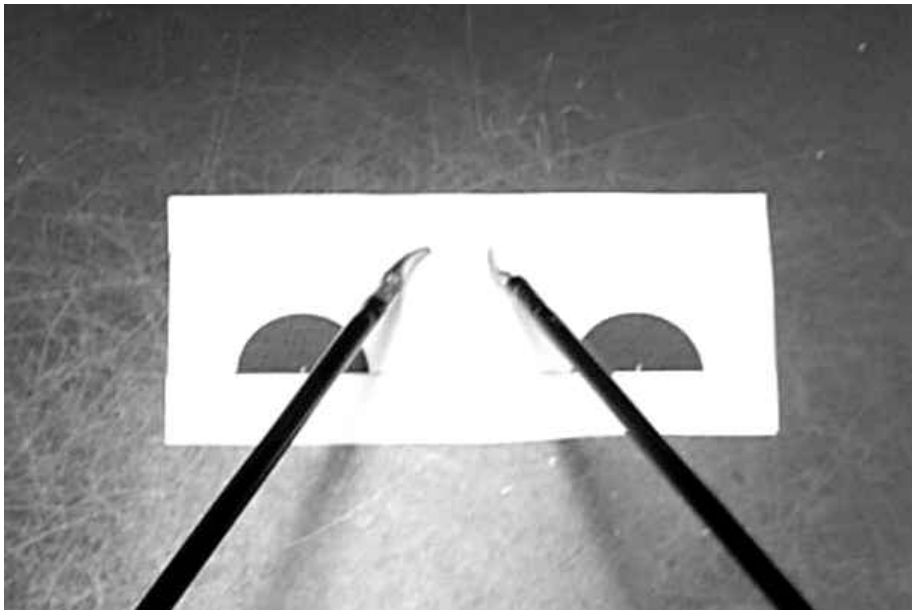


Figure 5.2: This picture shows the trainer with the lid removed. The inside is painted with matte black paint to prevent reflections inside the box from illuminating areas away from where the scope is pointing.

Figure 5.3.

Using either method, care must be taken to reduce the level of ambient light inside the trainer. If too much light is available inside the trainer, relative to ambient room light, areas within the trainer other than those directly illuminated by the laparoscope will become visible to the user. This can be prevented by controlling external light and by controlling internal reflections. External lighting can be controlled by keeping holes in the trainer as small as possible. Internal reflections can be controlled by making sure that objects inside the trainer are primarily dark colored and have a flat or matte finish.

The field of view restrictions due to the size of the synthetic "pit" may not be inherent in all future augmented reality systems. The "pit" has been a useful part of several test systems and should be considered. In the work on ultrasound guided breast biopsy, the pit is usually placed in a position that maximizes the view of the operative site, but it is also kept as small as possible so that the view of the real world is relatively unobstructed. While the pit can be moved and resized dynamically, in practice the location and size of the pit is usually selected early and then is not changed for the duration of the procedure. For the purpose of simulation, a small opening into the trainer is a stand-in for the synthetic pit.

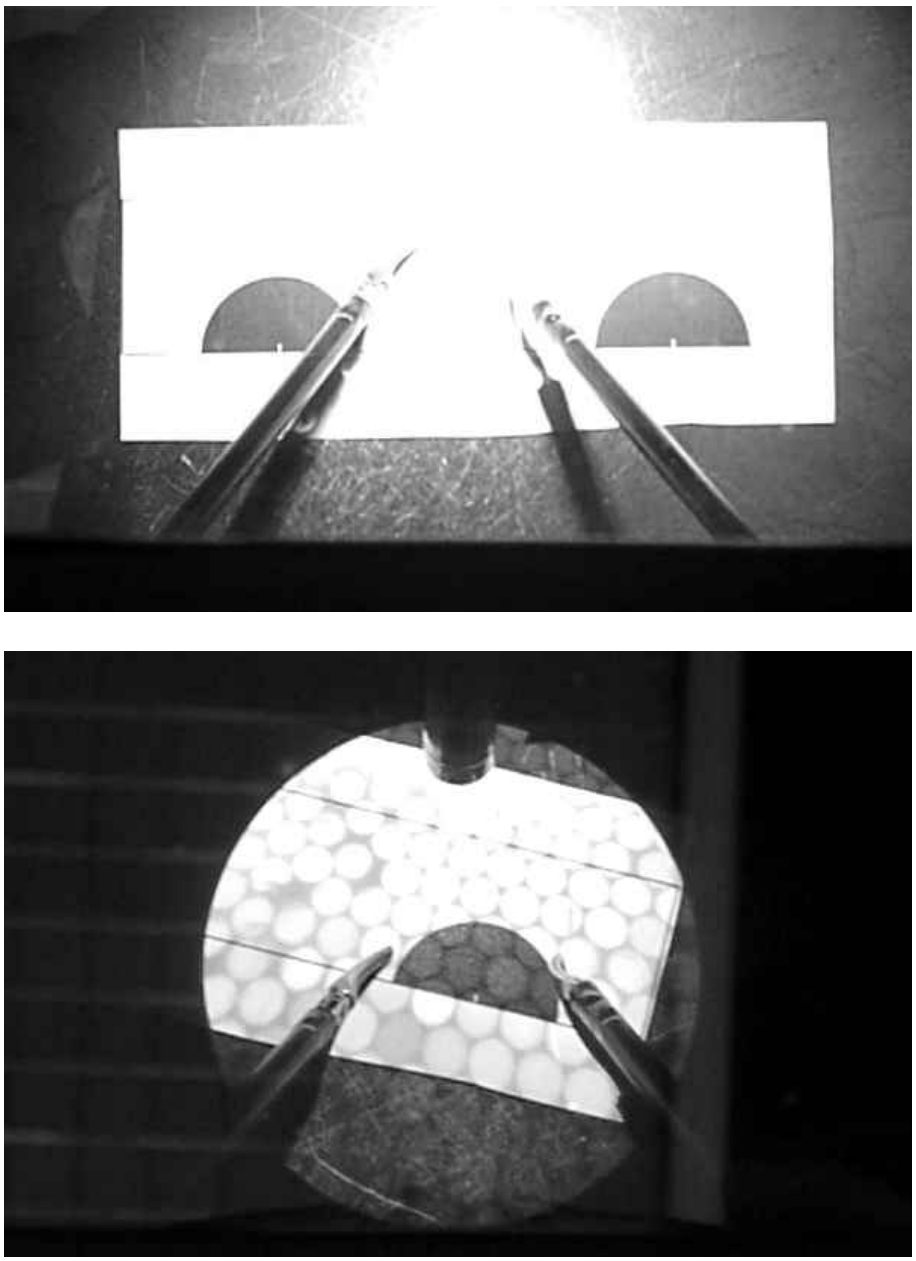


Figure 5.3: The effect of different illumination on the simulated augmented reality visualization. Top: the effect of using the illumination channel of the laparoscope. While the center of the image is saturated making it difficult to view, much of the inside of the trainer remains visible, including the shafts of the instruments and the whole strip of paper. Bottom: the effect of projecting light through the optical channel of the scope. Only the narrow area illuminated by the scope is visible. The regular pattern of circles visible in the illumination is the result of the individual fiber optic fibers in the light guide-cable.

5.1.1.2 Only front surface of structures visible

A second consequence of using a laparoscope is that it can only see visible surfaces. Rerendering of the laparoscope's view from a new perspective cannot show surfaces that are not visible to the laparoscope even though they might be visible from the novel perspective. This effect can also be simulated with the half-silvered glass and high-intensity light source. Surfaces not visible from the laparoscope will not be illuminated by the light source and therefore will not be visible to the user viewing them through the half-silvered glass. This is illustrated in Figure 5.3.

5.1.1.3 Limited image quality of visible surfaces due to the resolution of the acquisition device.

This can be achieved by distorting the light illuminating objects inside the trainer. Many high quality surgical light sources, if used without a diffuser, show the outline of each transmission fiber. Because some fibers are broken or are angled so they don't get any light, light coming out of the fiber projects an irregular pattern. This pattern in turn is projected onto the objects being viewed creating a view which visually distorts the object's surfaces. This effect is demonstrated in Figure 5.3.

5.1.1.4 Limited image quality in the display of visible surfaces.

Currently available display elements for head-mounted displays have resolution up to 1280 by 1024 pixels. The higher resolution display elements are significantly larger and heavier than lower resolution elements making them impractical to use in a head-mounted device. Currently, displays with resolutions of 800x600 pixels appear to be the best compromise between size and resolution. At this resolution, noticable pixelation occurs when observing fine structures.

Looking through a wire mesh from an appropriate distance can give a pixelated effect to the objects viewed through the mesh. If one gets too close to the mesh, or too far away, the effect is lost. Multiple levels of meshes (with different frequency and thickness of lines) compensate for moving closer or further away. Arbitrarily designed meshes can be easily created with drawing software. These meshes can be printed on transparency film. The effect of multi-resolution meshes, is shown in Figure 5.4

Figure 5.4: An approximation of the view a user would have while working under degraded simulated augmented reality conditions. The mesh makes focusing on the objects inside the box difficult.

5.1.1.5 Errors in tracking

Tracking errors take two major forms: complete loss of tracking and minor errors in tracking. In a complete loss of tracking, no current information about the position of one or more objects (including the user's head) is known making the subsequent correct rendering impossible. Minor errors are small errors in estimated position, and are usually seen by the user as slight misalignment of real and virtual objects or "jitter" of virtual objects.

Complete loss of tracking gives rise to a lack of information about the position of real and virtual objects relative to the user. This most frequently happens when the user inadvertantly leaves the tracked area or when a line-of-site constraint for a given tracker is violated. While more graceful modes of failure could be implemented, several simple methods can be deployed to simulate this effect.

Careful selection of the size and location of the "pit" can severely limit the positions from which the user can see into the trainer. The meshes described for limiting viewable image quality provide reduced visibility inside the trainer as the view angle becomes less perpendicular to their surface. Lights can be positioned around the trainer so that their illumination exceeds the illumination inside the trainer except from positions within the "trackable volume".

An additional method to simulate tracking errors relies on the use of an actual tracker. The tracker could be used to trigger an event (e.g., making an LCD plate over the "pit" opaque, turning off the lights inside the trainer) when the user is outside the volume.

Minor errors in tracking are more difficult to simulate. In this scenario, everything that the user sees inside the the trainer is a virtualized object. Small motion could be induced by placing a lenticular or slightly refractive sheet in the window over the trainer. "Jitter" could then be induced by moving this sheet.

5.1.1.6 Errors in range (depth extraction) information.

Errors in depth extraction might look to the user like distortions of real objects. Because depth extraction for this application is intended to be a continuous, dynamic process, one can imagine that a flat surface might appear more like a rippling wave. This effect can be generated using a similar technique to that described in the previous section for simulating tracker errors. Rather than using a single lenticular sheet, two lenticular sheets moved independently could be used.

5.1.2 Degrading simulated augmented reality

Some of the visual errors induced by technological limitations have some complex ways of correctly recreating their visual effects. While possible methods of recreating these effects are described in the previous section, in most cases a user, particularly a naive user, of an augmented reality system is most likely to perceive the effects in a fairly non-specific way. For example if the display resolution is too low they might say, "I can't see things clearly" or "it doesn't look right". For this reason I believe a very coarse representation of these errors, in a way which is very visually disturbing, is the best choice for approximating the effects of imperfect technological system components.

As described in the previous section, a multi-resolution mesh nicely simulates a variety of potential problems with augmented reality systems. Use of these meshes should provide sufficent degradation of the simulated augmented reality view to give a realistic way to measure the potential benefit, if any, of a less than ideal augmented reality environment.

5.2 The simulated augmented reality trainer

The simulated augmented reality trainer used in these studies is shown in Figure 5.5. It was constructed from the following parts:

- 36 Liter plastic storage container – *Sterilite* Model 1720
- matte finish epoxy spraypaint
- Four, 1 inch mirror hanging brackets
- 10cm x 20cm, 1/4 inch thick half-silvered glass
- 10cm x 20cm opaque material (black paper)
- 10cm x 20cm multi-resolution gride overlays (laser printed transparency film)
- 10cm x 20cm, 1/8 inch thick glass (clear)
- Gatorboard Gatorfoam^{TM} Arts Board cut to fit

The bottom of the *Sterilite* container was removed at approximately a 15° angle. A piece of the Gatorboard was cut to fit snugly into the bottom, completely closing the bottom. Remaining gaps in the bottom of the trainer were filled using caulk. An opening of 9cm by 19cm was cut into the lid of the trainer and holes for the mirror mounting brackets were punched 2cm from the edge of the opening on each edge. A hole 13mm in diameter was punched above the opening and two 8mm holes were punched below the opening. These are the openings for the laparoscope and the laparoscopic instruments.

The entire inner surface of the trainer, including the bottom of the lid and the mirror brackets, was then painted with a black, matte finish, epoxy paint.

The trainer can be set up for use as a conventional laparoscopic trainer by mounting the opaque material over the opening. It can also be set up for simulated augmented reality use by mounting the half-silvered glass over the opening, and it can be set up for degraded simulated augmented reality use by mounting the multi-resolution mesh under the half-silvered glass.

An additional small opening was cut in the back wall of the trainer. A miniaturized video camera was inserted through this hole so that actions within the trainer could be video taped without interfering with the performance of tasks or reliance on the video image from the laparoscope.

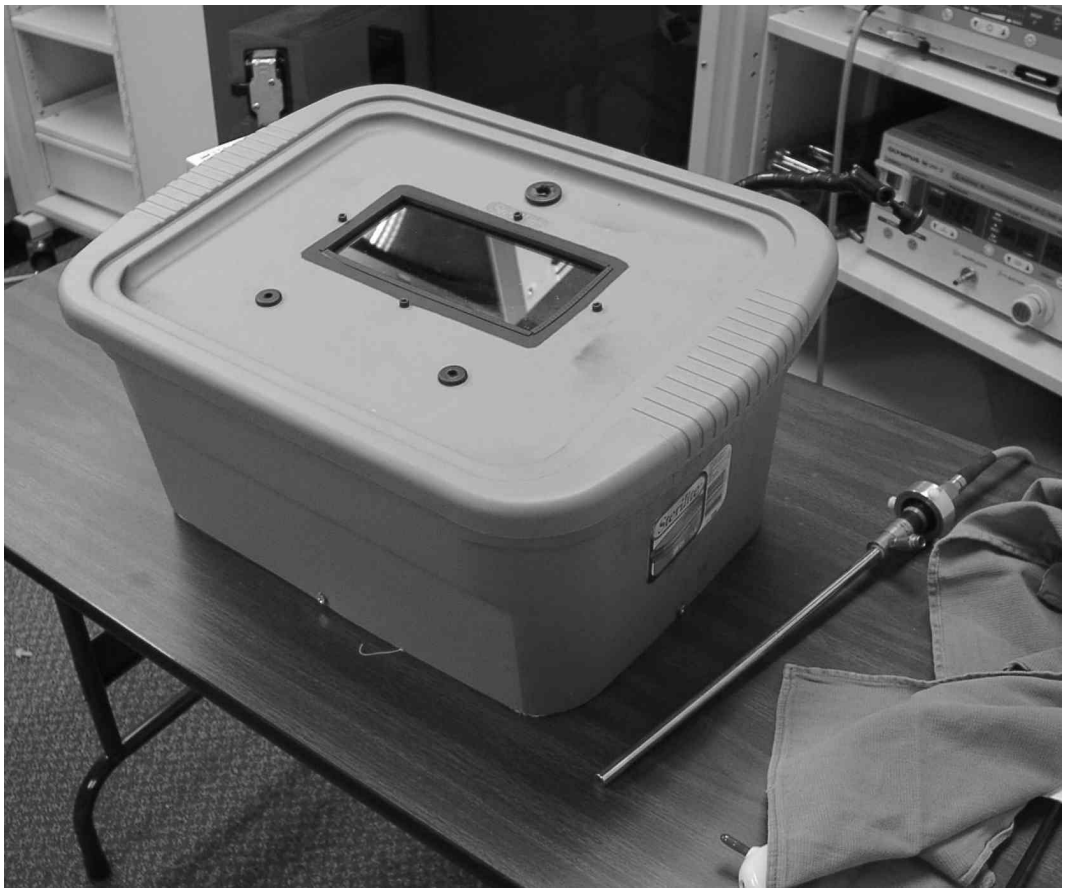


Figure 5.5: This is a photograph of the simulated augmented reality trainer. The glass pane can be replaced to transform the trainer into a conventional laparoscopic trainer, a simulated augmented reality trainer, or a degraded simulated augmented reality trainer. Instruments are inserted through the pair of holes below the window and the laparoscope is inserted through the single hole above the window.

Chapter 6

User study

"Medical Science becomes more perfect, its branches complement each other, and thus improved, it ascends the missing steps easier and faster to the final goal"

– Philip Bozzini, 1806

This chapter describes a user study designed to answer the primary question, "Is an augmented visualization useful for performing surgical tasks?" The choice of test task is discussed as are the findings of this study. For clarity, only the findings most pertinent to this question are discussed in this chapter. Additional findings are discussed in Chapter 7.

6.1 Study goals

The purpose of this study is to examine the effect of the use of an augmented reality visualization on surgical performance. In particular this study was designed to measure the effect of augmented reality visualization on performance of a task analogous to those performed during surgery.

6.2 Task for evaluation

The choice of the task and method of evaluation is critical. The task should be complex (e.g., involving multiple steps), require both hands for completion, and have multiple possible measures for evaluation. Possible measures include time to completion, successful completion, accuracy, and adequacy. The following are some suggested tasks, including some published batteries of tasks and some possible measures for evaluation:

- Simulated anastomosis, joining the ends of rubber tubes. Time to completion, pressure to failure.
- Paper marked with lines, cut between lines/on lines. Time to completion (per cut/entire paper), accuracy of cut placement
- Paper marked with lines, fold on lines. Time to completion (per fold/entire paper), accuracy of fold placement, (perhaps) finished object
- Multi-object (perhaps oriented) pick and place. Time to completion (per object/entire task), accuracy (precise orientation of objects on target, correct object placed)
- The McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS) battery[Derossis98] composed of the following seven tasks:
  - Two handed peg board transfer (lift, pass, place)
  - Pattern cutting
  - Placing clips between marks and cutting
  - Placing ligating loops
  - Placing and stapling a mesh
  - Simple suturing – intracorporeal tie
  - Simple suturing – extracorporeal tie

  These items include time and performance (penalties) for measures.
- The Yale Laparoscopic Skills and Suturing Program[Rosser Jr.98, Rosser97, Scott01] (also known as the "TopGun" program) which is composed of the following four tasks scored by completion time:
  - Cup drop – transfer a number of beans from a pile into a cup
  - Rope Pass – pass the length of a rope hand-over-hand through the instrument tips
  - Triangle transfer – pick up block using an eye-bolt on the block and a curved hook held in an instrument and transfer it to a marked position
  - Suturing – tie a simple suture using intracorporeal technique

It is important to remember that there is little consensus as to how to best evaluate a surgeon's skills in an ex-vivo environment. As such, there is no obvious choice for a task to measure a surgeon's performance using an experimental device or visualization.

In choosing a study task the significant considerations include: surgical relevance, difficulty, outcome measures and time to completion. A task which is not relevant to surgery is not appropriate because it is not easy to justify how performance on this task might relate to surgical performance. A task that is too easy may not provide sufficient resolution to differentiate levels of performance while a task that is too difficult may not be completed by study subjects. Time to completion is an important consideration for the practical reason of recruiting subjects. The outcome measures given by a task are significant because they influence the relevance of the task, the number of different types of data available for later analysis, and the types of analysis that can later be performed.

An additional consideration is prior validation. The Yale Laparoscopic Skills tasks have been extensively validated with participants in that training workshop. The MISTELs battery has been validated via several studies as well.

Tasks measuring suturing skill, as in the suturing and anastamosis tasks, are frequently pointed to as being the most surgically relevant. These tasks might be suited to studies in populations of trained surgeons, but are not appropriate for studies in other populations because of the amount of practice required to perform a simple suture task. Tasks, such as the ones used in the Yale Laparoscopic Skills and Suturing Program[1] are considered highly surgically relevant but are not something that can be completed without practice[2].

Of the tasks listed in the MISTEL battery, only three were found to differentiate between levels of training of surgeons with any statistical significance. Of these tasks the pattern cutting task has the highest degree of significance.

The pattern cutting task consists of cutting a pattern (a circle) from a marked piece of gauze held in place by a clip. The only measure of task performance used previously for this task was time to completion. Even though performance of early residents on this task was highly correlated to level of training, for practicing surgeons and advance residents, this correlation was not present. While performance on

[1] Also used for the American College of Surgeons *Top Gun* competition.

[2] These tasks are taught as part of two and a half day laparscopic skills and suturing workshops which use the three basic tasks as the basis for learning the suturing task

this task was previously only evaluated based on time and error score (mishandling of instruments), this task can easily be modified to use accuracy of cutting as its evaluation measure.

Ideally the accuracy of each individual cut could be measured and used as a separate data point, but this is not practical. A more practical approach is to over-sample the accuracy of all the cuts made (by measuring the final cut circle), and then use statistical methods (repeated measures) to estimate the effect of over-sampling. In the absence of these sophisticated statistical methods, accuracy measurements can be consolidated into single measures (e.g., mean radius, sum of squared errors) but with less statistical power for hypothesis testing.

Given these considerations, the circle cutting task from the MISTEL was selected as the basis to create a new task to be used in this study.

6.2.1 Paper circle cutting

The following modifications were made to the MISTEL circle cutting task:

- The material to be cut is paper rather than gauze
- The circle is printed with high precision (laser printed) with a smaller diameter (0.8 inches) rather than 10cm
- The material to be cut is not held by any clamps within the trainer
- Primary outcome measure is accuracy rather than time
- A practice period on a similar task is allowed for subjects to familiarize themselves with the task, instruments, and visualization
- A time limit for completion of the task (20 minutes) was instituted

6.3 Intended study subject population

Ideally this experiment would be done by trained surgeons who currently do laparoscopic surgery. In this population, any benefit found in the simulated augmented reality environment would be more significant because it would indicate a benefit to current surgeons. It is also likely that in a study of practicing laparoscopic surgeons, many subjects will not perform significantly better in a simulated augmented reality

environment than in the laparoscopic environment they are trained to use. However, trained surgeons may be difficult to recruit for participation in such a study due to the many demands on their time.

Surgical residents are a more accessible group of study subjects. They are presumably divided in skill by the number of years they have been surgical residents. They may see the study as an educational opportunity to gain certain skills sooner.

Medical students are likely to be more accessible than surgical residents. Laparoscopic skills will not be highly developed in any of these students regardless of their year in medical school. Familiarity with the basic concepts of laparoscopic surgery should be fairly uniform across all medical school classes.

Use of non-surgeons would likely increase the probability of finding significant positive improvement in the simulated augmented reality environment. It also raises larger difficulty in trying to make the results applicable to the ultimate target population. Use of non-surgeons also allows for the possibility of exploring the potential role of augmented reality in training new surgeons.

6.3.1 Data to be collected

For each experimental task, data on appropriate evaluation measures will be collected. The experiments (all viewing conditions) should be repeated during the same session as well as with some separation in time. This will allow evaluation of learning curves.

6.4 Pilot study

A pilot study was conducted to test the hypothesis that the modified direct visualization environment is beneficial to the performance of a task. The primary goals of the pilot study were to find flaws in the experimental protocol, determine if the "circle cutting" task is an appropriate task, and determine the sample size needed for statistically significant results.

Six study subjects (two surgeons and four non-surgeons) were recruited without compensation. The subjects were asked to complete the circle cutting task and the Rope Pass task from the Yale Laparoscopic Skills battery, described below, under both standard laparoscopic visualization and modified direct visualization. The subjects were randomly assigned which visualization method was used first.

The circle cutting task is a modification of the pattern cutting task from the

MISTEL battery. In this version of the task, study subjects cut out a one inch circle marked on a strip of paper. Completion of the task is video recorded from the laparoscope and from a secondary view inside the trainer box. The cut circle is scanned (flat bed scanner at 300 dpi) and then analyzed. The radii cut by the subject at two hundred evenly spaced angles is measured. Each of these measurements is compared to the known original radius of the marked circle and the magnitude of internal errors (where the subject cut inside the marked circle), and external (where the subject left excess paper) can be calculated. Both the remnant circle and recorded video can also be used for other analyses including measurement of time to completion, tool handling errors, and extraneous cuts (e.g., cuts penetrating the circle that do not effect the outer edge of the circle).

The circle cutting task is a complex and difficult task. The task must be completed using two instruments. Depth judgements are critical to the task in order to grasp or cut the paper.

The Rope Pass task simulates the surgical task of inspecting the small bowel. Study subjects begin at one end of a coiled piece of string and pass from one instrument to another in short segments until they reach the other end of the string. This task has been used by other researchers and is evaluated by time to completion.

The following information was learned from the pilot study:

- An improvement in performance was observed in all study subjects while performing the circle cutting task in the modified direct visualization environment
- Some task training (practice) should be done before subjects perform the circle cutting task
- Better (prerecorded video) instructions should be given to insure consistency of instruction for all subjects
- Consistent instrumentation, specifically adequately sharpened cutting tools, should be supplied for all study subjects
- A sample size of 8-10 surgeons and 16-20 non-surgeons should be sufficent to get high statistical power for the types of differences observed in the pilot study
- A computer program was developed to measure radius of cuts from scanned images of cut circles

6.5 Study design

Based on the results of the pilot study, a larger study was designed. Eight practicing surgeons (chief resident or more experienced) and 40 first, second, and third year medical students (third year students were excluded if they had already started the surgical clerkship) were recruited and were assigned randomly to the study groups. Seven randomization groups with differing sequences of visualizations were created. The sequence and sample size of each of these groups is shown in Figure 6.1.

Based on the results of the pilot study a sample size of sixteen medical students and eight surgeons were assigned to randomization groups for the primary hypotheses of the utility of augmented reality visualization. Half of these used simulated augmented reality visualization first, and half used laparoscopic visualization first. On their second circle they used the other method. This randomization was intended to provide a means of controlling for training effects.

Eight medical students were assigned to groups cutting their first two circles under either degraded simulated augmented reality or laparoscopic visualization. These groups were intended to provide data on the effect of less than ideal augmented environments. Based on the pilot study results it was anticipated that this would not generate statistically significant results.

Eight medical students were assigned to cutting circles first under simulated augmented reality and then under either degraded simulated augmented reality (four students) or the combination of degraded simulated augmented reality and laparoscopic visualization (four students). These groups were intended to gather data about the usefulness of simulated augmented reality and degraded simulated augmented reality and learning transfer to fully laparoscopic work.

Finally, eight medical students were randomized to cut all of their circles using laparoscopic visualization. This group was intended to serve as a control group for exploring the potential role of augmented reality in learning.

All medical students cut a third circle under laparoscopic guidance. This was intended as a mechanism to compare the effectiveness of different sequences for training students to perform the task laparoscopically.

Due to the small number of participants in each randomization group, an overall significance threshold of $p = 0.10$ was selected. Because of the exploratory nature of this study, results approaching this significance level are also reported although they are not statistically significant by this criteria.

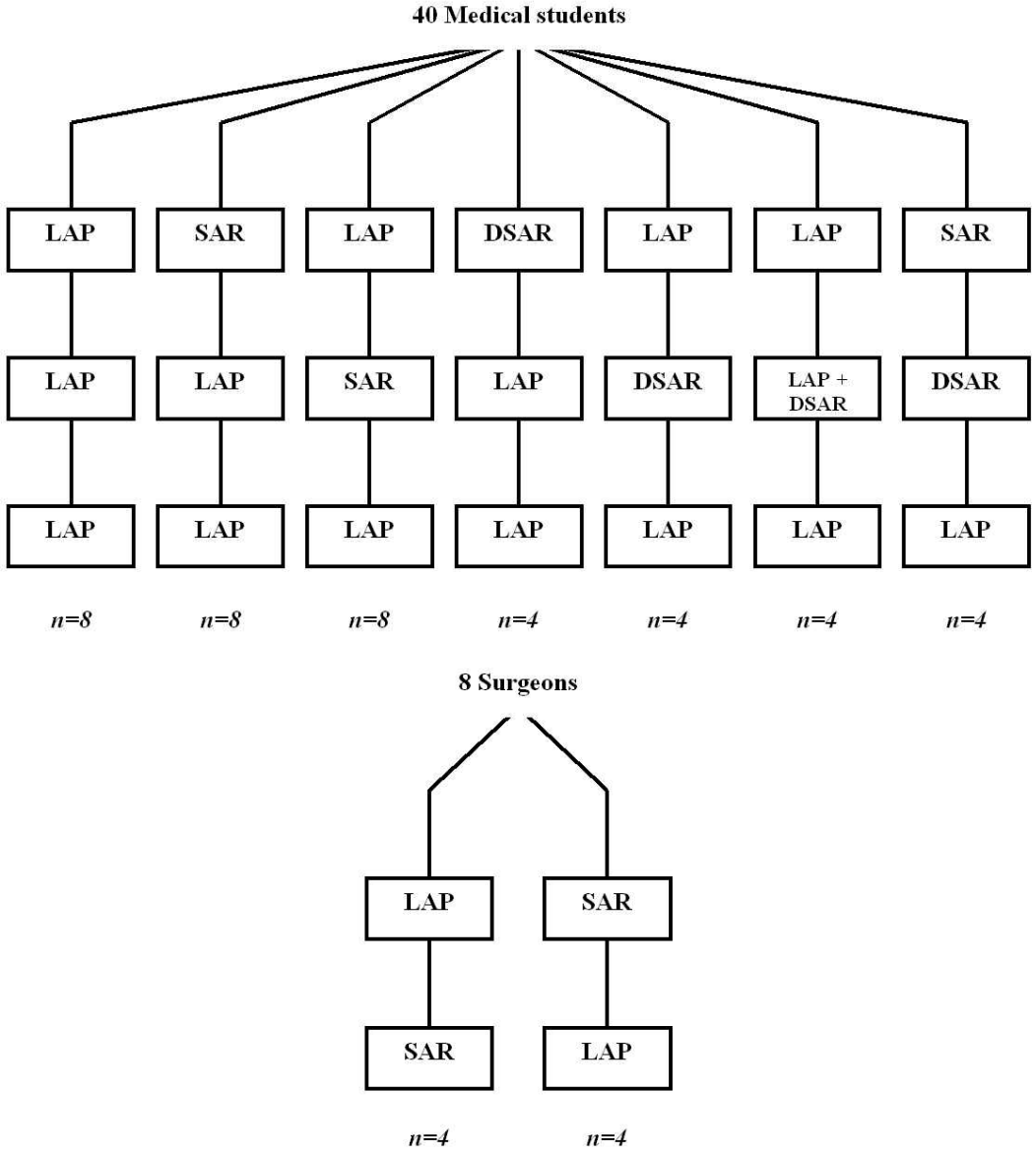


Figure 6.1: This chart shows the overall study design. Medical student and surgeon participants were randomly assigned to each of the branches shown here, "LAP" refers to cutting circles under laparoscopic visualizatoin, "SAR" refers to cutting circles under simulated augmented reality visualization, "DSAR" refers to cutting circles under degraded simulated augmented reality visualization, and "DSAR+LAP" to having both degraded simulated augmented reality and laparoscopic visualization available while cutting.

6.6 Study procedure

All study subjects were offered $10 for completion of the study. An additional $50 was given to the four (4) medical students with the best (lowest) total score. The two surgeons (25%) with the best total score were given a dinner for two ($175). These additional awards were offered as an incentive for participants to perform as well as they could.

After reading and signing the consent form, each study subject completed a short survey and was shown a brief instructional video explaining the use of the instruments and the study task. The video demonstrated basic use of the instruments and explained the time limits and scoring criteria.

Study subjects were then allowed up to 10 minutes to practice by cutting out semi-circles of the same diameter as the circles for the study task using the same visualization they had been assigned for their first task. After completion of a practice period the length of the practice period was recorded and the subjects were allowed to examine their work. Subjects were then given up to twenty minutes to cut their first circle. A video of their efforts was recorded from a video camera inside the trainer for future analysis. Time to completion of this task was recorded.

After completing the first task, subjects watched the instructional video a second time while the trainer was converted to the second visualization mode. Subjects were again allowed up to 10 minutes to practice cutting the half circles. Practice time was recorded and subjects were allowed to examine their work. Subject were then given up to 20 minutes to cut the circle. A video of their efforts was recorded from the video camera in the trainer and time to completion was recorded. Comments about the study task and visualizations as well as other observations were written down during the study.

Subject identification number and trial number were written on the back surface of all cut circles. Each circle was then scanned against a constrasting (red) background at 300 dots-per-inch and these images were saved in a non-compressed format. A computer program was used to measure the radius of the cut circle at 200 equally spaced positions around each circle. These radius measurements, along with survey responses, were then imported into a spreadsheet where the measurements were kept and descriptive statistics (described in more detail in Section 6.7) were generated.

6.7 Data analysis

There are three stages of data analysis:

1. Measuring circle radii
2. Calculating descriptive statistics
3. Analysis of survey data and descriptive statistics

6.7.1 Measuring cut circles

A computer program was written to load, display, and analyze scanned iamges of the cut circles.

The loaded image is displayed in a window with basic pan and zoom controls. The user must then move a cursor to the center of the cut circle (assisted by the cross marked on the paper) and select the center of the circle. The selected center is recorded to sub-pixel (with respect to the source image) accuracy.

Using a simple intensity and color threshold-based heuristic, each pixel in the source images is identified as belonging to the cut circle or the background. Although the program allows the threshold levels to be changed interactively, threshold levels were selected based on circles scanned for the pilot study and these values were used throughout this study. The program then searches for the edge of the circle along the requested number of evenly spaced angles (200 were used throughout the study), starting at the selected center. The search stops when a 'background pixel' pixel is reached. In order to prevent mis-identifying the edge in the event of small holes or noise in the scanned image, the search traverses slightly beyond a found edge and will continue if additional 'paper' pixels are found. The radius traversed (in integer pixels) is recorded in the output file for each ray. The program also records the name of the source image file, the threshold levels, and the position of the selected center of the circle.

6.7.2 Calculating descriptive statistics

After circles were measured, the measurements were imported into a Microsoft Excel spreadsheet. Subject responses from the survey were also included in this spreadsheet. The spreadsheet also contains the calculations for generating descriptive statistics about the circles. These statistics include mean and standard deviation of measured

radii, slice area, total area of errors, total area of errors cutting into the circle, total area of errors outside the circle, and counts characterizing the error observed at each radius.

The area, A of the n^{th} 'slice'(boardered by radius measurements r_n and r_{n+1} and swept by the angle ϕ) is

$$A_n = r_n r_{n+1} sin(\phi) \tag{6.1}$$

These area measurements can be converted into per-slice error measurements by subtraction from A_S, the area of a "standard" error-free slice:

$$E_n = A_n - A_s \tag{6.2}$$

A slightly more accurate result could be reached by making some assumption about the way radius varies between the measured points and then integrating the difference between the radius (based on measurements) and the ideal value over ϕ.

Total absolute error, E_{abs} is then

$$E_{abs} = \sum_{n=1}^{m} |E_n| \tag{6.3}$$

Calculations of the amount of error that occured inside the circle (that is, too much material was cut away) or outside (excess material left on) can be easily done by conditional sum (by sign) of E_n.

6.7.3 Overall analysis

The Microsoft Excel spreadsheet containing the summary statistics from each circle as well as randomization information and survey responses was then imported into the SPSS statistical package. All subsequent statistical analyses were performed using this package.

6.8 Internal validation of outcome measure

Every measurement method has associated sources of error. An important part of using measurement techniques for scientific work is to characterize the magnitude and source of expected errors. In general it is important to characterize both the accuracy (how close to the actual value repeated measurements will be) and repeatability (how

much a single measurement will vary from the mean). In this case accuracy is not important because these measurements need not be compared to an external standard. Repeatability is quite important because it effects our ability to make claims about correlations of our measure with independent variables.

While a measure of repeatability of the entire task would be nice, a wide variety factor, most notably learning effects, make it difficult to measure the overall repeatability of the task. It is possible to measure the variablity in our ability to measure the outcome measure given a cut circle. Given the automated methods for processing data, it is reasonable to assume that there are two primary sources of variability, due to variation in human judgment and sensors. These are differences in scanned images, and selection of the center point of the circles.

Seven cut circles were selected at random (5% of the available circle population), and were each rescanned three times. Each scanned circle was reanalyzed by the program three times, requiring separate selections of each circle's centers.

Based on this data we can conclude: The average standard deviation between scans is 156.6 (300dpi pixels) (0.21%). The average standard deviation between analyses of a given scanned circle is 53.1 (300dpi pixels) (0.072%) The average standard deviation of any scan/analysis as compared to a large sample is 165.4 (300dpi pixels) (0.22%)

That means that with 95% certainty, we could expect variation of less than 306.9 (300dpi pixels) (0.41%) by rescanning the circle, 104.1 (300dpi pixels) (0.14%) by reanalyzing the circle, and 324.1 (300dpi pixels) (0.437%) by both rescanning and reanalyzing the circle.

Because it is possible to scan the circles at much higher resolution, three circles were scanned at 600dpi and the scan was analyzed three times. The standard deviation of the error between the results of these scans and those from identical (lower resolution scan captured without moving the circle) was (in 300dpi pixels) 126.9 (0.17%). The standard deviation of the differences measured that are due to different selection of the center at this higher resolution was (normalized to 1/300th inch pixels for comparison) 35.6 (0.048%).

These results indicate that there is a benefit to scanning at higher resolution (in terms of repeatability of this measure), but little benefit to overall accuracy.

Figures 6.2-6.5 show the range of variability of the cut circles. These figures show a scanned image of the cut circle on the left, a drawing of the circle segmentation in the middle, and some of the descriptive statistics generated by the segmentation on the right. In the segmented drawings, gray area represents area of the circle that

the subject correctly left in place, and black represents area where the subject either cut too much material or left extra material. A faint circle is visible around these drawings for scale.

6.9 Results

The results in the following sections were tested for statistical significance by repeated measures analysis of variance (REPM ANOVA) using the SPSS package. Data from the randomization groups of the study that performed tasks under the same visualization conditions but in different order were pooled for these analyses.

Using the REPM ANOVA rather than a standard ANOVA improves the accuracy of the estimate of significance. It is the appropriate measure of significance to use in randomized cross-over trials, that is subjects are tested under experimental or control conditions and then are retested under the opposite condition, when there is an expectation that the order of exposure may have a signifcant impact on the dependent variables. This method gives significance levels for each between- and within-subject sources of variation. In this case the between-subject effects are due to the randomization group, while the within-subject effects are due to the visualization and the cross-over effect of visualization given randomization group.

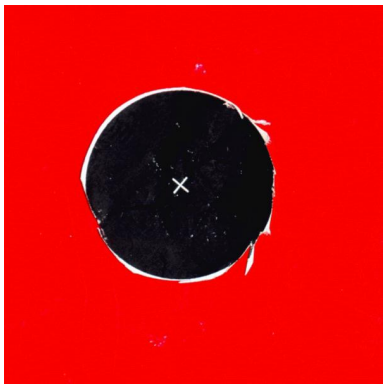
A multiply-multi-variate analysis can be performed when several outcome measures are available. In this case the total area left on the circle, outside errors, and the amount of material inoorectly cut away, inside errors, can be modeled separately. This analysis is useful for identifying any biases in the types of errors and for identifying more subtle differences between performance with respect to visualization and randomization groups.

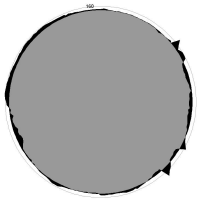
While some of the reported results are below acceptable levels of significance ($p > 0.10$), they should still be carefully considered due to their relatively high significance in the face of small sample sizes.

While significance levels for the differences between randomization groups are given here, their relevance is discussed in 7.1.

6.9.1 Performance in simulated augmented reality

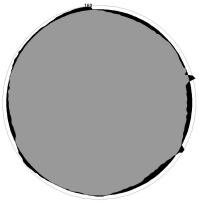
Medical students using simulated augmented reality visualization performed better than those using laparoscopic visualization ($p = 0.011$, $n = 16$). The group random-





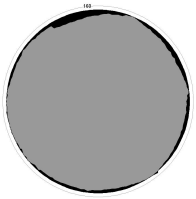
Mean radius	154.25
Std. dev.(radius)	4.02
Total Error Area	3040
Inner Error Area	1228
Outer Error Area	1812

Representative image of the 1st decile of circles cut.



Mean radius	151.95
Std. dev.(radius)	4.1
Total Error Area	3554
Inner Error Area	2587
Outer Error Area	967

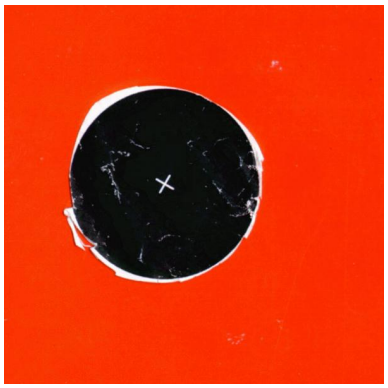
Representative image of the 2nd decile of circles cut.

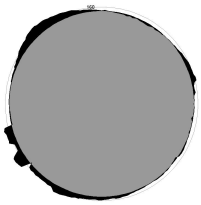


Mean radius	149.36
Std. dev.(radius)	3.20
Total Error Area	4142
Inner Error Area	4115
Outer Error Area	27

Representative image of the 3nd decile of circles cut.

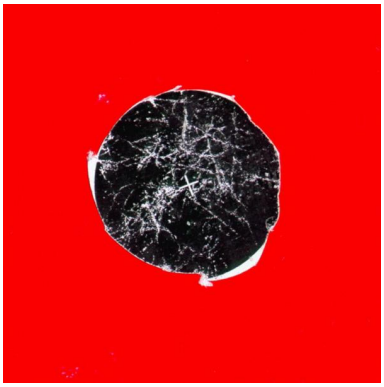
Figure 6.2: This figure shows representative images of the 1st through 3rd decile of circles cut. The images on the left show the scanned circle (slightly magnified), the middle image is a rendering based on the found radii with gray area representing correctly cut areas and black representing errors. The corresponding numerical scores are shown on the right. All measurements are given in 1/300ths of an inch (pixels).

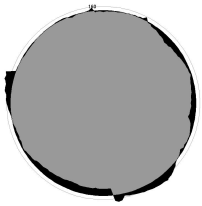




Mean radius	157.75
Std. dev.(radius)	5.28
Total Error Area	4694
Inner Error Area	318
Outer Error Area	4376

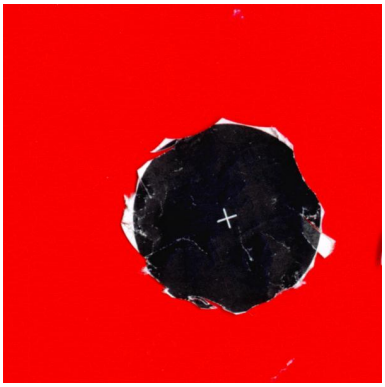
Representative image of the 4th decile of circles cut.

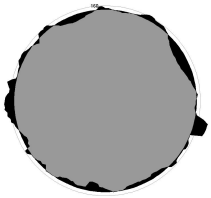




Mean radius	151.7
Std. dev.(radius)	6.72
Total Error Area	5235
Inner Error Area	3502
Outer Error Area	1733

Representative image of the 5th decile of circles cut.

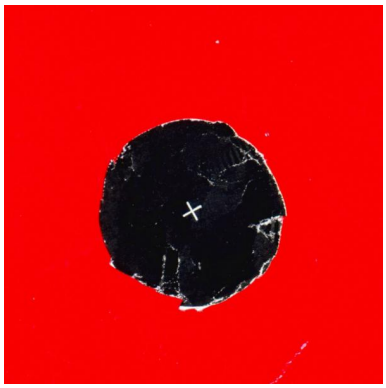


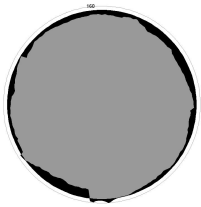


Mean radius	152.78
Std. dev.(radius)	7.95
Total Error Area	5855
Inner Error Area	3272
Outer Error Area	2583

Representative image of the 6th decile of circles cut.

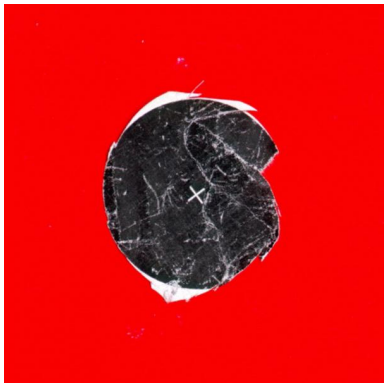
Figure 6.3: This figure shows representative images of the 4th through 6th decile of circles cut. The images on the left show the scanned circle (slightly magnified), the middle image is a rendering based on the found radii with gray area representing correctly cut areas and black representing errors. The corresponding numerical scores are shown on the right. All measurements are given in 1/300ths of an inch (pixels).

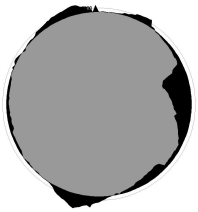




Mean radius	146.77
Std. dev.(radius)	4.46
Total Error Area	6818
Inner Error Area	6645
Outer Error Area	173

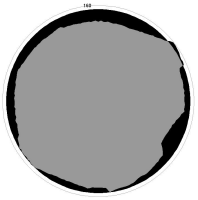
Representative image of the 7th decile of circles cut.





Mean radius	150.44
Std. dev.(radius)	13.26
Total Error Area	8653
Inner Error Area	5606
Outer Error Area	3047

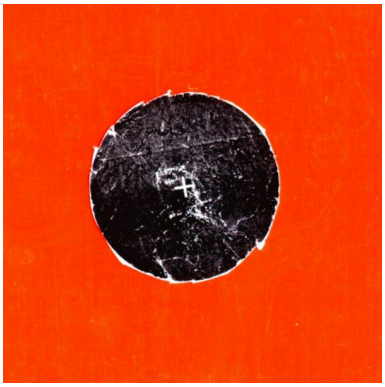
Representative image of the 8th decile of circles cut.

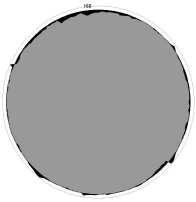


Mean radius	142.09
Std. dev.(radius)	7.19
Total Error Area	10911
Inner Error Area	10764
Outer Error Area	147

Representative image of the 9th decile of circles cut.

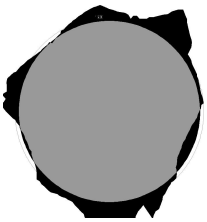
Figure 6.4: This figure shows representative images of the 7th through 9th decile of circles cut. The images on the left show the scanned circle (slightly magnified), the middle image is a rendering based on the found radii with gray area representing correctly cut areas and black representing errors. The corresponding numerical scores are shown on the right. All measurements are given in 1/300ths of an inch (pixels).





Mean radius	152.485
Std. dev.(radius)	2.37
Total Error Area	1879
Inner Error Area	1508
Outer Error Area	371

The best of circles cut.



Mean radius	167.76
Std. dev.(radius)	18.49
Total Error Area	18320
Inner Error Area	1534
Outer Error Area	16786

The worst of the circles cut.

Figure 6.5: Images of the best and worst of circles cut. The images on the left show the scanned circle (slightly magnified), the middle image is a rendering based on the found radii with gray area representing correctly cut areas and black representing errors. The corresponding numerical scores are shown on the right. The segmented image has been cropped for the worst circle to fit more neatly on the page. All measurements are given in 1/300ths of an inch (pixels).

ized to perform the tasks using simulated augmented reality first performed better than the group that used laparoscopic visualization first ($p = 0.115$). A crossover effect with comparably low significance was observed ($p = 0.184$). The difference in inside errors and outside errors was significant ($p = 0.021$). The mean scores for these groups are shown in Table 6.1 and breakdowns of means by visualization and randomization groups are given in Table 6.2.

		Overall	Lap	SAR	SL	LS
Medical Students	Inside Error	3786.5	3624.0	3948.9	2971.9	4601.1
	Outside Error	2050.4	2975.1	1125.7	1874.5	2226.3
	Total Error	5836.9	6599.1	5074.6	4846.4	6827.3
Surgeons	Inside Error	4013.7	3893.1	4134.3	3401.6	4625.7
	Outside Error	801,4	295.0	506.4	359.7	441.7
	Total Error	4414.3	4188.1	4640.6	3761.3	5067.4

Table 6.1: Mean performance scores for surgeons and medical students using laparoscopic versus simulated augmented reality visualization. Overall is for all subjects, "Lap" is for trials done under laparsocopic visualization, "SAR" is for trials done under simulated augmented reality visualization, "SL" for those randomized to do simulated augmented reality and then laparoscopic, and "LS" for those randomized to cut laparoscopically and then using simulated augmented reality .

		Inside error		Outside error		Total error	
		LAP	SAR	LAP	SAR	LAP	SAR
Medical Students	SL	3088.8	2855.0	2155.4	1593.6	5244.1	4448.6
	LS	4159.3	5042.8	3794.8	657.8	7954.1	5700.6
Surgeons	SL	3264.3	3538.9	190.0	529.3	3454.3	4068.2
	LS	4521.8	4729.6	400.0	483.4	4921.8	5213.0

Table 6.2: Mean performances of medical students and surgeons arranged by randomization group and visualization. "SL" indicates randomization groups using simulated augmented reality followed by laparoscopic visualization, "LS" randomization groups using laparoscopic visualization followed by simulated augmented reality . "LAP" indicates performance using laparoscopic visualization and "SAR" indicates performance using simulated augmented reality visualization.

Surgeons using laparoscopic visualization performed better than when using simulated augmented reality visualization ($p = 0.458$, $n = 8$). This difference is more significant for the outside errors ($p = 0.216$). The group randomized to perform the tasks using simulated augmented reality first performed better than the group that used laparoscopic visualization first ($p = 0.103$). No significant crossover effect was

observed ($p = 0.787$). The difference in performance between inside errors and outside errors was significant ($p = 0.001$). The mean scores for these groups are shown in Table 6.1 and breakdowns of means by visualization and randomization groups are given in Table 6.2. Both sets are given in the same chart as equivalent data from medical students to aid in comparision.

6.9.2 Performance in degraded simulated augmented reality

Slightly worse performance was observed under degraded simulated augmented reality visualization as compared to laparoscopic, but it was not significant ($p = 0.948$). Subjects assigned to use degraded simulated augmented reality followed by laparoscopic visualization performed slightly worse those who worked in the opposite order ($p = 0.324$). However, this effect is more pronounced when inside errors ($p = 0.272$) and outside errors ($p = .114$) are considered seperately. The means contributing to these calculations are shown in Table 6.3. No significant crossover effect ($p = 0.870$) was observed. The difference in performance between inside errors and outside errors was not significant ($p = 0.249$). The means representing these effects are shown in Table 6.4.

	Overall	Lap	DSAR	DL	LD
Inside Error	4223.6	3917.0	4530.2	3330.2	5117.0
Outside Error	2338.5	2416.9	2260.4	3332.1	1345.1
Total Error	6562.1	6333.9	6790.6	6662.4	6462.1

Table 6.3: Mean performance scores for medical students using laparoscopic versus degraded simulated augmented reality visualization. "Overall" is for all medical students randomized to use only laparoscopic visualization and degraded simulated augmented reality , "Lap" is for trials done under laparsocopic visualizatio, "DSAR" is for trials done under degraded simulated augmented reality visualization, "DL" for those randomized to do degraded simulated augmented reality and then laparoscopic, and "LD" for those randomized to cut laparoscopically and then using degraded simulated augmented reality .

	Inside error		Outside error		Total error	
	LAP	DSAR	LAP	DSAR	LAP	DSAR
DL	3122.5	3537.9	3055.3	3609.1	6177.8	7147.0
LD	4711.5	5522.6	1778.2	911.9	6489.7	6434.5

Table 6.4: Mean performances of medical students arranged by randomization group and visualization. "DL" indicates randomization groups using degraded simulated augmented reality followed by laparoscopic visualization, "LD" randomization groups using laparoscopic visualization followed by degraded simulated augmented reality . "LAP" indicates performance using laparoscopic visualization and "DAR" indicates performance using degraded simulated augmented reality visualization.

Chapter 7

Secondary findings

"Though the sense of touch provides ample knowledge, vision will support its findings, and the more of our senses that are concentrated on an object, the less likely will we be deceived."

– Philip Bozzini, 1806

While the primary goal of this user study was to evaluate the benefit of augmented reality visualization for laparoscopic surgery, benefits and disadvantages beyond those seen by a simple comparision of performance between simulated augmented reality and laparoscopic visualization are likely to exist. The study was designed to address several other important questions. In particular, is there any difference in learning between those who had exposure to a simulated augmented reality environment and those who did not? Are there other correlations with improved performance? These questions may give a more complete picture of the usefulness of augmented reality technology for laparoscopic surgery.

7.1 The role of augmented reality in learning

Many in the surgical community are concerned about how to most efficiently educate surgeons, both residents and those already in practice, to perform laparoscopic procedures. There are three components of this training. The first is learning technical skills such as camera navigation, manipulation, and suturing. The second is learning to perform basic procedures such as cholecystectomy. Finally, the third component is learning advanced procedures and techniques including complex procedures, procedures on obese and high-risk patients, and use of intraoperative imaging.

Augmented reality may have a useful role in all of these areas. This study was designed to detect this effect of augmented reality on learning. The circle cutting

task is primarily a test of basic manipulation skills. However, the task requires the use of basic manipulation skills that a laparoscopic surgeon already has mastered but applies them to an unfamiliar situation. In this sense it also examines learning associated with the second type of training, albeit not in a way directly pertinent to operating on patients.

The cross-over and randomization effects detected by the statistical analysis in the previous chapter may be a result of learning differences between the randomization groups, as a direct result of the order and type of exposure to simulated augmented reality environments. The third circle cutting task, completed by all medical students under laparoscopic visualization, provides some additional data to support the theory that exposure to a simulated augmented reality environment, and therefore also an augmented reality environment, has a beneficial effect on trained performance using laparoscopic visualization.

7.1.1 Experimental Design

The overall experimental design for the medical students, presented in Figure 6.1, was intended to provide control groups for the randomization arms. The pairs of randomization groups who used the same visualizations but in opposite order provided controls for the effect of order in the results reported in the previous chapter. For evaluating learning effects a control group with no exposure to any simulated augmented reality visualization was used. The third trial, which all the medical students completed using laparoscopic visualization, provides a mechanism to compare how well they learned to complete the task laparoscopically.

The reported differences and significance values in this section are based on linear regression and analysis of variance (ANOVA) except where otherwise noted. Laparoscopic visualization and randomization to the group using the sequence of visualizatons Laparoscopic-Laparoscopic-Laparoscopic (LLL) were used as the basis of camparision for all models using medical students. Laparoscopic visualization and randomization to the group using the sequence Laparoscopic-Simulated augmented reality were used as the basis for comparision for all models using surgeons.

7.1.2 Findings

A linear regression model and analysis of variance (ANOVA) of the third trial data (using only the total absolute error area) was created to test the hypothesis that

exposure to a simulated augmented reality visualization improves performance in the third trial.

Based on this model (overall significance for the model, $p = 0.189$), significant or near significant (given the small available sample sizes) differences in performance on the third trial were observed in the groups randomized to the sequences Simulated augmented reality-Laparoscopic-Laparoscopic (SLL) ($p = 0.085$), Laparoscopic-Degraded simulated augmented reality-Laparoscopic (LDL) ($p = 0.360$), and Simulated augmented reality-Degraded simulated augmented reality-Laparoscopic (SDL) ($p = 0.163$).

Total scores for surgeons in the Simulated augmented reality-Laparoscopic (SL) group were lower than those in the Laparoscopic-Augmented reality group ($p = 0.103$). This suggests that the learning effect observed in the medical students may also be present in the surgeons.

7.1.3 Transitioning to laparoscopy from simulated augmented reality

If an augmented reality visualization is useful training for a laparoscopic task, is there a better way to transition trainees from using augmented reality to working fully laparoscopically? The groups randomized to the sequences Simulated augmented reality-Degraded simulated augmented reality-Laparoscopic (SDL) and Simulated augmented reality-Both-Laparoscopic (SBL) were created to explore this question. In terms of real surgical training, the first sequence might be implemented by starting trainees learning a task using simulated augmented reality and then transitioning them to an actual augmented reality system, with corresponding degraded visualization, and finally having them work fully laparoscopically. In the second sequence both the augmented reality system and conventional laparoscopic visualization would be made available to the trainees.

The SDL(4100) group performed better than the SBL (7269) group in the third trial ($p = .003$, by t-test, $n = 8$). The SDL group performed better ($p = 0.163$) than all other groups except for the one with the sequence Simulated augmented reality-Laparoscopic-Laparoscopic (SLL) ($p = 0.085$), while no significant effect was observed in the model for the SBL group ($p = 0.618$).

One possible interpretation of this result is the inferior visualization of degraded simulated augmented reality without the laparoscopic view may train students to

use the laparoscopic view better when it became available. Observational findings generally support this hypothesis in that subjects in the SBL group said that they primarily used the laparoscopic monitor but were observed to regularly use or shift their eyes to the degraded simulated augmented reality window. This observation suggests that they may have been using the best of both available visualizations but did not learn the skills needed to work entirely laparoscopically.

7.2 Time to completion

A linear regression model and ANOVA were completed to test for effects on time-to-completion. The overall significance of this model was $p < 0.001$.

The following were associated with a shorter time to completion: simulated augmented reality ($p < 0.01$), degraded simulated augmented reality ($p < 0.01$), later trials ($p < 0.01$), prior experience ($p = 0.02$), right handedness ($p = 0.02$), and having corrected vision ($p = 0.04$)

The following factors were associated with a longer time to completion but were not statistically significant: age ($p = 0.250$) and male gender ($p = 0.147$).

The high levels of significance within this model suggest that time-to-completion is significantly affected by many factors. Additional useful results may be found by a more careful analysis of time-to-completion.

7.3 Incidental correlations

A large regression model was constructed to examine the collected data for incidental correlations.

Video game playing experience, both peak hours-per-week and number of years experience were correlated ($p < 0.10$) with improved performance. A video-game statistic, defined as the product of peak hours-per-week and number of years experience with video games, and a subject's being defined as a "video-gamer", by having a video-game statistic greater than 15, were both strongly associated with improved performance ($p < 0.05$).

Taking more time to complete the task was associated with worse scores ($p < 0.10$). Increased practice time was associated with decreased time to completion ($p = 0.10$).

Performance on any one task was strongly associated with the total score on all three tasks ($p < 0.05$).

Surgeons' scores were better (lower) while medical students' scores were worse (higher) ($p < 0.10$). Surgeons also took less time to complete the study tasks.

7.4 Confounding factors

There are many potentially confounding factors in this study. These factors include explained differences between the randomization groups and possible systematics differences.

Having small sample sizes increases the likelihood that characteristics of subjects assigned to randomization groups are not evenly assigned. However, comparision of other data gathered about each subject, such as gender, handedness, corrected vision, video game playing, prior experience, and age revealed few significant variation between any of the groups. Video game playing experience varied most between groups.

Unexplained and variation is most noticable in the LLL randomization group. This group had the worst performance of any groups, even when comparing first trial performance with other groups performing this trial laparoscopically. The only clear difference between this group and others is that none of the subjects in this group were identified as "video-gamers" and the group had correspondingly low average video-game playing experience. Given the small sample size, this difference is not statistically significant.

Systematic factors may have been introduced via fatigue factors that may vary with the visualization used. In 7.5.1 observations about differences in posture for different visualizations were noted. Additionally, as noted in 7.2, time-to-completion was strongly associated with each visualization so fatigue may have had differing cummulative effects for subjects in each randomization group.

7.5 Observational findings

The findings reported in this section are based on notes recorded during trials. These findings are anecdotal in nature, but may contribute to a better understanding of how to build an effective augmented reality environment to assist laparoscopy.

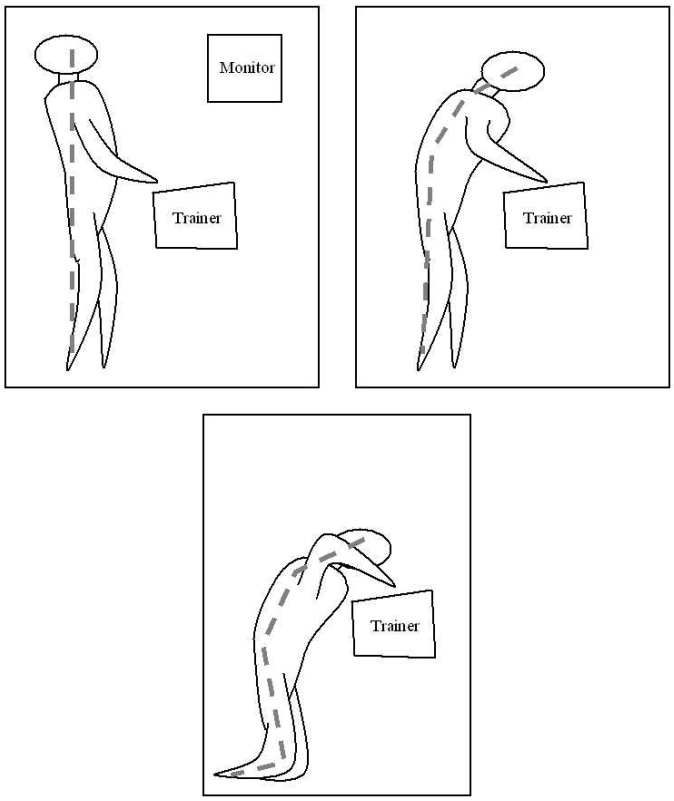


Figure 7.1: Upper left: typical posture for laparoscopic visualization. Upper right: typical posture for simulated augmented reality visualizations Bottom: one of the stranger postures adopted by a study participant.

7.5.1 Posture

Figure 7.1 illustrates the typical body postures of study participents as well as one of the more creative positions adopted. In most instances use of simulated augmented reality was associated with a hunched (lordotic) posture in both the back and neck. Some subjects positioned themselves so that their arms and hands were positioned high relative to their shoulders. These postures are a sharp constrast to the upright alignment of the neck and back and low hand and arm position adopted by most participants under laparoscopic visualization.

7.5.2 Laparoscopic magnification

Early in the study it was observed that some medical students used the magnification effect provided by the laparoscope to maximum effect while others did not. Students who realized, whether consciously or not, that they could get greater magnification

by either moving their work closer to the laparoscope or the laparoscope closer to their work did so. Approximately 25% of the medical students used the maximum magnifcation available by moving their circle or the laparoscope so that the circle nearly entirely filled the camera's view.

I noted that students who used this effect were more likely to have made positive comments about the study task. All of the surgeons used the available magnification afforded by the laparoscopic visualization. More rigorous validation of this observation might be possible through analysis of the video recordings taken from inside the trainer while each subject completeded their tasks.

7.5.3 Circle cutting strategies

Most participants had some difficulty initially picking up the paper. Once they did, they then proceeded to use one of two strategies to cut the circle. In the first approach, subjects first trimmed away large areas of excess material and then successively cut away smaller and smaller areas. In the second approach subjects made an initial cut near the border of the circle. From this cut they began making cuts along the circle's boundary.

Most subjects, upon achieving an approximate cutout of the circle then attempted to trim away small areas of excess material. As an observer, it appeared that in many instances this extra trimming increased error scores by cutting away interior portions of the circle rather than successfully removing outer portions.

Two subjects, both medical students, attempted a unique cutting strategy. They attempted to accurately fold the circle in half through the marked circle and then proceeded to cut out the circle by cutting out the half circle on the folded paper. The first subject who began this approach rapidly abandoned it. The second used this approach for two of his three trials. In the second trial he did not notice that he folded the paper inaccurately and was not satisfied with the result.

7.5.4 Complaints

Of all the particiapnts, six (four medical students and two surgeons) complained of back or neck pain. Of those, five reported the discomfort during or immediately following one of the simulated augmented reality conditions. Both of the surgeons, upon feeling uncomfortable during practice periods, took time before continuing to find more comfortable positions.

Four medical students complained of hand pain and were shown correct hand positions that alleviated most of the discomfort.

One student reported dizziness during a trial which she attributed to standing for an extended period of time. She completed her practice period and took a short break after which she completed the trial uneventfully.

Two students reported that they believed fatigue degraded their performance in the second and third trials.

Three of the surgeons asked if there was some way to see a magnified view while using the augmented reality visualization.

Ten subjects had recorded comments about the the loss of three-dimensional cues under laparoscopic visualization.

7.6 Discussion

This study was designed with the primary goal of identifying whether augmented reality, using simulated augmented reality as a surrogate, is beneficial for performance of a laparoscpoic task and to determine if there is improved learning associated with use of simulated augmented reality . Secondary goals included exploring the affect of low fidelity augmented reality systems via degraded simulated augmented reality and identifying patterns that might spur further research in augmented reality and training methods for laparoscopic surgery.

In order to meet these goals a previously used task was modified. Most validated measures of laparoscopic performance rely primarily on time-to-completion. While these types of measures have been statistically validated, the utility of time-to-completion as a performance goal in surgery is questionable. Accuracy is a measure which makes intuitive sense for relevance to surgery, as it would seem to be more likely to affect success. The redesigned task used a semi-automated measuring system to reach a reproducible score. The two parts of the measuring system requiring human intervention, scanning the cut circle and selecting the center of the circle, were demonstrated to create only small variations in the resultant score.

A significant bias in the type of errors made (inside versus outside errors) was observed under most measurement conditons. Inside errors were more common than outside errors. This may be because leaving extra white material on a black circle is readily apparent while cutting black away from a black circle is not.

For this task, time-to-completion produced comparable results as the error-based

metric. While not truly validated, this suggests that the task might have similar characteristics as the original task described in [Derossis98, Derossis99].

As expected, a significant improvement in performance was observed in medical students using simulated augmented reality visualization as compared to those working laparoscopically. A slight, but not statistically significant, decrease in performance was observed in trained surgeons using simulated augmented reality visualization. Surgeons with significant experience operating laparoscopically are trained to use the beneficial aspects of laparoscopic visualization, primarily magnification, and have overcome many of the disadvantages. This theory is supported by the observational finding that more than one third (three of eight) surgeons, without prompting, specifically mentioned that magnification would have been useful while using the simulated augmented reality visualization.

The findings on the role of simulated augmented reality on learning are not as clear cut. There are three data points the support the hypothesis that use of simulated augmented reality improves learning. First, the SLL randomization group had the best performance in the third trial. Secondly, in the comparision of the SLL and LSL groups on their first and second trials controlling for visualization, the SLL group performed better in these trials. Thirdly, the SL randomization group of surgeons had better performance that those in the LS group. The first two points, taken together, could just as easily indicate some bias present within the SLL group. However, the same effect being present in a second, albeit much smaller, otherwise unrelated group suggests that this effect is due to the randomization order rather than a bias. Additional studies are needed to study this question more carefully. In particular, the hypothesis that simulated augmented reality, and therefore agumented reality, will be more benefical to surgeons completing a more complex task should be explored.

The results from using degraded simulated augmented reality indicate only one thing with any confidence: the advantages of an augmented reality system can easily be lost if the system does not have good enough visual fidelity. Because the degraded simulated augmented reality visualization was created merely to represent a significant degree of degredation without any attempt to quantify the effect, further research is needed to determine what visual quality is needed to preserve the benefits seen using degraded simulated augmented reality .

This study achieved most of the goals set out in the design. While levels of statistical significance were not always found, the relatively high levels of significance in spite of small sample sizes suggest many areas for future research.

Chapter 8

High speed structured light depth extraction

"To visualize the internal processes taking place in the cavities and interstices of the living animal body, it is necessary that
(1) a sufficient amount of light be introduced;
(2) the light rays be reflected back to the eye."

– Philip Bozzini, 1806

This chapter shifts the focus from measurement of the benefits of augmented reality to the problem of whether or not an augmented reality system for laparoscopic surgery is technically feasible. As noted in Chapters 3 and 4, there are currently many augmented reality systems for a variety of applications. The one piece of technology that prevents building an augmented reality system for laparoscopy is the lack of a continuous, real-time, three-dimensional data source. In many ways this problem is similar to the one faced by Bozzini: how to get enough light into, and out of the body.

A literature search on the term "three-dimensional laparoscope" and its variations turns up a suprisingly large number of published articles. Nearly all of these articles describe various methods for getting a *stereo* view of the operative site. It is important to note that while the results of these systems are perceived as being "three-dimensional", no explicit three-dimensional information is recovered. Explicit three-dimensional information, that is the location in three-dimensional space of pixels or groups of pixels in the laparoscopic image, is needed in order to correctly redraw the laparoscope's view from the surgeon's perspective.

In this chapter I will give a general overview of three-dimensional scanning techniques that give explicit three-dimensional information with an emphasis on the trade-offs involved in picking a particular method for using augmented reality in laparoscopy. A more detailed overview of structured light techniques will be given.

Then a new method for high-speed image processing will be presented. Results from simulation and a prototype implementation will be shown.

8.1 Considerations for image acquisition in laparoscopy

For the application of scanning the interior of a patient's abdomen during surgery, there are several important considerations which guide the decision of which method to use.

1. Non-contact – the method should not touch or disturb tissues
2. Small working volume – the entire working volume bounded by a working distance between 2cm and 10cm and a 90° field of view
3. Curved surfaces – abdominal organs have mostly smooth contours with few sharp angles or large flat surfaces
4. Light penetration – light penetrates to varying degrees into tissues depending on light wavelength and tissue composition
5. Light reflection – tissue surfaces tend to have areas of strong specular reflection. Specular highlights obscure inherent surface characteristics and move with movement of the camera or the light source causing errors in several methods
6. Surface texture – abdominal organs have little surface texture. Some surface texture is seen due to superfical blood vessels, but this usually looks like a subtle diffuse streak on the surface
7. Movement – abdominal organs are almost always moving. Pulsations due to arterial pulse (between 50 and 120 per minutes), respiration (5 per minute) are expected with additional movement caused directly by the surgery
8. Adequate refresh – user performance for tasks performed under visual guidance drops rapidly at refresh rates less than 20 per second

8.2 Three-dimensional scene acquisition

A large variety of techniques have been proposed and implemented to create computer models of three-dimensional scenes. This includes specially built scanners using sound

or light, as well as algorithms which can be applied to one or more images. The following are some of the broad categories of methods:

1. Sonic/ultrasonic rangefinder
2. Laser rangefinder: time of flight
3. Depth from stereo/multibaseline stereo
4. Shape from shading
5. Structured light
   (a) Triangulation/multibaseline
   (b) Focus/Defocus
6. Depth from focus/defocus
7. Shape from motion

A summary of some of the characteristics of these methods relative to the design criteria is shown in Table 8.1.

Of these possible methods, only depth from stereo, structured light (both triangulation based and defocus based), and shape from motion have been successfully implemented for real-time operation.

Depth from stereo is perhaps the most widely used method used to generate three dimensional information from two or more sets of images of the same object from slightly different perspectives. The mathematics of triangulating position are quite similar to the techniques used for structured light and are discussed in 8.4.1. The difficulty in these methods comes from the identification of corresponding features or regions in both images. These techniques therefore work best on highly textured surfaces. Curves can present problems to these methods because the visual meeting point between a curved surface and a surface it occludes is a readily identifiable feature, but its location on the curved surface changes depending on the camera view. A depth from stereo method has been applied to depth extraction of cerebral surfaces through a surgical microscope by matching intensity windows in the image pairs [Fleig01]. This system is not real-time, requires some manual intervention, and fails in regions with few features or with specular highlights. The surfaces of the brain, in contrast to surfaces in the abdomen, have many distinct features because of foldings and blood vessels visibile on the surface.

Real-time shape from motion methods have been demonstrated [Chiuso02]. These methods identify features in sequential video frames and then use the small changes in position of features to calculate the three-dimensional motion of the features and the camera. These algorithms are not likely to be successful in laparoscopy because they depend on having many repeatably identifiable surface features. These methods also will fail in the presence of a high degree of specularity because specular highlights move across surfaces as the camera and light source are moved. An additional problem with these methods is that unless objects of known scale are introduced into the video sequence, the scale of objects is indeterminant, and only relative scale information can be determined[1]. These techniques are most frequently applied to problems where the video sequence comes from static scenes with multiple objects of known size.

Depth from focus methods (with or without structured light) have been implemented for real-time use [Nayar96]. Depth from focus and defocus systems that are not real-time frequently use a single camera with a calibrated focus adjustment. The depth value for every region of a scene is selected based on which image in that portion of the scene is most in focus. The real-time systems work by comparing the degree of focus observed in two or more optically aligned cameras observing the same scene. Because methods of measuring the degree of focus (or defocus) in a picture depend on some a priori knowledge about the expected spatial frequencies in the scene, projection of a light pattern with know frequencies into a scene can be very useful. The utility of projection of these structured patterns into the scene can be enhanced if the projected light is focused between the focal planes of the two cameras. This technique is not believed to be easily applicable to laparoscopy because of the scale needed for laparoscopy, the potential for large quantities of specular reflection, and because tissue surfaces are semi-tranparent which creates an extra blurring of the imposed patterns.

A remarkable method for real-time structured light depth extraction has been developed by Olaf Hall-Holt and Szymon Rusinkiewicz [Hall-Holt01][Rusinkiewiczi01]. Their system uses a standard video camera and video projector and produces a fairly dense range estimate at 60Hz (every frame at the speed of the camera) with a claimed accuracy of 0.1mm within a 10cm deep scanning volume. Unlike other structured light triangulation methods, their method can compensate fairly well for motion. Their method uses the previous three captured video frames to unambiguously identify

[1]For example such systems would be unable to distinguish between a video sequence of cars in a parking lot and model cars scattered on the floor

Technology	*Constraint*						
	Size	Curved Surfaces	Light Penetration	Light Reflection	Surface Texture	Movement	Refresh Rate
Sonic/ultrasonic	No	No	No[a]		Yes	Yes	No
Laser, time of flight	No	Yes	No	No	Yes	Yes	Yes
Stereo/multibaseline	Yes	No	Yes	No	No	Yes	No
Shape from shading	Yes	Yes	No	No	No	Yes	No
Light Triangulation	Yes	Yes	No	No	Yes	No	Yes
Focus/Defocus	Maybe	Yes	No	No	Yes	Yes	Yes
Shape from motion	Yes	No	Yes	No	No	No	No

[a]Sonic methods are disrupted by materials that strongly absorb or reflect the sonic beam

Table 8.1: This table identifies the problems of applying different three-dimensional scanning methods to laparoscopy.

stripe boundaries. The position of found stripe boundaries is used to triangulate depth values. A more complete discussion of this method appears in the following section.

8.3 Structured light triangulation methods

The term "structured light" is often used to specifically refer to structured light triangulation methods, but there are other structured light methods. These include the depth from defocus method mentioned in the previous section. Many laser scanning techniques can be considered structured light methods except that the structure may only be a single point of light or a stripe. Section 8.4.1 describes the mathematics of calculating depth in depth-from-stereo or structured light triangulation methods. Some of the information in that section is assumed in this discussion.

Structured light is a widely used technique. It seems likely that it would be useful for the following reasons:

- It works on curved surfaces
- It works independent of surface texture[2]
- It can work with some specular components in the scene

[2]Structured light methods usually make some assumptions about how much variation in surface properties can be present in the scene. If this tolerance is exceeded, the method may not be able to recover the structure from the scene and will then fail.

Light structures can be any pattern that is readily (and ideally unambiguously) recovered from the scene. The difference between the pattern injected into the scene and the pattern recovered gives rise to the depth information. Stripe patterns are frequently used because of the easy way that they can be recognized and because of some mathematical advantages discussed in Section 8.4.2. Many methods have been proposed to distinguish projected stripes from one another. These include color coding [Boyer87, Caspi96, Davies96], pattern encoding [Posdamer82] (like barcodes for each stripe), and temporal encoding. Color coding schemes rely on the ability of the projector to accurately produce the correct color, the surface to reflect back the correct color, and the camera to record the correct color. While projectors and cameras can be color calibrated, this puts serious constraints on the appearance and reflective properties of objects to be scanned. Pattern encoding schemes assume enough spatial coherence of the object being scanned that enough of the code can be found to identify a stripe. Temporal encoding of stripe patterns is very popular because it makes few assumptions about the characteristics of the surface, including possibly working with surface texture. Motion during the scanning sequence can be a significant problem.

A simple and effective way to temporally encode many stripes is to project multiple patterns into the scene that use a binary encoding for each stripe. Assuming no movement of the scene or the scanner, the stripe that each pixel of the image belongs to is encoded by whether or not the pixel in each image is illuminated or not. This method reduces the number of projected patterns needed to encode n stripes to $log_2 n$. A variety of sequences have been developed to reduce the potential for error by misidentification of a pixel's status in one or more images.

Several systems use the position of the edge of stripes rather than the position of the stripe itself for triangulation. One of the advantages to doing this is that subpixel accuracy for this position can be achieved. Another advantage of this method is that the edges may be easier to find in images with variations in texture than the stripes themselves. If the reflective properties of the anticipated scene are relatively uniform, the point of intersection of a plot of intensity values with a preset threshold, either in intensity value or in the first derivative of intensity, can be sufficient to find edges with subpixel precision. If a wider range of reflective patterns are anticipated, structure patterns can be designed to ensure that, in addition to labeling pixels as to which stripe they belong to, each of these edges is seen in two images but with the opposite transition (that is the transition from stripe one to two is on-off in one image and

off-on in a later image). The point where a plot of intensity values at this edge cross each other can be used as a subpixel measure of the position.

The real-time system developed by Rusinkiewicz and Hall-Holt uses edges in an even more clever way. They recognized that a stripe must be bounded by two stripes and, over a sequence of four projected patterns the status of both stripes can be changed arbitrarily. This means that 256 (2^{2^4}) different coding sequences can be projected over the sequence of four frames. After removing sequences in which the edge cannot be found (because the neighboring stripes are "on-on" or "off-off") in two sequential frames and where a stripe remains "on" or "off" for the entire sequence, the latter sequence is removed to avoid the possible confusion of their edge-finding method of these stripes with texture on the surface of the object. After such considerations, one hundred ten (110) stripe edge encodings may be used. This method could potentially be applied using more neighboring stripes to encode a particular stripe edge or a larger number of frames. For n frames and m stripes to encode each boundary the number of possible stripes that can be encoded is approximately proportional to $(m^2)^n$. It should be noted that to project x stripe patterns $O(\log x)$ frames are needed as in the case of binary encoded stripes except the scalar multipliers on this limiting function are generally much smaller for this technique as compared to binary encoding.

This system uses the rate of change (first derivative) of intensity to find boundary edges with the sign determining what type of transition is observed. Efficient encoding demands that many edges be encoded by sequences in which the edge is not visible in every frame. The existence and approximate position of these "ghost" edges is inferred and matched to the nearest found edge or hypothesized edge in the previous frame. The encoding scheme limits the number of frames in which an edge can exist as a ghost, but the matching process limits the motion of objects in the scene to not more than half of the average stripe width between frames. This system may have a lower threshold for motion if objects have high-frequency textures or any sudden changes in surface texture that would be misinterpreted as an edge. The method might also fail for objects that are moving in or out of shadows (that is, the camera sees a point on the surface, but the projector cannot illuminate it). The primary application for this system is building high-resolution three-dimensional models of objects. A modified iterated closest point (ICP) algorithm is used to match and register the range images acquired from frame to frame in order to build up the dataset.

The speed of structured light triangulation systems is limited by three major fac-

tors. These are the rate at which structured patterns can be projected, the rate at which an image of the scene with the structured pattern present can be captured, and the rate at which the captured images can be processed. The method of Rusinkiewicz reduces the demands on the projector and camera by use of sophisticated processing and judicious use of assumptions of spatial and temporal coherence. An alternative approach explored here is to use simple, high-speed, hardware acceleratable algorithms, while placing greater demands on the camera and projector portion of the system both to compensate for motion and to achieve a useful rate of output for augmented reality.

Many currently available displays including LCD and micro-electro-mechanical systems (MEMS) (e.g., Texas Instrument's DMD^{TM}, digital micromirror display) based devices are used to project images with refresh rates over 60Hz. These devices are frequently used to project color imagery by using small time slices to project component colors of the total displayed imagery. The underlying display in these devices is capable of projecting monochrome images (black and white, not grayscale) at over 40kHz (for a DMD based device)[Hornbeck88]. High-speed digital cameras are available that capture images at over 10kHz[3]. While most of these devices are simply used for recording short high-speed sequences, image capture and analysis cards are available which are capable of handling the data stream and performing simple analysis at high data rates.

Given the devices that are currently available, the remaining challenges include the development of algorithms that can be executed at high speed, alogrithms that can deal with texture or the absence of texture, specularity, and small amounts of motion. An additional challenge is to build a device that satisfies these conditions and is compatible with laparoscopic surgery. The remainder of this chapter focuses on the software issues, while the next chapter examines the issues in building the hardware device.

8.4 Mathematical model for recovering depth

This implementation of structured light can be thought of as a special case of depth from stereo. Instead of a second camera, a projector is used so that disparity can

[3]While cameras are advertised that capture at these extremely high rates, most cannot capture "full frame" images at that rate. Currently advertised cameras can record sequences at 1kHz with a resolution equivalent to standard video (640x480, non-interlaced).

be defined as the difference between where the projector puts a feature and where the camera finds the feature (rather than being the difference between where the two cameras find a scene feature). First a simple mathematical model of depth from stereo will be discussed, then the changes needed to apply this model to a camera and projector will be discussed.

8.4.1 Depth from stereo

A schematic diagram of a single scan line depth from stereo depth extraction system is shown in Figure 8.1. In such a simple case of depth from stereo, the cameras are coplanar, are directed parallel to each other, and are ideal "pin-hole" cameras with identical internal parameters.

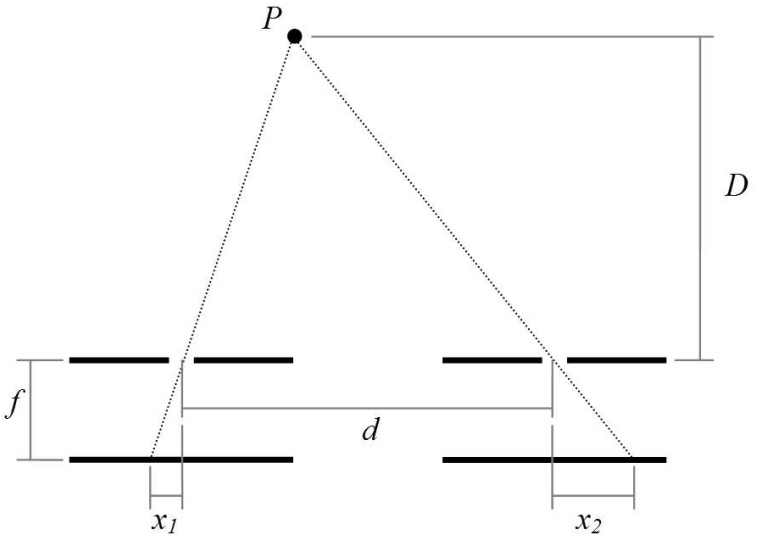


Figure 8.1: This figure shows a schematic diagram of two one-dimensional pinhole cameras observing a point in space. P is the point in space observed, d is the separation between the cameras, D is the distance to P, f is the distance between the pinhole and the imaging plane of the cameras, and x_1 and x_2 are the positions of the observed object on the imaging planes of the two cameras.

In this case the relationship between the depth of a point seen by both cameras (D), focal distance (f), camera separation (d), and the observed position of the point on the image planes of both cameras (x_1 and x_2) (by similar triangles) is

$$D = \frac{df}{x_1 - x_2} \tag{8.1}$$

The denominator contains the term $x_1 - x_2$, which is the disparity or difference

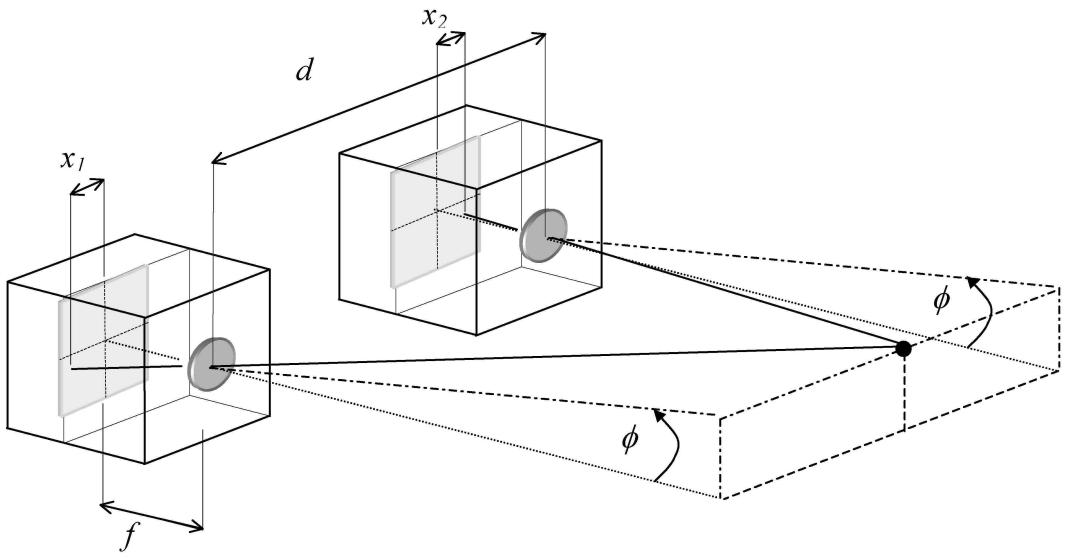


Figure 8.2: Schematic diagram of two two-dimensional pinhole cameras observing a point in space. d is the separation between the cameras, D is the distance to the observed point, f is the distance between the pinhole and the imaging plane of the cameras, and x_1 and x_2 are the positions of the observed object on the imaging planes of the two cameras. The angle of elevation, ϕ, must be identical in the two cameras.

of position of the object in the two cameras.

This relationship holds if the cameras have two-dimensional imagers rather than the one-dimensional imagers that were initially assumed. This can be demonstrated by observing that the angle of elevation between the ray starting at the pinhole and ending at the point in space (ϕ), is the same if the earlier assumptions about the cameras are maintained. The implication of this observation is that if the point is observed at a position (x_1, y_1) in the first camera, its position in the second camera must be at (x_2, y_1) – that is its position in the second camera is limited to a single horizontal line. If some a priori information is known about the distance to the point, the portion of the image that needs to be searched to find that point in the second image can be further reduced. This constraint is also known as the epipolar constraint for rectified images.

When the assumption of parallel and identical imaging planes is broken, the relationship becomes more complicated. Fortunately, points projected onto a two-dimensional plane can be reprojected onto a different plane using the same center of projection. Such a reprojection can be done in two or three-dimensional homeogenous coordinates and takes the following form

$$\begin{bmatrix} x_r \\ y_r \\ w \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{8.2}$$

Since this is a projection in homeogenous coordinates, the correct coordinates on the reprojected plane (x_c, y_c) will then be

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} \frac{x_c}{w} \\ \frac{y_c}{w} \end{bmatrix} \tag{8.3}$$

Reprojection of this type is performed very efficiently by current graphics and image processing hardware. In the special case where reprojection is done onto parallel planes, w becomes a constant for all coordinates (x, y) in the original image. An important implication is the observation that, because the operation of perspective projection preserves straight lines as straight lines in the resulting image, that an object observed in a particular location in one camera's view must be located somewhere along a straight line (not necessarily on a scan line) in the other camera's image. This constraint is known as the epipolar constraint.

Further simplification can be achieved if reprojection is done to planes that are nearly parallel. This assumes that the original planes of projection are very similar to the rectified ones. That is, the center of projections are approximately aligned. In this case, w varies little in the region being reprojected. Therefore, rough approximation of the depth calculation, without using any reprojection, is simply (where C_n, k, and K are appropriate selected arbitrary constants)

$$D = \frac{K}{C_1 x_1 + x_2 + C_2 y_1 + C_3 y_2 + C_4} \tag{8.4}$$

This approximation will be used in Section 10.4.1 as the basis of a simple calibration for the prototype system. As it is a rough approximation, it should be applied with great caution if a high degree of accuracy is needed.

While rectification, particularly the reprojection in homogeneous coordinates can be done in hardware, it is not always advisable to do so. Practical implementation of reprojection usually entails a loss in image quality. Minor corrections usually cause little loss of information, but large differences in the position of the planes of projection can cause a loss of a large number of pixels in some portions of the resultant image and smearing of pixels across wide areas in other portions of the

resultant image. In these cases it is wise to find positions of objects in the original images, and reproject only the found (x, y) coordinates of those objects.

Violation of the assumption that the camera is a pin-hole camera can also be corrected. There are a variety of methods to measure the distortion in camera images by imaging known objects. Once the nature of the distortion is known there are several techniques that use graphics or image processing hardware to rapidly correct for these distortions to create a distortion-free image. In most cases it is easier to correct for camera distortion before looking for objects in the images so that the epipolar constraints are straight lines and not the curved lines that are possible in camera images with distortions.

8.4.2 Projector and a camera

When one of the cameras from the model discussed in the previous section is replaced with a projector, the meaning of some of the steps and variables changes. These changes have several important implications.

Camera one in Figure 8.2 is replaced with a projector that is essentially an inverse of a pinhole camera, like a *camera obscura* projecting on the wall of a darkened room. The projector by itself gathers no data about the point in space. If a single ray of light is projected out into space, and no other light is available, then the second camera can now unambiguously see the point where that ray strikes a distant surface. In this situation, the coordinate of the light being projected is x_1 and the point where the camera sees that point of light is x_2. If we then move to two-dimensional cameras and projectors, the ray (x_1, y_1) is illuminated by the projector, illuminating the surface at a point which is then seen by the camera at the point (x_2, y_2). The mathematical relationships of the previous section then hold.

Reprojection and rectification, as described in the previous section, take on a new meaning for the projector. Rather than "undistorting" an acquired image, the projected image can be "predistorted" to account for the projection optics and variations in the position, orientation, and characteristics of the relative position of the camera and projector combination. This predistortion should be the inverse of the undistortion that would need to be applied if it were a camera rather than a projector.

Another implication of the epipolar constraints in the case of the projector-camera combination is that more than one ray can be illuminated simultaneously and they can easily be distinguished. In the simplest case, where the camera and projector have identical internal characteristics, are pointed in a parallel direction, and have

coplanar planes of projection, the epipolar lines run parallel to the x axis. A ray projected to (x_1, y_1) will only be seen along the y_1 scanline in the camera. A second ray projected through (x_2, y_2) will only possibly be seen on the y_2 scanline in the camera. In practice it is usually simplest to project vertical stripes (assuming that the camera and projector are horizontally arranged). Because each scanline in the camera corresponds to a single horizontal line in the camera, the found x position of the stripe projected through position x_s can be converted to depth using Equation 8.1, after appropriately substituting x_s

$$D = \frac{df}{x_s - x} \tag{8.5}$$

8.5 Summary

Table 8.1 and the text of Section 8.2 describe the strengths and weaknesses of several different methods for acquiring the three-dimensional structure of a scene. Of these methods, structured light-based triangulation techniques appear most promising for real-time acquisition in the laparoscopic environment. Temporally encoded light stripes are a logical choice for structure patterns because of the ease which with these features can be recovered.

Several challanges remain for applying this technique to the laparoscopic environment. Table 8.1 lists three barriers to using these techniques. These are penetration of light into tissues, dealing with shiny, highly specular surfaces, and movement in the scene. The following two chapters describe hardware and software methods to overcome these barriers. Because light penetration into tissues can irregularly blur the edges of stripes, the center of each stripe should be used to recover depth rather than the edges.

Chapter 9

Depth extraction hardware[1]

"A main requirement is that the flame remains of constant intensity and in a fixed position. A good wax candle installed in a sheath of iron or brass as follows seems adequate:"

– Philip Bozzini, 1806

In 1806 Bozzini was concerned with the question of what type of light source to use to illuminate the interior of patients. He suggested two other sources of illumination, sunlight and an Argand lamp. He rejected both in favor of a candle due to the complexity they would add. In the device described here there are many similar design considerations.

This chapter begins with a discussion of the major overall design considerations and early prototype. Later sections will discuss the overall design of the depth extracting laparoscope, the miniaturized projector, and the optical portion of the laparoscope.

9.1 General design considerations

In addition to the constraints discussed in Chapter 8, two additional design constraints on the laparoscopic depth extraction system should be considered. These are operating room compatability and fail-safe mechanisms.

Operating room compatibility has many aspects. An operating room compatible device should have small enough size and weight for intraoperative use. The device, or at least appropriate portions of the device, must be sterilizable. Finally, the device should not produce excessive heat or make loud noises.

[1]**I am particularly indebted to Kurtis Keller for his contributions to this section.**

While one hopes that a surgical device will never fail, when it does it is important to design a device so it fails gracefully. For a depth extracting laparoscope it should always be possible to simply remove the device and insert a conventional laparoscope. A fail-safe mode in which this device functions like a conventional laparoscope would be ideal.

9.2 General system design

From the background given in Section 8.3 it is apparent that three major system components are needed. These are a camera, a computer for image acquisition and processing, and a projector. A fourth component is a specialized optical laparoscope. Each of these components, while dependent on the function of others, should be considered as a separate element. A schematic diagram of these elements in a modular system is shown in Figure 9.1.

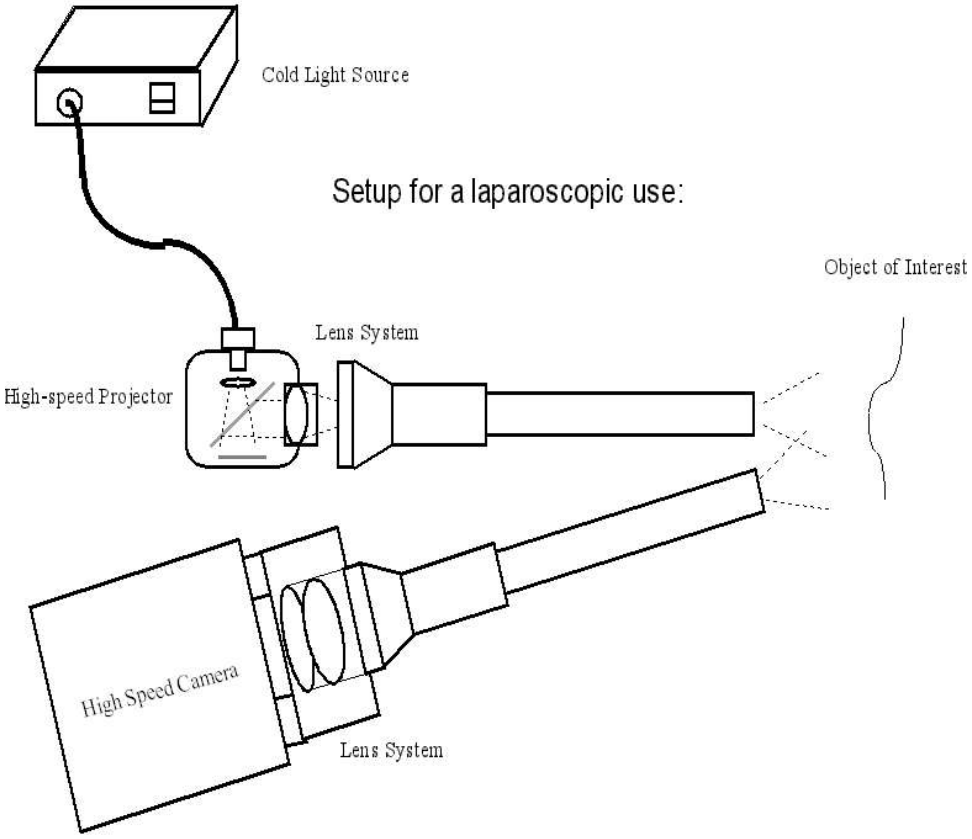


Figure 9.1: Schematic diagram of the needed components for a structured light triangulation system for laparoscopy.

Modularity is an important aspect of this system from both a design perspective as well as from an operating room compatibility point of view. Some system components, particularly those containing sensitive electronics, may not be sterilizable. It is therefore important to be able to easily disassemble the device into sterilizable components. The projector should be designed to use a separate light source, like laparoscopes currently in use, and be connected by wire and fiber optic cables to make it easy to reposition.

Synchronization is an important component of the overall design. Synchronization is needed between the camera and projector, between the camera and image processing hardware, and between the projector and image processing hardware. Our past prototypes have delayed acquisition to achieve synchronization resulting in a system that operates much slower than it might run with better synchronization.

It seems likely that this system is best built with two cameras rather than one. Only monochrome information is needed to find and process structure, while full color imagery is the standard for viewing surgery laparoscopically. Full color images can be used as texture maps for the extracted and rendered three-dimensional surfaces. Using two separate cameras also provides a simpler fail-safe mode. If the depth-extraction camera fails, the "texture" camera should continue to operate and the whole system should perform like a standard laparoscope.

In order to ensure correct alignment of imagery with the projected texture, the system should have a software mechanism to allow rotation of the full color imager. On 0°-angled laparoscopes this rotation is equivalent to rotating the scope and camera together, but for off-axis scopes, camera rotation is an extra degree of freedom in positioning the scope that would be lost if the orientation of the camera were fixed.

Future prototypes using a single-tube design will likely need a combined projector and camera module that can be used with a separable laparoscope. Addition of non-volatile RAM storage on both of these components could be a useful way to store and access calibration information. Positive mechanical interlocks should be used on both components to insure reproducible relative positioning.

9.3 Prototype systems

A first prototype system was constructed in 1998. The purpose of this system was to demonstrate that range acquisition could be integrated into an augmented reality visualization aimed at augmented reality. This system used a large standard DMD-

based video projector that had been modified to project through a steel tube 1.5 inches in diameter(Figure 9.2). A range image was generated approximately every four seconds. Images were acquired by a "lipstick" camera and captured on a SGI O2. After a sequence of images was captured and processed, the range image was then transfered to the SGI Infinite Reality system running the augmented reality environment via ethernet. This system was never successfully tested with biological tissues. An example of the system's output in the augmented reality environment is shown in Figure 9.3.

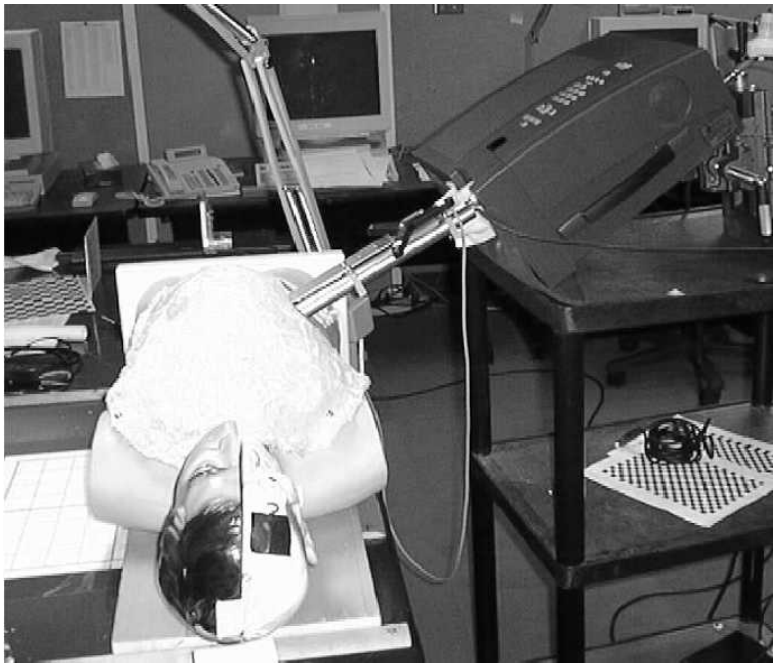


Figure 9.2: Prototype structured light depth extraction system built for an augmented reality environment. Steel tubes, 1.5" in diameter, were used as stand-ins for laparoscopes. A conventional DMD-based video projector was modified to project through the tubes. A small camera acquired imagery through a second steel tube.

The next group of prototypes constructed were aimed at building a depth-extraction system that would be more operating room compatible. These systems used a large standard digital micro-mirror display (DMD) based video projector. The projector was modified to project light into a flexible endoscope. We hoped to make it easier to repositition the laparoscope by projecting structured light patterns through the flexible endoscope. In the first version, images were captured through a standard laparoscope. Light throughput was very poor and the system was then modifed to collect images through a large aperture camera at the tip of the instrument. These

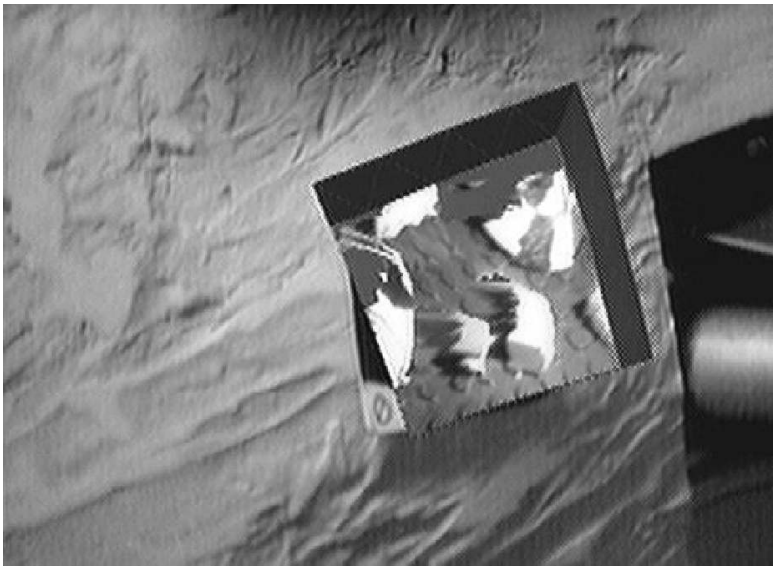


Figure 9.3: View through the head-mounted display of small toys acquired using structured light techniques and rendered in an augmented reality environment using structured light techniques.

two versions are shown in Figure 9.4.

Using this system we began to collect data. Our primary concern was whether it would be possible to recover projected structured light patterns inside the abdomen. To test this we projected simple structure patterns (a sequence of squares) through our system and compared the locations of the recovered squares when the tissue was fresh, and after it was coated with a thin layer of talcum powder. The talcum powder makes all surfaces appear uniform in color and reflective properties. Talcum coated tissue has nearly Lambertian reflective properties and produces no specular highlights. We found that the structured patterns could be found at nearly identical positions in both the coated and uncoated images (An example of these images is shown in Figure 9.5). Measurement of the intensity of the viewed surface under both the illuminated and non-illuminated conditions was needed to automatically locate structured patterns on the uncoated surface.

This system had significant problems. Despite the addition of the flexible endoscope to transmit light from the projector to the laparoscope, this device was still difficult to reposition. The flexible endoscope also had very low optical efficency, so very little light was projected onto the tissue surfaces. The final problem with this system was that the flexible endoscope had a relatively small number of optical fibers. The small number of fibers meant that despite the high resolution of the projector,

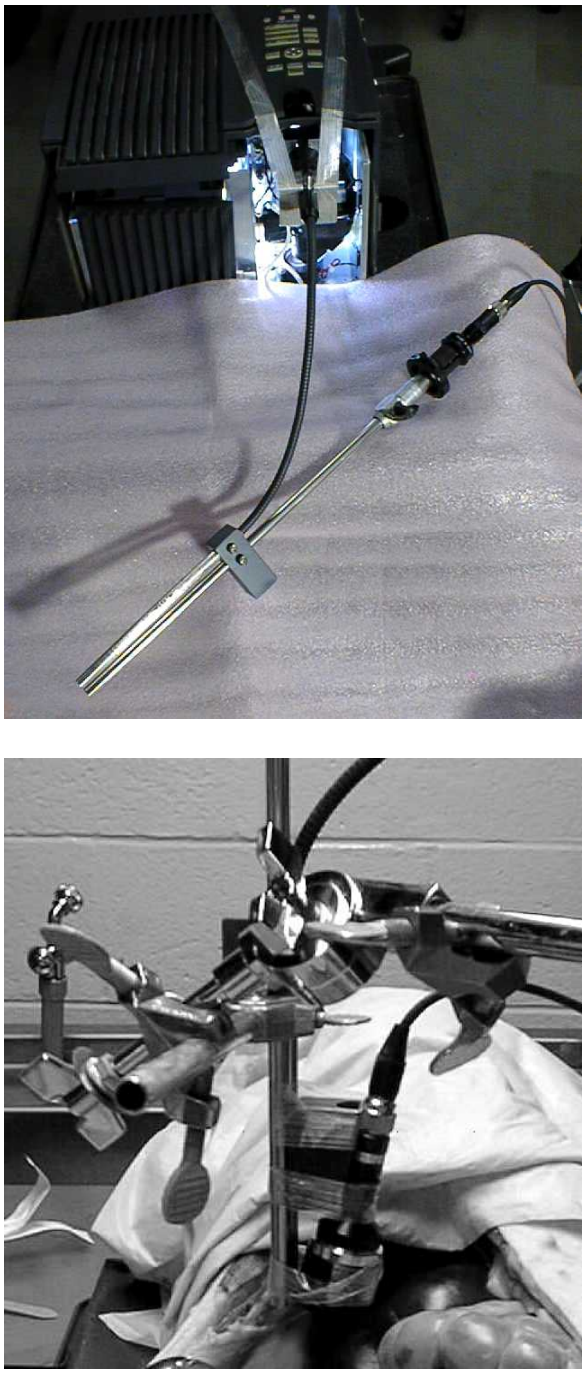


Figure 9.4: These are two prototype depth extraction systems used in early tests of depth extraction for laparoscopy. Top: prototype using standard laparoscope for the acquisition camera and a flexible endoscope to convey the structured light from the DMD-based projector. Bottom: a prototype using camera with a large lens system near the tissues for acquisition, shown in use for experiments in animal cadavers. This system was used to acquire the images in Figure 9.5.

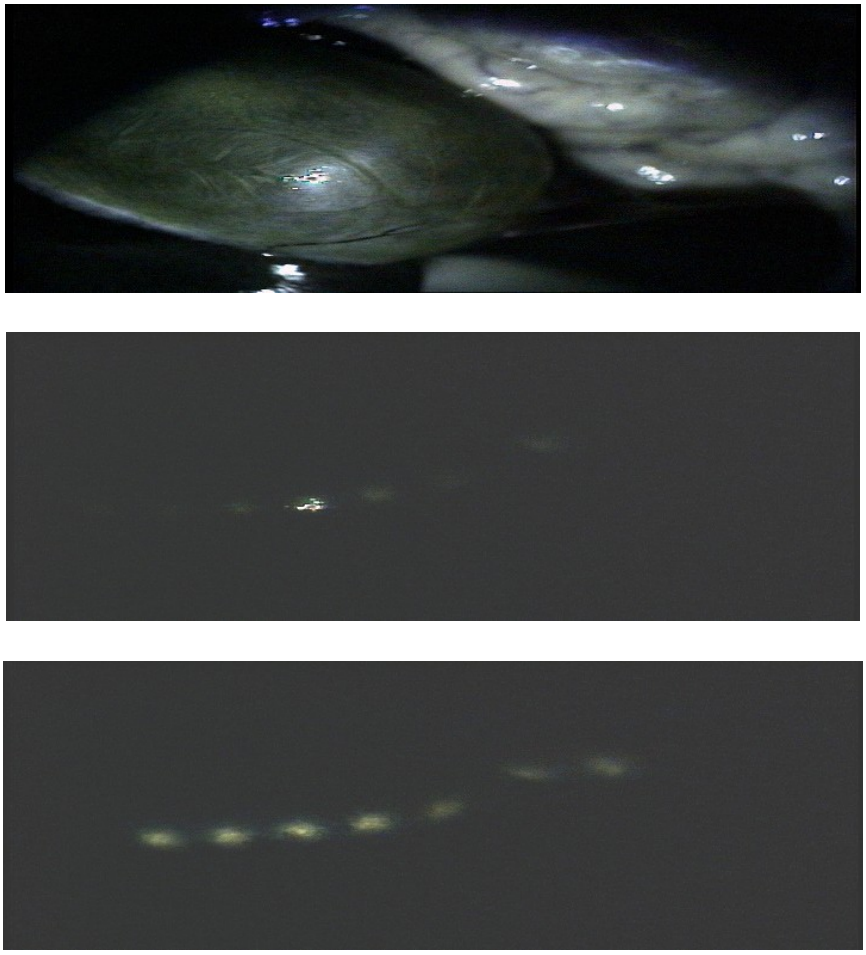


Figure 9.5: Top: fully illuminated raw tissue, Middle: tissue with a sequence of squares as a structure pattern projected onto the same tissue. The squares are barely visible except near a specular highlight which partially obscures one square, but are detectable in the raw image data. Bottom: the same pattern on the same tissue after a thin coating of talcum powder was applied. The specular highlight is no longer visible, and all of the squares are easily located.

only very low resolution images could be projected. The low resolution images that could be projected were usually significantly blurred as well. This system used a SGI O2 to provide the video source for projection and also for video capture.

After completing these experiments, the image processing methods described in 10.2.1 and 10.2.2 were developed based on the tissue properties observed in these experiments. In order to test the effectiveness of these algorithms we built an interim prototype. This system used a standard DLP-based projector, and images captured through a laparoscope. In order to make the system more portable than previous systems the projected patterns were supplied by a video signal from a laptop computer and video was captured to MiniDV tape. The collected images were than taken back to the lab and analyzed offline in order to test the image processing algorithms.

The final and current prototype makes use of a miniaturized, high-speed projector. The projector and image capture and processing are all controlled by a Matrox Genesis card in a PC. This system has gone through several revisions. Initially we both projected the images and captured images through a laparoscope, and used a continuous illumination source. We then switched to a strobed light source for the projector. Finally, we currently only project images through the laparoscope. This system has been calibrated and tested on biological tissues.

The miniature projector in this system was custom-built from a VGA (640x480) ferro-reflective liquid crystal display. This display originally was part of a head-mounted display that showed imagery at 30 frames per second in frame-sequential color. The illumination inside the display was achieved with colored LEDs that were turned on only while the appropriate color field was being displayed on the ferro-reflective display element. Ferro-reflective liquid crystal displays need to switch rapidly or the crystals migrate across the pixel resulting in unequal illumination. The driving electronics of the display deal with this problem by switching every pixel to the opposite polarity after each color field. Thus this display operates at 180Hz (30Hz times 3 colors 2 polarities) for monochrome imagery, as long as images and their inverses are useful. When a continuous light source is used the projected imagery is not apparent to the eye. Most people are aware that the light is pulsing and some structures may become evident with rapid eye movement or blinking. However, when the projected imagery is viewed with a synchronized camera with short exposure time, the projected imagery becomes apparent.

The intensity of the projected images as well as the intensity of images captured by the video camera have been an ongoing problem. Initially we used a continu-

ous illumination source (200 Watt halogen), but debugging was difficult because the projected images could not be seen directly. The camera we were using did not synchronize reliably with the projector, and very low intensity images were captured. Because the camera only operated at 30Hz, we obtained and integrated a strobe light as the illumination source for the projector. This resulted in projected imagery that is visible to the unaided eye and simplified synchronization although the timing must be manually adjusted every time the system is used. While the strobe light delivers less total light per second, it is delivered in very short flashes (20μsec) and only when the display element is displaying a particular image. The strobe makes a loud noise when it is active. Despite adding the strobe, structure patterns were barely above the background noise. We removed the laparoscope from the camera and captured significantly higher contrast images.

9.4 Miniaturized projector design

Most currently available video projectors are designed to project on relatively large and static display surfaces. Consequentially, they are built with large optical apertures and narrow depth of field. Size and weight considerations only affect portability. A projector custom-built for laparoscopic depth extraction should have a very different set of design criteria. Optical apertures should be small to match those of laparoscopes. Size and weight are very important concerns for usability.

Most available displays also integrate a high intensity light source into the projector. For most applications this helps the projector work as efficiently as possible. By separating the light source from the projector we can derive several benefits. The light source, heat sinks, and power supply take up a large amount of space and add significant weight to the device. By separating the light source and display engine the weight and heat sinks needed on the projector can be dramatically reduced. A more powerful light source can be built into its own container with filters to remove excessive infrared and ultraviolet light and heat sinks. A fiber optic cable can be used efficiently to bring the light to the display engine. Similar fiber optic cables and light sources are already used in the operating room to light the laparoscope's view.

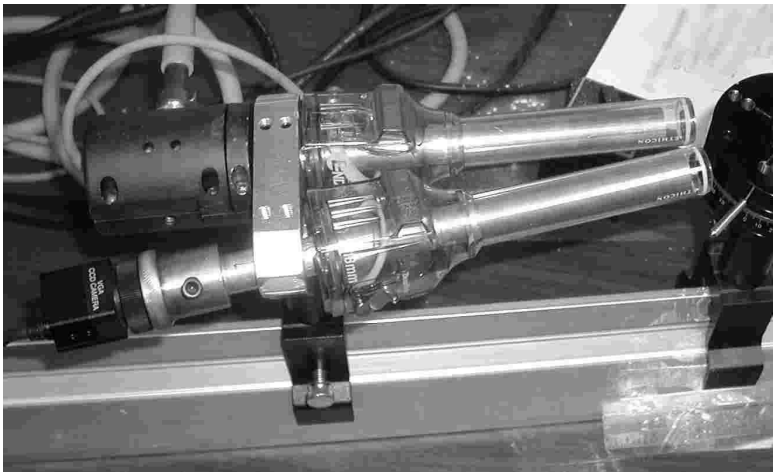


Figure 9.6: Current prototpye depth extracting laparoscope. Separate tubes are used for projecting structure patterns and acquiring imagery.

9.4.1 Current prototype

The current prototype miniaturized projector is a cylinder 60mm in diameter and 75mm long and weighs 380 grams. The display engine core is a ferro-reflective LCD display with a resolution of 640 by 480 pixels. It was designed to operate as a 30Hz display with frame sequential color with two bits per color. Inverted images are displayed after the positive image is displayed to prevent the liquid crystal matrix from migrating. While the display is switched at least 360 times per second (30 Hz, 3 colors, positve and inverse, two bits per color), timing of the second bit is not reliable so synchronization above 180Hz is not currently possible. The driving electronics are sensitive to environmental factors.

The projector is designed to be rigidly mounted to a standard laparoscope eye-piece. It has an adaptor for a standard laparoscopic light source as well. In order to be able to see the projected imagery the light source is a synchronized strobe light (Figure 9.6).

The projector has a contrast ratio of 80:1 which is much lower than the 300:1 to 1000:1 contrast ratios available in projector displays. For this reason the projector produces low intensity imagery without enough contrast. The projected imagery is brighter on the periphery than in the center, and the driving electronics fail intermittently. The timing of the driving video signal, and, to a lesser degree, the timing within the display itself drifts, making synchronization more difficult.

9.4.2 Future prototype

There are many important design decisions for building a better minaturized projector for the laparoscopic high-speed depth extraction system. The most important of these are the choice of display engine technology and light source. These choices cannot be considered entirely independently.

9.4.2.1 Display engine

There are currently four available technologies that are possible choices for a display engine. These are transmissive LCD, reflective LCD, MEMs array, and scanning MEMs devices.

LCD technologies have a significant inherent loss of optical efficiency because light projected through these devices must be polarized twice. Polarizing the light may be a useful technique to dimish the effect of specularity, however polarization costs approximately 60% of the illumination. Collimation of the light source is not essential, but the projected image must be focused. Ferroelectric LCD, a reflective liquid crystal on silicon technology, offers switching rates up to 10kHz, while other LCD technologies (e.g., nematic LC) are considerably slower. Transmissive LCDs are usually associated with significant light leakage in the "off" state while reflective LCDs have a low level of leakage primarily due to reflection off the front surface of the display engine[Underwood00, Underwood97].

MEMs devices have been developed that can rapidly scan a beam across a surface. These devices can be used as projective displays if the beam is scanned fast enough and its intensity is simultaneously modulated. Because the whole image is not simultaneously displayed, long exposures relative to the time needed to sweep a single scanline are needed. This property limits flexibility about exposure times for high speed acquisition for extremely high-speed devices and limits usefulness of lower speed devices This property limits flexibility about exposure times for high speed acquisition for extremely high-speed devices and limits usefulness of lower speed devices.

MEMs array devices such as a digital micro-mirror display (DMD) are an attractive alternative because of their high optical efficiency and extremely fast switching. Use of these devices can be problematic for a miniaturized application because they require a collimated light source. The optics needed to collimate an uncollimated light source may be too large for use in a system like this one. The individual micromirrors in a DMD array shift from $+10^\circ$ to -10° depending on the state of the associated

pixel. While the reflective surface of the micromirror has over 80% optical efficiency, almost no light is reflected in the direction associated with the "on" state when the pixel is switched to the "off" state.

Of these alternatives, reflective LCD and DMD engines are the best choices. Both offer switching speeds of at least 5kHz and can be integrated with memory and driving electronics into a single silicon chip. DMD engines have better optical efficiency but require a collimated light source. Reflective LCDs have much lower optical efficiency and cannot safely be left in the same state for too long, but have the advantages both of outputing a polarized light source and of not requiring collimation.

9.4.2.2 Light sources

There are two primary choices in the selection of types of light sources for this depth extraction system. The first is whether to use a broad spectrum light source delivering white light, or to instead use a narrow spectrum, possibly single wavelength, source. The second choice is whether to use continuous illumination or to use an intermittent source like a strobe light.

Broad spectrum illumination is attractive because only a single light source might be required for the system to provide structured light and full color video for textures. From the perspective of optical design broad spectrum illumination can be more difficult because of the need to avoid chromatic separation. For structured light, different wavelengths of light have different reflectance and absorbence properties and indiscriminatantly using all wavelengths, as is the case when a conventional monchrome camera is used, captures wavelengths with the best characteristics and the worst characteristics. If high intensity lighting is needed, potentially damaging levels of infrared and ultraviolet light may be delivered to illuminated tissues.

Narrow spectrum illumination has the benefit that a single wavelength or range of wavelengths can be selected with ideal reflectance and absorbence properties in tissue. Projector optics can be simplified without risk of color separation. A second, broad spectrum, unstructured light source is still needed to acquire full-color imagery. When the second light source is introduced it must either be on intermittently or there must be appropriate filters on the source and the structure detecting camera to avoid losing the structure in the broad spectrum light.

At the speeds used for current prototypes a strobe light, an intermittent light source, has had several advantages over continuous illumination. While total light output is lower than a continuous source, illumination is only supplied when an ap-

propriate structure pattern is being displayed. The illumination supplied during that time, less than 1 millisecond, is much greater than a comparable wattage continuous source could provide. A camera synchronized with the projector to only have the shutter open while a particular structure pattern is being displayed therefore sees more light with the intermittent source. As acquisition speeds increase this advantage is lost at the point when equivalant amounts of illumination is provided for the time when the shutter is open. At low acquisition speeds synchronized illumination allows the camera to be unsynchronized and makes the structured patterns visible to the eye, thereby simplifying debugging. The power supplies for many xenon strobe lights are quite loud and may present a fire hazard in an operating room with flamable anesthetic gases. Strobed light sources can be uncomfortable to directly observe.

In the current prototypes, not enough light was transmitted through the system to reliably find the structured patterns. At anticipated projection speeds structure patterns are hidden to both the eye and an unsynchronized longer exposure camera[2].

9.4.2.3 Proposed design

The best results are anticipated in a device making use of a digital micro-mirror display engine, a single-wavelength continuous light source, and a filtered continuous broad-spectrum light source. A schematic view of this prototype is shown in Figure 9.7.

A laser diode light source is a relatively inexpensive single wavelength source. These devices are small, lightweight, efficient, and can easily produce collimated light. Having a monochromatic collimated light source allows the use of a digital micro-mirror display without large optics for collimation. Digital micro-mirrors are also ideal for use with a laser because these displays are the smallest for equivalent pixel counts. This means that the beam does not have to be expanded very much nor recondensed very much for projection. Because the laser beam may be readily condensed into an extremely narrow beam, smaller optical apertures may be used for the projected light, leaving room for both a larger optical path for the cameras and for fiber-optics for broad spectrum illumination. Care must be taken to avoid interference patterns that occur when light is projected through apertures near its wavelength.

[2]While hidden to the eye, the patterns remain detectable under some conditions and would not be acceptable as an illumination source for long term work under direct vision. Likewise, a camera should be synchronized to ensure that it never exposes for periods that might make the structure apparent.

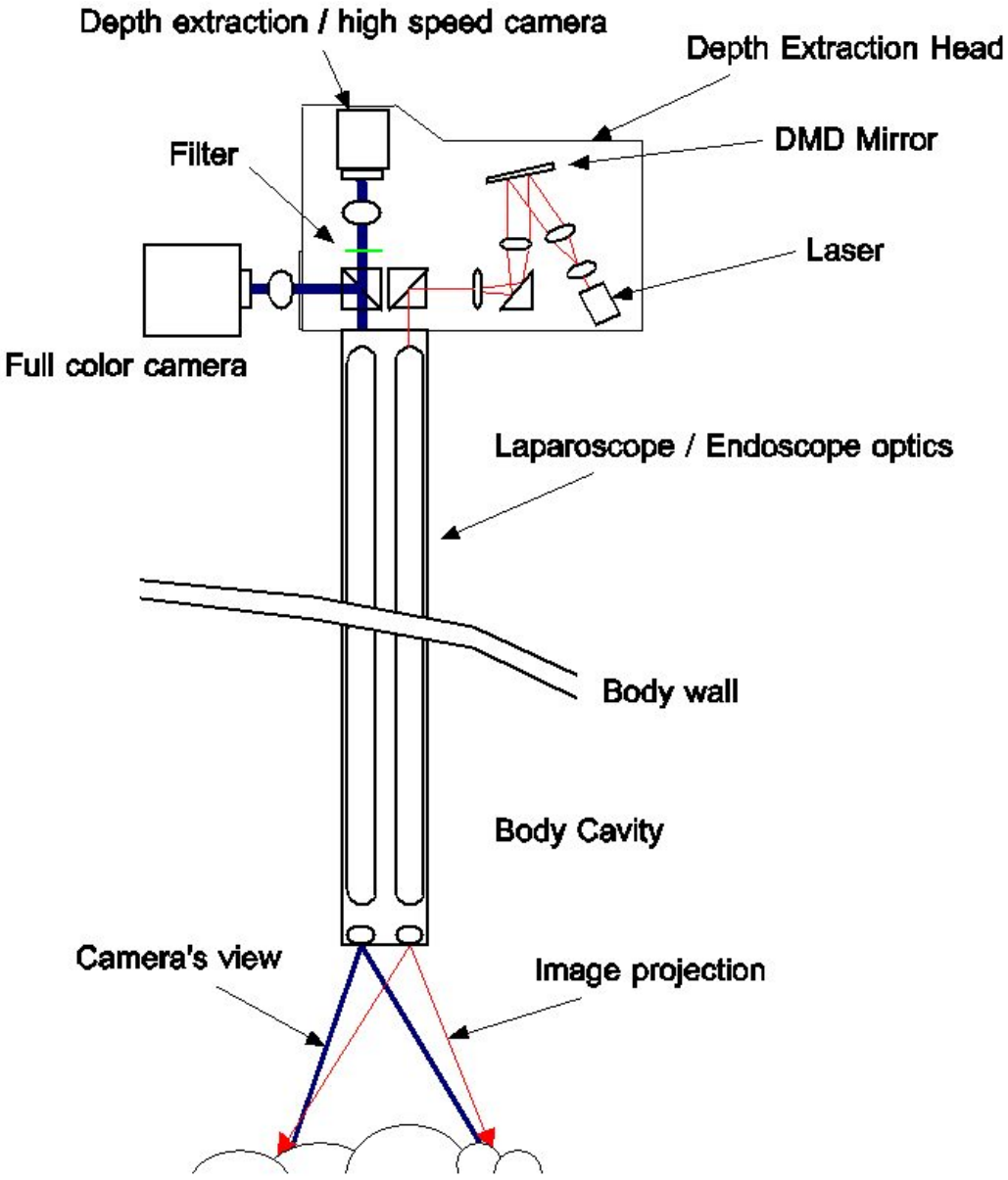


Figure 9.7: Schematic plan for a new prototype. The broad spectrum illumination source was not included in this cartoon for clarity.

A desktop model using a laser pointer (3mW, 650nm) and printed films for modulation was built as a proof-of-concept for a miniaturized laser projector. This source produces patterns that are visible in a lit room and are visible to a camera using relatively short exposures. A source with higher output, 10-20mW, would be better and would not require additional size. A green source (500-570nm) would have the best tissue interaction characteristics.

A continuous broad-spectrum source should be used to generate color imagery. This light should not be structured so unstructured fiber filaments could be used to deliver this light to the instrument's tip as is done in conventional laparoscopes. A very narrow filter should be used to remove light in frequencies very close to that of the laser, but wide enough to account for temperature sensitive shifts in wavelength. A dichroic mirror could be used to split the wavelength of the structure pattern from the other wavelengths for the structure detection camera and the color camera respectively.

It is likely that this design will be larger than the current prototype. Given a DMD 17mm wide and assuming a light incident perpendicularly to the DMD's surface, nearly a 100mm distance is needed from the light source aperture to the exit aperture in order to provide adequate separation between "on" and "off" reflected light.

9.5 Laparoscope design

A design for the laparoscope, a single integrated optical device rather than the pair of laparoscopes needed in the prototypes, similar in many respects to that used in stereo laparoscopes should be used. Unlike stereo laparoscopes, having two nearly identical optical paths may not be needed depending on the structure illumination source. Also, the side-by-side configuration of the lenses used for stereo laparoscopes is not needed.

Given the advantages of an angled laparoscope, a design with an angled view that could be used as a conventional angled scope, should be used rather than one with a 0°-angle design. In order to allow easy transition between use as a conventional or depth-extracting scope, the optical opening for the camera should be placed at the laparoscope's tip, while the aperture for the projector should be placed slightly up the shaft of the laparoscope.

Both the field of view of the cameras and of the projector will primarily be set by

lenses at the exit aperture of the laparoscope. The field of view of the camera should be selected to match comparable conventional laparoscopes. The field of view for the projector has several design considerations. If the field of view of the projector is too narrow, it will be unable to completely illuminate the camera's field of view for areas within the anticipated working volume. This reduces the area where effective depth extraction can occur. Choosing too large a field of view for the projector increases the power needed to provide adequate illumination for the entire field of view. This may also increase the number of patterns that need to be projected in order to gain adequate spatial resolution.

The choice of fields of view combined with the distance between the center of projection of the camera and projector have important ramifications for the accuracy and spatial resolution of the system. Use of numerical simulation, such as the simulation of depth extraction methods described in Appendix B, will be needed to make these design decisions.

The choice of separation distance between the apertures for the camera and projector a have a number of design considerations. If the apertures are too close together very low accuracy depth extraction will be possible for distant objects. If the separation is too large, shadows, areas seen by the camera but not illuminated by the projector are more likely to occur. In a worst-case scenario a projecor placed on the shaft of the scope could be placed against a tissue surface while leaving the camera a completely unobscured view of an area. A separation of 5mm appears to be a good choice for a working range of 15-150mm (error[3] of ± 0.04 to ± 2.0mm, increasing with distance). The projector fully illuminates the camera's field of view in this range.

[3]Error based on simulation results assuming 90° field of view camera with 600 pixels per line and 120° field of view projector viewing a flat surface using 128 stripes. A description of this simulation is given in Appendix B.3

Chapter 10

Image processing methods

"As the light conductors guide the entering light to the intended target, it is the purpose of the reflection conductors to return the reflected light to the eye"

– Philip Bozzini, 1806

A series of protoype systems were built to use structured light for laparoscopic acquisition. The hardware implementation of these devices is discussed in greater detail in Section 9.3. The results of these tests with respect to the image processing methods are discussed here. Later in this chapter a novel image processing method to rapidly and efficiently find stripes is described along with results from its implementation.

10.1 Initial animal cadaver studies

Two pig cadaver studies using two camera/projector configurations were performed. The hardware used for these studies is discussed in Section 9.3. The first of these studies was intended to test standard depth extraction techniques, using miniaturized hardware, on biological tissues. The second was aimed at collecting data to test the effectiveness of proposed algorithms.

As noted in the discussion of hardware, the result of the first study was that projected structure patterns could be recovered from biological tissues. While this result was encouraging the structure patterns had extremely low constrast and varied greatly in intensity between different tissue types. Specular highlights, both due to the projector as the light source and due to other sources in the room were much brighter than the projected patterns. Extraneous light sources, including light leaking out of the projector through the cooling vents, had to be systematically eliminated before structure patterns could be found.

The goal of the second study was to collect test data for new image processing algorithms. Using a large, high-constrast (1:300) video projector to collect this

data was based on the optimistic assumption that a miniaturized projector might be able to produce similar light output. With this system specular reflections from extraneous sources were not a concern. Because laparoscopic surgery is performed with the abdomen closed, and therefore not illuminated by light sources other than the laparoscope, this appears to have been a reasonably accurate dataset to use. Because locating structure patterns on uniformly colored Lambertian surfaces, like biological tissues coated with talc, is a simpler problem, images of patterns projected on both coated and uncoated tissues were used.

The ability of different algorithms to accurately detect the stripes from both the coated and uncoated tissue could then be compared. Because the goal of this experiment was to compare stripe segmentation, not calibrated acquisition, data was gathered, but an approximate model based on the ideal scenario (Equation 8.1) was used. The following equation was used for this reconstruction with coefficents selected only to visually match the scene but not to provide an accurate reconstruction:

$$D = \frac{1}{K_1 + K_2 s - x} \tag{10.1}$$

where K_i are arbitrary coefficents, s is the stripe number (numbered sequentially), and x is the found position of the stripe. Compared with simple thresholding, thresholding after scaling to full and no-illumination, and a simple edge finding method, only the method described here reliably provided similar results on fresh tissue in both the talcum powder coated and uncoated states (see Figure 10.1).

For all methods spatial aliasing of high frequency structure patterns was a problem, and other segmentation failures were more frequency in patterns with higher spatial resolution. This problem was worse on the uncoated tissue, most likely due to the smaller amount of light scattered back to the camera and the defocusing effect of the tissue surface.

10.2 Image processing

The following images are needed as source data for this algorithm:

m images, P, with projected stripe patterns

m images, N, with the inverse stripe patterns projected

m is the number of bits of striping (for binary encoded structured light)

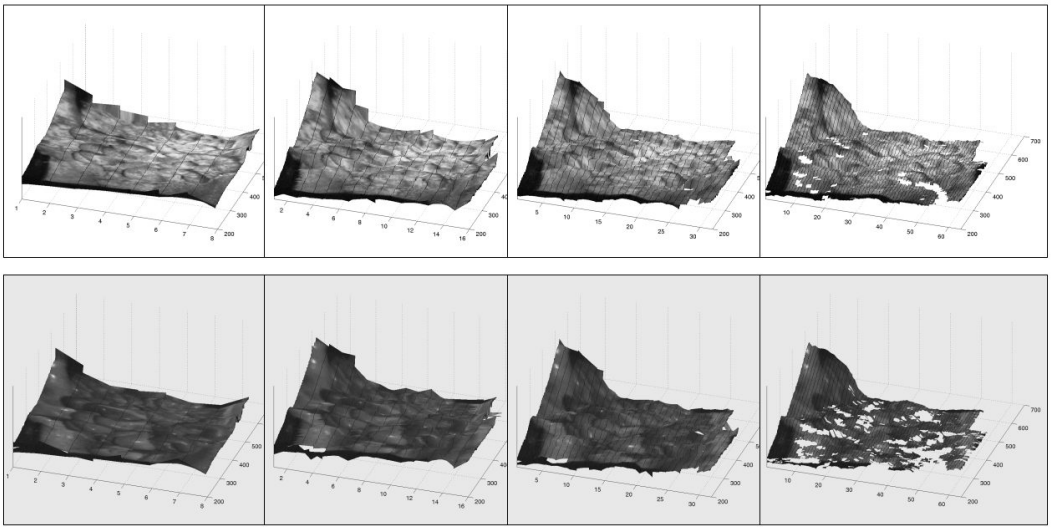


Figure 10.1: Rendered texture-map range images based on recovered structured light and an approximate (uncalibrated) model for depth. Top: raw tissue. Bottom: Talc coated tissue. Left to right, lower to higher number of stripes. Areas where patterns could not be identified are not rendered. Note similar shape in all renderings and surface detail visible with greater numbers of stripes. More holes in the surface are present when there are more stripes.

10.2.1 Non-optimized

Subtract the nth negative image, N_n, from the nth positive image, P_n

$$C_n = P_n - N_n \tag{10.2}$$

Label each pixel that belong to an 'off' stripe as 0, as 1 for an 'on' stripe, and leave pixels that cannot be identified unmarked. The 'on' stripes can be identified because their intensity value in the difference image C_n will be greater than a threshold level r, while 'off' pixels will have an intensity level less than $-r$. Pixels with intensity values between r and $-r$ cannot be identified (as in the case of shadows, specular reflections, or low signal-to-noise ratio).

$$L_n = \begin{cases} 0 & if \quad C_n \leq 2^{b-1} - r \\ 1 & if \quad C_n \geq 2^{b-1} + r \\ undefined & if \quad 2^{b-1} - r < C_n < 2^{b-1} + r \end{cases} \tag{10.3}$$

Record which pixels cannot be identified in this image pair.

$$U_n = \begin{cases} 0 & if \quad 2^{b-1} - r < C_n < 2^{b-1} + r \\ 1 & otherwise \end{cases} \tag{10.4}$$

In the case of binary encoding of the stripes, an image with each pixel labeled with the stripe encoded, S can be created by

$$S = (1 + \sum_{i=1}^{m} 2^i L_i) \prod_{i=1}^{m} U_i \tag{10.5}$$

The multiplication by the product of all images U_n insures that any pixel in which the stripe pattern could not be identified in any image pair is not labeled as belonging to a stripe. Thus the image S contains stripes labeled from 1 to $2^m + 1$ with zero value pixels indicating that a stripe could not be identified for that pixel in one or more image pairs.

10.2.2 Optimized

Equation 10.2 is appropriately normalized so the result fits into the image data type (b bits). This will most likely result in the loss of the low order bit.

$$C_n = \frac{P_n}{2} + 2^{b-1} - \frac{N_n}{2} \tag{10.6}$$

The next step is to apply thresholds to identify and mark the 'on' and 'off' pixels. First threshold to find the 'on' pixels by setting bit $n-1$ if the intensity of that pixel is greater than 128 plus a preset level r. Next threshold 'off' pixels by setting bit $n-1$ to zero for pixels with intensity greater than 128 minus a preset level r. Pixels that are not part of these two ranges are labeled non-identifiable by setting the mth bit.

$$L_n = \begin{cases} 0 & if \quad C_n \leq 2^{b-1} - r \\ 2^{n-1} & if \quad C_n \geq 2^{b-1} + r \\ 2^m & if \quad 2^{b-1} - r < C_n < 2^{b-1} + r \end{cases} \tag{10.7}$$

In this way, Equation 10.7 combines Equations 10.3 and 10.4 from the non-optimized version into a single operation. Likewise two images, U_n and L_n in the non-optimized method can be combined into a single L_n.

The set of images L may be combined with a bitwise OR as they are generated or as a separate step creating a stripe encoded image, S.

	Convolution	$n = 7$	$n = 5$	$n = 3$
Camera capture off	ON	100ms	73ms	47ms
	OFF	88ms	64ms	41ms
Camera capture on	ON	2203ms	1602ms	967ms
	OFF	2201ms	1568ms	966ms

Table 10.1: Performance on a PII/400Mhz with a Matrox Genesis (1998) board for capture and processing. Performance is based on the mean time of 20 uninterrupted runs. The value n refers to the number of bits of striping used. For each run $2n$ images were processed. Runs performed with camera capture turned off were completed by pre-loading sample images into buffers on the Matrox Genesis card and copying them to new buffers for processing when needed.

$$S = \bigcup_{i=1}^{m} L_i \tag{10.8}$$

The resulting image S from the optimized method differs from that obtained in the non-optimized method in that non-identified pixels have values greater than 2^m rather than having a value of 0.

The stripe labeled image S may be converted to a range image using normal techniques used in structured light as discussed in Section 8.3. Pre-calibrated look-up tables and regression fitted functions may be useful methods to rapidly progress from segmented images to renderable surfaces.

10.3 Performance

Table 10.1 shows the time to complete the image processing phase in the prototype system. This implementation is capable of finding stripes for 7-bits of stripe patterns at a rate of ten times per second. Significant optimizations still remain that result in much faster operation. The current implementation uses the Matrox Imaging Library (MIL) as a high-level programming interface with the Matrox Genesis digital signal processing board. Using MIL greatly simplifies programming the digital signal processors, but adds overhead and loss of fine level tuning. Additionally, image processing commands through MIL are issued to the Genesis board via the system bus while low level programming of the processors allow the processors to run autonomously from the rest of the PC. Other processes on the PC cause delays up to 50 milliseconds.

When camera capture is enabled, the process is dramatically slowed. A large portion of the time is used for synchronization. Each time a camera capture is needed, the current protype waits at least one and a half frames of video. This is done to ensure that the updated stripe pattern is actually being projected by the projector and to make sure that the camera is at the beginning of the frame when capture starts. Some of these delays can be reduced by triggering the camera only when a capture is needed. In the current implementation, no image processing is performed while waiting for the next images to be captured. A lower level implementation of the system should have the processors on the capture and processing board operate while the next image is being captured.

10.4 Results from prototype

Two experiments were performed on animal tissues to test these methods. The first used a porcine cadaver and an uncalibrated system with off-line processing. The second experiment was performed on chicken organs with on-line segmentation and off-line rendering.

10.4.1 Prototype calibration

Further simplification of Equation 8.4, is useful to get a reasonable calibration for the prototype system. If the projected patterns and captured images are nearly rectified, then we assume that a scan line in the camera images implies only a single scan line from the projector. Ideally the y terms could all be dropped, but instead the y value from the camera image is maintained as in the following equation:

$$C_2 y_1 + C_3 y_2 = B y_2 + k \tag{10.9}$$

This is the assumption that a y position in the first camera implies (linearly) a y position in the second camera. This implies that the original center of projection is very similar (except for scaling) to that used for reprojection (mostly to account for differences in camera characteristics) that might be needed.[1]. Under these assump-

[1]Equation 10.9 could be written as $C_2 y_1 + C_3 y_2 + x_2 = B_1 y_2 + B_2 x_2 + k$ as a very rough accomodation for greater variation from the ideal model. While this approach is more correct if one were to measure what the coefficents should be based on camera parameters, the result, Equation 10.11, is the same for achieving a good fit with a regression model using either model.

tions

$$D = \frac{K}{C_1 x_1 + x_2 + B y_2 + C_4} \tag{10.10}$$

which is equivalent to

$$\frac{1}{D} = c_1 x_1 + c_2 x_2 + c_3 y_2 + c_4 \tag{10.11}$$

where c_n are appropriate coefficents to fit the model.

A rough calibration of the prototype can then be achieved by linear regression with $1/D$ as the independent variables, the stripe number, the stripe's position, the scan line in the camera image, and a constant. A linear model also lends itself to rapid calculation on processed images.

For the results shown in the remainder of the chapter the scheme that follows was used for calibration. A series of data were collected from planes placed at known distances, every 0.5cm between 5.0cm and 15.0cm, from the camera and projector. The stripe patterns in several images, those acquired every 1cm between 7.0cm and 15.0cm, were segmented and the centers of each stripe on each scan line was found. The approximation given in Equation 10.11 was applied by performing a linear regression of one over the distance ($1/D$) to the plane with the stripe number, the center of the stripe, the scanline number, and a constant. The coefficents found using this regression could then be used to recover depth from the location of found stripes in images of arbitrary objects.

In order to verify the calibration, planes imaged at depths not used in the regression were converted to range images so the resulting depth information could be compared to the known range. Figures 10.2 and 10.3 (top) demonstrate that a good result was achieved near the center of the range of depths regressed. However, as the depth varies from near 11.0cm, significant distortions become increasingly apparent (Figures 10.2 and 10.3, middle and bottom). While many factors may be responsible for these distortions, these distortions are consistent with the effects of radial distortion in the camera and projector which was not corrected for in these datasets.

For these data corrected x and y positions were also not recovered. A relationship similar to the one given in Equation 10.11 can be developed to implicitly find the relationship between the found position of the feature and the x and y components of its three-dimensional coordinates. The calibration grid in the images was not well designed for automated location of know coordinates so this was not attempted.

This calibration, while far from perfect, is aedequate for demonstrating the depth

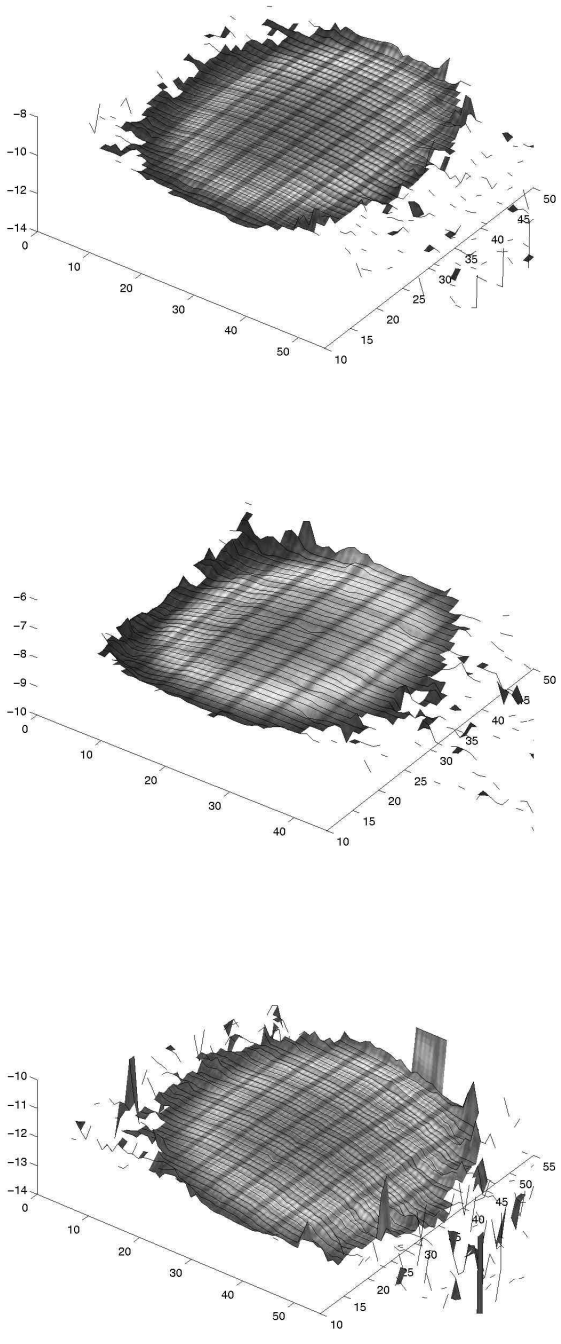


Figure 10.2: Top: a texture-mapped rendering of a plane 10.5cm from the camera which appears nearly planar and at 10.5cm. Middle: rendering of a reconstructed plane 8.5cm and Bottom: 12.5cm from the camera showing significant distortion.

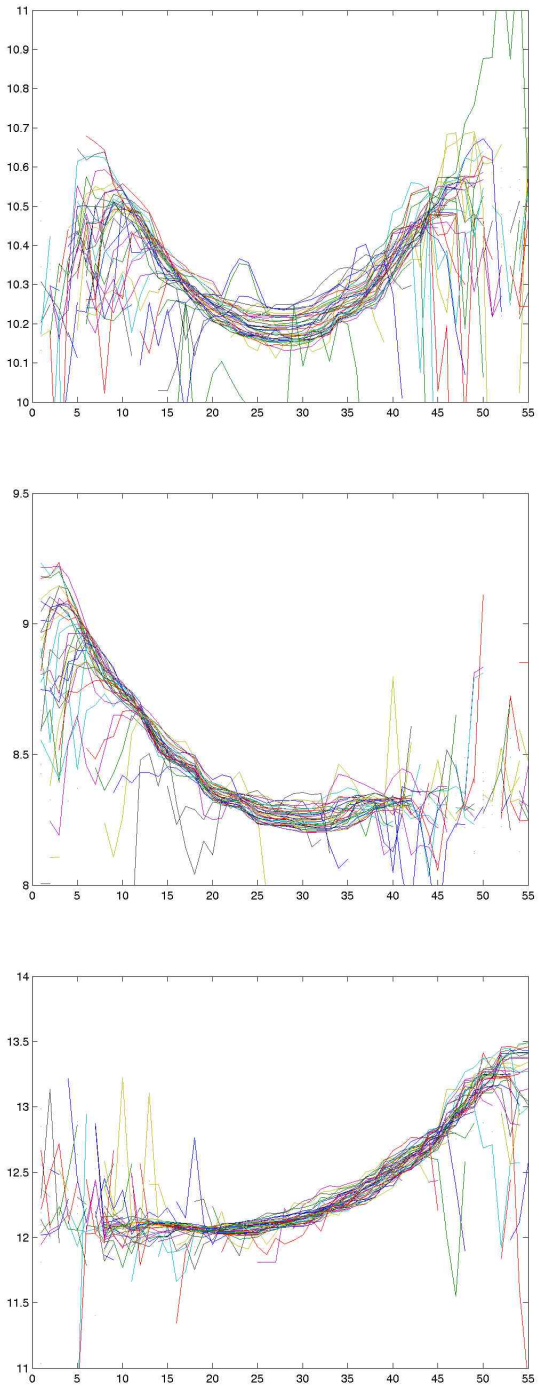


Figure 10.3: Top: plots of depth at scan lines through a reconstructed plane 10.5cm from the camera. Middle: plots of depth across a reconstruction of a plane at 8.5cm and Bottom: 12.5cm from the camera showing significant distortion. In all of these plots a signifcant degree of noise is evident.

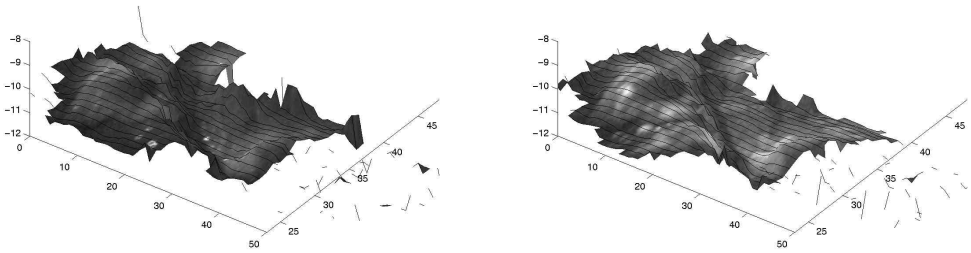


Figure 10.4: Texture-mapped rendering of the reconstructed surface of raw chicken organs. Left: uncoated surface, Right: talc coated surface

can be recovered using the prototype apparatus.

10.4.2 Depth aquisition

The apparatus and algorithms were tested on fresh animal tissue. Chicken necks and livers were used to provide some variety in albido. These pieces were arrayed in front of the device, imaged, coated with talcum powder, and imaged again. Tissues were coated with talc by gently blowing talc onto them. While every effort was made to create a fine, uniform coating, variation in the coating was unavoidable. Following imaging, the tissue was removed, carefully rinsed to remove all of the talc, and rearranged for another capture sequence.

The software saved all recorded images as well as the segmented stripe images in uncompressed TIFF format. Subsequent processing into depth and rendering was completed with several MatLab programs.

The results of two sets of reconstructions and renderings are shown in Figures 10.4 and 10.5. The reconstructions from raw and talc-coated tissue are shown side-by-side for comparision. The reconstructed areas have irregular borders due to distortion at the edge of the projector's field-of-view and shadowing effects.

In Figure 10.4, 947 range points were identified in both the raw and talc data sets. The mean difference between recovered depth values was 0.48mm with a standard deviation of 3.16mm. The mean absolute difference was 1.52mm. In this dataset less depth information was recovered from the talc coated tissue than from the uncoated (964 versus 1023). In Figure 10.5, 1028 points were found in common (1158 in the raw set and 1165 in the coated) with a mean difference of 0.09mm. The standard

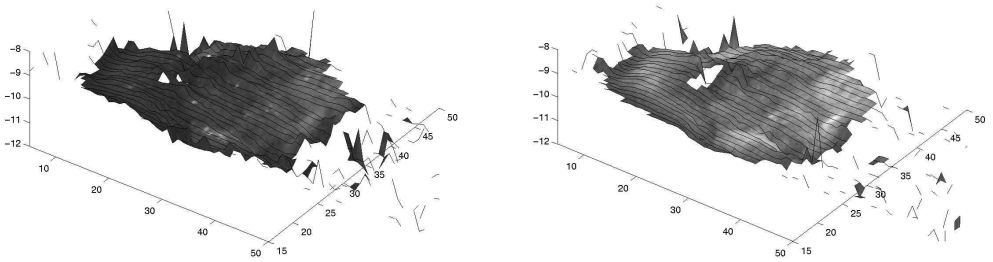


Figure 10.5: Texture-mapped rendering of the reconstructed surface of raw chicken organs. Left: uncoated surface, Right: talc coated surface

deviation of the difference was 2.3mm and the mean absolute error was 1.39mm in this dataset.

10.4.3 Discussion

Despite median filtering and combining data from multiple scanlines, significant noise is evident in the reconstruction. Strategies like combining scanlines earlier in the reconstruction process (before conversion to depth rather than after) may help, but further reduction of noise by improving the quality of acquired images is a better solution. The segmentation threshold used detected stripes by an intensity diference of only 3 ($r = 1$ in Equation 10.7) intensity levels (of 255) in the images. The somewhat contradictory finding that more stripes were found in the uncoated case for Figure 10.4 is perhaps explained by the increased albido causing an increased sensitivity to ambient lighting when the tissue was coated that was not overcome by the low contrast ratio of the projector.

Earlier data collected with this apparatus was abandoned because the projected stripe patterns were only slightly above the noise level in the camera. In any given pair of images, stripe patterns were correctly identified for approximately 80% of the pixels. Because the pixels correctly identified were randomly distributed, as the pairs of images were consolidated to be able to identify stripe patterns, the number of correctly recognized stripes dropped off dramatically.

Thus, the alogrithms presented here appear quite promising in that they perform well despite a lower contrast ratio projector and very low image intensity. These algorithms now operate at near real-time (10Hz) with high-level code on aging hard-

ware, so further improvements in performance should be expected. The calibration methods used here are easily applied to create nice renderings and may be applied more accurately if the underlying assumptions are more closely observed.

Chapter 11

Future work

"Ulcers . . . fistulas, tumors, . . . all these can be visualized easily and distinctly without causing any injury, detriment, or discomfort."

– Philip Bozzini, 1806

11.1 Depth extraction

The results shown here are a good demonstration that the described algorithms and proposed hardware are likely to be successful. The current system is limited in its performance by both hardware and software issues.

A new prototype built with components and properties described in Chapter 9 is needed. This system should include a new miniaturized projector, a high speed camera, and a custom laparoscope.

The software, currently written using a high-level programming language, will need to be coded directly, in low level code, on the image processing board. This low level coding should allow maximum efficiency in processing to enable high speed capture, processing, and output.

After this prototype has been extensively tested on the bench-top, it should be tested under progressively more realistic conditions. As with past prototypes, testing using animal cadavers is a good first step. A next step would be to test the devices capture and processing in a live animal without using processed data to guide any interventions.

At some point this device would be ready for testing in the operating room with actual patients. At first, it should be used as a three-dimensional recording device and, as approvals are granted, a first three-dimensional application should be considered.

A system that produces output like a stereo laparoscope, a relatively trivial rendering of the three-dimensional data from the view points of the right and left eye,

is a promising first choice. This system, unlike conventional stereo-laparoscopes, create left and right eye views that are corrected for the user's intra-pupillary distance (IPD) and distance from the monitor. This feature might reduce the eye strain, and associated headaches, dizziness, and nausa, sometimes reported in users of stereo laparoscopes. More importantly, this system, with a mechanism for head tracking, could provide head-motion depth cues. Research using alternative techniques for coupling a monoscopic laparoscope's view to the user's head have shown an associated doubling in accuracy of performance of a spatial exploration task [Voorhorst97, Voorhorst99]. Use in augmented reality is a long term goal with many additional obstacles that are discussed in 11.4.

Other application areas including industrial inspection applications should be considered. A small system for real-time acquisition of three-dimensional shape could be important for many applications. Inspection of the interior of pipes in chemical plants or quality control of internal components in automobile manufacturing are two of many possible applications outside of medicine.

11.2 Simulated augmented reality

Many of the results from the user study are not statistically significant due to the small sample size. For the results with a relatively high level of significance further investigation is warranted. In particular, additional work should be done on the possible benefits of the simulated augmented reality environment on learning to perform tasks with full laparoscopic visualization.

This simulated augmented reality environment was designed to simulate a augmented reality environment for laparoscopic surgery that only provides a different visualization. One of the anticipated benefits of have augmented reality systems is the possibility of intuitively merging preintraoperative and intraoperative imagery. A simulated augmented reality system could conceivably be constructed using techniques similar to those used in Stetten's Sonic Flashlight[Stetten00]. Other possibilities to achieve data fusion in the simulated augmented reality environment should be explored in order to assess the benefits of these features before a fully realized augmented reality system to assist laparoscopic surgery is built.

Other tasks may have additional benefits in the simulated augmented reality environment. The task used here, while difficult and complicated to complete, makes no use of complicated geometry. Tasks involving more complex geometry, more analo-

gous to anatomy encountered surgically, or more complex instrumentation may show larger benefits both in terms of improved performance in simulated augmented reality and in terms of subsequent performance under fully laparoscopic conditions. Navigation of an angled laparoscope, a considerably more difficult task than navigation of a 0°-angle scope, is one example[Eyal01].

11.3 Assessment task

The task used for this study has a number of aspects that could be improved further. With a single task requiring up to 20 minutes, this task is quite time consuming making it impractical to collect a large number of samples from a single subject. The consequence of this is that a study to truly examine learning, by measuring changes in task performance until a plateau in performance is reached, would be very difficult because of the time needed by each subject to complete the activity.

Measurement of the circle radius at 200 points was intended to provide a mechanism for increasing the sample size. Analysis using the 200 radius measurements separately is difficult because the order they were cut is not known. Finally, some extra variation in the final outcome measure is introduced in the steps requiring human intervention.

Except for these issues, the experimental task functioned as intended. The following modifications could be made to correct these issues:

1. Measure accuracy of cut along a straight line

2. Direct subjects to cut from a starting point to a finishing point so data points can be ordereexperimentalem Mark paper with a parallel line. Straight lines can be accurately and reproducably found by image processing methods so a fully automated analysis could be used

3. The marked line should be a thin black line on white paper so that subjects will not be as likely to be biased towards errors in one direction

This task should be take much less time to complete while retaining most of the difficulties associated with the original task. Having a shorter task will allow future studies to gather the additional samples needed to accurately assess learning without requiring additional time from subjects.

11.4 Augmented reality for laparoscopic surgery

There are a large number of problems that remain for creating an augmented reality system to assist laparoscopic surgery. The technical problems include designing system architecure to support it, design of usable displays, and design of useful visualizations, particularly for data fusion. On the medical side, it is not totally clear how the technology should best be used, and changes in workflow may be needed to integrate augmented reality systems in the operating room.

Current stereoscopic medical augmented reality systems use graphics workstations designed for scientific visualization. These computers, whose performance is now superseded in graphics performance by personal computers (PCs), offer video input and output capabilities, specifically multiple channel video input directly into the graphics pipeline, not currently available in PCs. PC components that, if used together, might match the features of the workstations frequently are not compatible. Creative use of multiple PCs and high-speed networking may give one possible solution.

In some of the local experiments with mock laparoscopic visualization in an augmented reality environment several visualization questions have been raised. When multiple data sources are available, how should they best be optimally displayed?. Figure 11.1 demonstrates how quickly such a visualization becomes confusing. The field of scientific visualization has a growing body of knowledge on how to approach displaying multiple datasets. There are many possible approaches.

The design of useful displays for laparoscopic surgery is a very significant problem. Current displays offer relatively low angular resolution given the size of features a surgeon might want to see. Software methods to adapt displays to the demands of the surgical environment (such as in Appendix A) can provide usable systems for short-term use despite some serious deficiencies in the displays. Most research has only had subjects using augmented reality systems for short time periods. The long term consequences of using various display devices in augmented reality, both for repeated use and for continuous use, are not well known. Demonstrated efficacy of the augmented reality systems may allow less than optimal displays to be adopted, but significant research and development will be needed to develop better displays. Recent developments in microdisplay technologies give reason to believe that these issues as well as size and weight can be addressed.

If an augmented reality system to assist laparoscopic surgery were available, it is not clear how it should be deployed for maximum benefit. The results given in

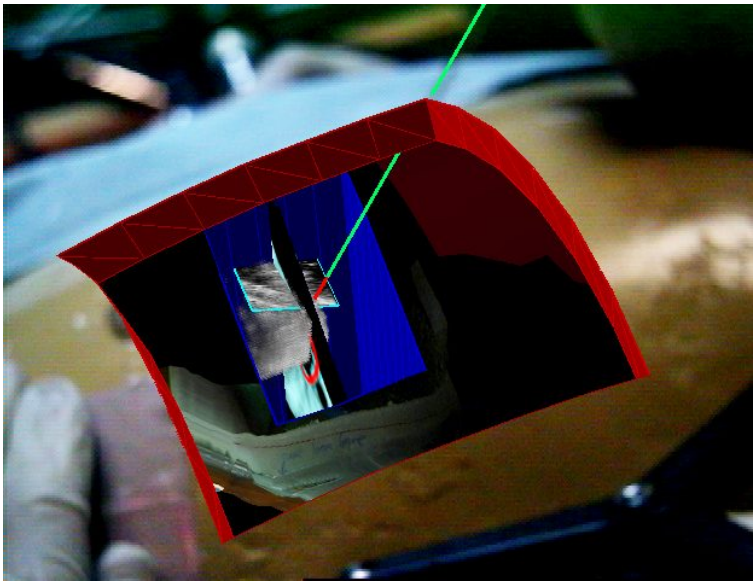


Figure 11.1: An attempt to show a single visualization combining real-time (simulated) laparoscopically acquired surfaces, real-time live ultrasound, tracked instruments, and a preoperative CT scan.

Chapter 7 demonstrate a benefit for untrained personnel and suggest a disadvantage for trained surgeons. The performance of trained surgeons might be improved if magnification of the operative site were available to them in the augmented reality view. Perhaps of greater importance to the future of augmented reality systems is the suggestion in the data that use of an augmented reality visualization confers a learning advantage both among untrained personnel and in students. Perhaps surgeons learning or observing a novel laparoscopic procedure using an augmented reality visualization would get a similar boost in their rate of learning the procedure.

Several researchers are exploring the effectiveness of teaching medical students and residents anatomy from a laparoscopic perspective[Fitzpatrick01, Park01, Cundiff01]. The hope is that by combining conventional anatomy training by open cadaveric dissection with laparoscopic dissection, students and residents will more quickly learn to understand laparoscopic procedures. A similar benefit could perhaps be realized by observing laparoscopic procedures using an augmented reality environment. The augmented reality environment more closely resembles open surgery and open dissection and would perhaps help as a transition to fully laparoscopic work.

In the event that augmented reality systems are sufficently useful to warrant their regular use in the operating room, significant changes in the positioning of the patient and the surgeon may be needed. As noted in Section 7.5 many of the study

participants who used the simulated augmented reality environment commented that the positioning of the view window, the task objects, and the instruments was less than ideal. New operating room layouts for performing these procedures, including modified instrument handles may be needed in order to provide better ergonomics for use of augmented reality systems. Recent interest in the ergonomics of laparoscopic surgery has shown that current technique and instruments are not very good and future studies may indicate significant changes in the conventional laparoscopic operating room[Emam02, Berquer02, Berguer01, van Veelen01].

Chapter 12

Summary and Conclusion

"If the light transmitter is constructed according to these specifications and the acoompanying drawings, it will work as described"

– Philip Bozzini, 1806

In this dissertation I examine the usefulness and buildability of an augmented reality visualization system for laparoscopic surgery.

The basic principles of laparoscopic surgery have been known for nearly 100 years. The development of laparoscopic instrumentation and technique goes back approximately 200 years. Development of these instrumentes and techniques has been an interdisciplinary process across medical and research disciplines. The interdisciplinary nature of this is evident in ongoing developments. The recent recent widespread acceptance of laparoscopy has occured because of the benefits to most patients, and despite the additional difficulty for surgeons.

Early results from computer graphics, going back to the early days of interactive and immersive graphics, demonstrated the potential of augmented reality. The ability of this technology to add virtual objects or to insert real, but usually invisible objects into a visualization, has many compelling applications. In medicine augmented reality is already in use to assist image-guided and stereotactic procedures. Laparoscopy is a compelling application area for augmented reality because it might be applied to restore the natural hand-eye coordination lost in laparoscopic technique. More importantly, augmented reality can provide an intuitive way for surgeons to visualize additional data including preoperative and intraoperative imagery in the context of a procedure underway.

Two major questions were addressed regarding this technology. The first is, if an augmented reality environment to assist laparoscopic surgery could be built, would it actually be useful? The second question is, can it be built?

Testing the usefulness of something that does not yet exist is tricky business. This question is addressed using *simulated augmented reality*, a simulation of an idealized augmented reality system created using physical mock-ups. Mechanisms for simulating known deficencies of actual augmented reality systems are described. Some of these methods were also employed to create a degraded simulated augmented reality environment to explore differences in effectiveness between the idealized and a perhaps more realistic model of what an augmented reality environment will be like.

A study of 40 medical students and eight surgeons using simulated augmented reality was conducted to compare performance of a surgically relevant task under laparoscopic and simulated augmented reality visualization. The simulated augmented reality enviroment provided a significant performance benefit for medical students but not for practicing surgeons. Performing the task using the simulated augmented reality environment before performing the task laparoscopically was beneficial to both medical students and to surgeons. This finding suggests that an important role for augmented reality systems may be for training residents to perform laparoscopic procedures and for assisting practicing surgeons to learn new procedures. The potential benefit of combining preoperative and intraoperative imagery using augmented reality was not addressed by this study. Degraded simulated augmented reality did not have a benefical effect on task performance, but may have had some benefits in training.

A more detailed exploration of the effects of having a less than ideal augmented reality environment should be performed. In particular, the hypothesis that greater benefits will be seen, even among trained surgeons, for tasks of greater complexity should be explored. If this is true, the usefulness of augmented reality for trained and experienced surgeons will be well established.

Augmented reality systems have a large array of technological requirements. The short version of this list is that a display, a source of three-dimensional data, and mechanisms to correctly register the data with the display are the primary requirements. Current augmented reality environments demonstrate that usable displays and registration techniques are feasible. However, for laparoscopic surgery, a real-time way to acquire surface geometry inside a patient – the needed source of three-dimensional data – has not been demonstrated.

A proof-of-concept real-time range acquiring laparoscope has been constructed. This system uses specialized hardware and algorithms for a near real-time implementation of structured light triangulation-based depth extraction. Results from preliminary studies on biological tissues show the effectiveness of the highly paral-

lelizable algorithms. The current prototype demonstrates that this system can be miniaturized. However, significant improvements are needed. The design of a system using a digital micromirror device (DMD), laser light source, specialized optics, and a high-speed camera is proposed to overcome the deficiencies of the current prototype.

These results demonstrate the potential usefulness and feasibility of augmented reality for laparoscopy. The long term potential of this technology is to allow surgeons an intuitive way to combine multiple data sources intraoperatively. This may give rise to new and improved surgeries in the future.

Appendix A

Dynamic Virtual Convergence for Video See-through Head-mounted Displays: Maintaining Maximum Stereo Overlap throughout a Close-range Work Space[1]

A.1 Introduction and motivation

Most commercially available head-mounted displays have been manufactured for virtual reality applications, or, increasingly, as personal movie viewing systems. Using these off-the-shelf displays is very appealing because of the relative ease with which they can be modified for video see-through use. However, depending on the intended application, the characteristics of the displays frequently are at odds with the requirements for an AR display.

In one of our applications, ultrasound-guided needle breast biopsy (Figure A.1, a physician stands at an operating table. The physician uses a scaled, tracked, patient-registered ultrasound image delivered through our AR system to select the optimal

[1]This appendix is taken with only slight modification from the paper Andrei State, Jeremy Ackerman, Gentaro Hirota, Joohi Lee, and Henry Fuchs. Dynamic virtual convergence for video see-through head-mounted displays: Maintaining maximum stereo overlap throughout a close-range work space. *Proceedings of the International Symposium on Augmented Reality (ISAR)*, pages 137–146, 2001.

approach to a tumor, insert the biopsy needle into the tumor, verify the needle's position, and capture a sample of the tumor. The physician wears a VST-HMD throughout the procedure. During a typical procedure the physician looks at an assistant a few meters away, medical supplies nearby, perhaps one meter away, the patient half a meter away or closer, and the collected specimen in a jar twenty centimeters from the eyes.

Most commercially available HMDs are designed to look straight ahead. However, as the object of interest (either real or virtual) is brought closer to the viewer's eyes, there is a decreasing region of stereo overlap dedicated to this object (on the nasal side). Since the image content being presented to each eye is very different, the user is presumably unable to get any depth cues from the stereo display in such situations. Users of our system have been observed to move either the object of interest or their head so that the object of interest becomes visible primarily in their dominant eye – from this configuration they can apparently resolve the stereo conflict by ignoring their non-dominant eye.

In typical implementations of video see-through displays, cameras and displays are preset at a fixed angle. Researchers have previously designed VST-HMDs while making assumptions about the normal working distance. In one design, discussed in the following section, the video cameras were preset to converge slightly in order to allow the wearer sufficient stereo overlap when viewing close objects. In another design, the convergence of cameras and displays could be selected before using the system to an angle most appropriate for the expected working distance. Converging the cameras, or both the cameras and the displays, is only practical if the user need not view distant objects as there is often not enough stereo overlap or too much disparity to fuse distant objects.

This appendix discusses an alternative to physically modifying convergence of either the cameras or the displays. Our technique does not require moving parts within the HMD and is implemented fully in software.

A.2 Related work

In the pioneering days of VST AR work, researchers used to improvise (successfully) by mounting a single lipstick camera onto a commercial VR HMD. Even then careful consideration was given to issues such as calibration between tracker and camera[Bajura92].

In 1995, our team assembled a stereo AR HMD [State96b]. The device consisted of a commercial VR-4 unit and a special plastic mount (attached to the VR-4 with Velcro!), which held 2 Panasonic lipstick cameras equipped with oversized C-mount lenses. The lenses had been chosen for their extremely low distortion characteristics, since their images were digitally composited with perfect perspective CG imagery. Two important flaws of the device emerged: (1) mismatch between the fields of view of camera (28° horizontal) and display (ca. 40° horizontal) and (2) eye-camera offset or parallax (see [Azuma92] for an explanation), which gave the wearer the impression of being taller and closer to the surroundings than she actually was. To facilitate close-up work, the cameras were not mounted parallel to each other, but at a fixed 4° convergence angle, which was calculated to also provide sufficient stereo overlap when looking at a collaborator across the room while wearing the device.

Today many video-see-through AR systems in labs around the world are built with stereo lipstick cameras mounted on top of typical VR (opaque) or optical see-through HMDs operated in opaque mode (for example, [Kanbara00]). Such designs will invariably suffer from the eye-camera offset problem mentioned above. (The design described in this paper is no exception, even though our new technique is not limited to such designs.)

[Fuchs98] describes a device that was designed and built from individual LCD display units and custom-designed optics. It had two identical "eye pods". Each pod consisted of an ultra-compact display unit and a lipstick camera. The camera's optical path was folded with mirrors, similar to a periscope, making the device "parallax-free" [Takagi00]. In addition, the fields of view of camera and display in each pod were matched. Hence, by carefully aligning the device on the wearer's head, one could achieve near perfect registration between the imagery seen in the display and the peripheral imagery visible to the naked eye around each of the compact pods. Thus this VST-HMD can be considered *orthoscopic* [Drascic96]. Since each pod could be moved separately, the device (characterized by small field of view and high angular resolution) could be adjusted to various degrees of convergence (for close-up work or room-sized tasks), albeit not dynamically but on a per-session basis. The reason for this was that moving the pods in any way required inter-ocular recalibration. (The "head tracker" was rigidly mounted on one of the pods so there was no need to recalibrate between head tracker and eye pods.) The movable pods also allowed exact matching of the wearer's inter-pupillary distance (IPD).

Other researchers have also attacked the parallax problem by building devices in

which mirrors or optical prisms bring the cameras "virtually" closer to the wearer's eyes. Such a design is described in detail in [Takagi00], together with a geometrical analysis of the stereoscopic distortion of space and thus deviation from ortho-stereoscopy that results when specific parameters in a design are mismatched. For example, there can be a mismatch between the convergence of the cameras and the display units (such as in the device from [State96b]), or a mismatch between inter-camera distance and user IPD.

While [Takagi00] advocates rigorous ortho-stereoscopy, other researchers have investigated how quickly users adapt to dynamic changes in stereo parameters. [Milgram92] investigated users' judgment errors when subjected to unannounced variations in inter-camera distance. The authors determined that users adapted surprisingly quickly to the distorted space when presented with additional visual cues (virtual or real) to aid with depth scaling. Consequently, they advocate dynamic changes of parameters such as inter-camera distance or convergence distance for specific applications.

[Ware98] describes experiments with dynamic changes in virtual camera separation within a fish tank VR system. They used a z-buffer sampling method to heuristically determine an appropriate inter-camera distance for each frame and a dampening technique to avoid abrupt changes. Their results indicate that users do not experience "large perceptual distortions," allowing them to conclude that such manipulations can be beneficial in certain VR systems.

Finally, [Matsunaga00] describes a teleoperation system using live stereoscopic imagery (displayed on a monitor to users wearing active polarizers) acquired by motion-controlled cameras. The results indicate that users' performance was significantly improved when the cameras dynamically converged onto the target object (peg to be inserted into a hole) compared to when the cameras' convergence was fixed onto a point in the center of the working area.

A.3 The dynamic virtual convergence system

The [Fuchs98] device described above had two eye pods that could be converged *physically*. As each pod was toed in for better stereo overlap at close range, the pod's video camera and display were "yawed" together (since they were co-located within the pod), guaranteeing continuous alignment between display and peripheral imagery. Our new technique deliberately violates that constraint but uses "no moving parts", since it is implemented fully in software. Hence there is no need for recalibration as

convergence is changed. It is important to note that sometimes VR or AR implementations mistakenly mismatch camera and display convergence, whereas our method intentionally decouples camera and display convergence in order to allow AR work in situations where an ortho-stereoscopic VST-HMD doesn't reach (because there are usually no display pixels close to the user's nose).

The implementation requires a VST HMD whose video cameras have a much larger field of view than the display unit (Figure A.2). Only a fraction of a camera's image (proportional to the display's field of view) is actually shown in the corresponding display via re-projection (Figure A.3). The cameras acquire enough imagery to allow full stereo overlap from close range to infinity (parallel viewing).

Thus, by enlarging the cameras' field of view, we remove the need to physically toe in the camera to change convergence. But what about the display? To preserve the above mentioned alignment between display content and peripheral vision, the display would have to physically toe in for close-up work, together with the cameras, as with the device described in [Fuchs98]. While this is doubtlessly desirable, we have determined that it is in fact possible to operate a device with fixed, parallel-mounted displays in this way, at least for a majority of our users. This surprising finding might be easier to understand by considering that if the displays converged physically while performing a near-field task, the user's eyes would also verge inward to view the task-related objects (presumably located just in front of the user's nose). With fixed displays however, the user's eyes are viewing the very same retinal image pair, but in a configuration which requires the eyes to not verge in order for stereoscopic fusion to be achieved.

Thus virtual convergence always provides images that are aligned for parallel viewing. By preventing (relieving?) the user from converging her eyes, it allows stereoscopic fusion of extremely close objects even in display units that have little or no stereo overlap at close range. This is akin to wall-eyed fusion of certain stereo pairs in printed matter (including the images in this paper, Figures A.3-bottom, A.9, and A.10-bottom, HMD image only) or to the horizontal shifting of stereo image pairs on projection screens in order to reduce ghosting when using polarized glasses. It creates a disparity-vergence conflict (not to be confused with the well-known accommodation-vergence conflict present in most stereoscopic displays [Drascic96]). For example, if we point converging cameras at an object located 1m in front of the cameras, then present the image pair to a user in a HMD with parallel displays, the user will not converge his eyes to fuse the object but will nevertheless perceive it as being much

Table A.1: Depth cues and depth cue conflicts for close-range work: Enabling virtual convergence maximizes stereo overlap for close-range work, but "moves" the vergence cue to infinity.

Virtual convergence settings	Available close-range stero overlaps	Where are depth cues, accomadation (A), disparity, (D) and vergence, (V)?		Conflicts between depth
		Close range	2m through ∞	
OFF	partial	D, V	A	A-D, A-V
ON	full	D	A, V	A-D, D-V

closer than infinitely far away due to the disparity present in the image pair. This indicates that the disparity depth cue dominates vergence in such situations; our system takes advantage of this fact. Also, by centering the object of interest in the camera images and presenting it on parallel displays, we all but eliminate the accommodation-vergence conflict for the object of interest, assuming that the display is collimated. In reality, HMD displays are built so that their images appear at finite but rather large (compared to the close range we are targeting) distances to the user, for example, two meters in the Sony Glasstron device we use (described below). Even so, users of a virtual convergence system will experience a significant reduction of the accommodation-vergence conflict, since virtual convergence reduces screen disparities (in our case, the screen is the virtual screen visible within the HMD). Reducing screen disparities is often recommended [Akka92] if one wishes to reduce potential eye strain caused by the accommodation-vergence conflict. Table A.1 shows the relationships between the three depth cues accommodation, disparity and vergence for our VST-HMD with and without virtual convergence, assuming the user is attempting to perform a close-range task.

By eliminating the moving parts, we are now in a position to dynamically change the virtual convergence. Our implementation allows the computer system to make an educated guess as to what the convergence distance should be at any given time and then set the display (re)projection transformations accordingly. The following sections describe our hardware and software implementation and present some application results as well as informal user reactions to this technology.

A.3.1 Hardware implementation

We use a Sony Glasstron LDI-D100B stereo HMD with full-color SVGA (800x600) stereo displays, a device we have found to be very reliable, characterized by excellent image quality even when compared to considerably more expensive commercial units. (Unfortunately, it is no longer on the market.) It has a horizontal field of view of $\alpha = 26°$. The display-lens elements are built $d = 62mm$ apart and cannot be moved to match a user's inter-pupillary distance (IPD). However, the displays' exit pupils are large enough [Robinett92] for users with IPDs between roughly 50 and 75mm. Nevertheless users with extremely small or extremely large IPDs will perceive a prismatic depth plane distortion (curvature) since they view images through off-center portions of the lenses; we do not address this issue here any further. We have mounted two Toshiba IK-M43S miniature lipstick cameras on top of this device (Figure A.4). The cameras are mounted parallel to each other. The distance between them is also 62mm. There are no mirrors or prisms, hence there is a significant eye-camera offset (about 60-80mm horizontally and about 20-30mm vertically, depending on the wearer). In addition, there is an IPD mismatch for any user whose IPD is significantly larger or smaller than 62mm.

The head-mounted cameras are fitted with 4mm-focal-length lenses providing a field of view of approximately β=50° horizontal, nearly twice the displays' field of view. It is typical for small wide-angle lenses to exhibit barrel distortion, and in our case the barrel distortion is non-negligible and must be eliminated (per software) before attempting to register any synthetic imagery to it. The entire head-mounted device, consisting of Glasstron, lenses, and an aluminum frame on which cameras and infrared LEDs for tracking are mounted, weighs well under 250 grams. (Weight was an important issue in this design since the device is used in extended medical experiments and is often worn by our medical collaborator for an hour or longer without interruption.)

Our AR software runs on an SGI Reality Monster equipped with InfiniteReality2 (IR2) graphics pipes and DIVO digital video I/O boards. The HMD cameras' video streams are converted from S-video to a 4:2:2 serial digital format via Miranda picoLink ASD-272p decoders and then fed to two DIVO boards. HMD tracking information is provided by an Image-Guided Technologies FlashPoint 5000 opto-electronic tracker. A graphics pipe in the SGI delivers the stereo left-right augmented images in two SVGA 60 Hz channels. These are combined into the single-channel left-right alternating 30Hz SVGA format required by the Glasstron with the help of a Sony

CVI-D10 multiplexer.

A.3.2 Software implementation

Our AR applications are largely single-threaded, using a single IR2 pipe and a single processor. For each synthetic frame, we capture a frame from each camera via the DIVO boards. When it is important to ensure maximum image quality (considering that we will end up looking in close-up at a (re-projected) portion of an NTSC-resolution image) we capture two successive NTSC fields, even though that may lead to the well-known visible horizontal tearing effect during rapid user head motion.

DIVO-captured frames are initially deposited in main memory, from where they are transferred to texture memory. Before any graphics can be superimposed onto the camera imagery, it must be rendered on textured polygons. We use a 2D polygonal grid which is radially stretched (its corners are pulled outward) to compensate for the above mentioned lens distortion (Figure A.5), analogous to the pre-distortion technique described in [Watson95]. The distortion compensation parameters are determined in a separate calibration procedure; we have found that both a third-degree and a fifth-degree coefficient are needed in the polynomial approximation [Robinett92]. The stretched, video-texture-mapped polygon grids are rendered from the cameras' points of view (using tracking information from the FlashPoint unit and inter-camera calibration data acquired during yet another separate calibration procedure).

In a conventional video-see-through application one would use parallel display frustums to render the video texturessince the cameras are parallel (as recommended by [Takagi00]). Also, the display frustums should have the same field of view as the cameras. However, for virtual convergence, we use display frustums that are verged in. Their field of view is equal to the display's field of view α. As a result of that, the user ends up seeing a reprojected (and distortion-corrected) sub-image (Figure A.6) in each eye.

The maximum convergence angle is $\delta = \beta - \alpha$, in our case approximately 24°. At that convergence angle, the stereo overlap region of space begins at a distance

$$z_{over,min} = 0.5d \tan(90° - \beta/2) \tag{A.1}$$

in our case approximately 66mm, and full stereo overlap is achieved at a distance

$$z_{over,full} = \frac{d}{(\tan(\beta/2) - \tan(\alpha - \beta/2))} \tag{A.2}$$

in our case approximately 138mm. At that latter distance the field of view subtends an area that is $d + 2z_{over,full} \tan(\alpha - \beta/2)$ wide, or approximately 67mm in our case.

After setting the display frustum convergence, application-dependent synthetic elements are rasterized using the very same verged, narrow display frustums. For some parts of the real world we have registered geometric models (Figure A.7), and so we can rasterize those in Z only, thereby priming the Z-buffer for correct mutal occlusion between real and synthetic elements [State96b]. As shown in Figure A.7, only part of the patient surface is known. The rest is extrapolated with straight lines to approximate the size of a human. There are static models of the table and of the ultrasound machine (Figure A.1), as well as of the trackeed handheld objects [Lee02]. Floor and lab walls are modeled coarsely with only a few polygons.

A.3.3 Sheared vs. rotated display frustums

One issue that we considered early on during the implementation phase of this technique was the question of whether the verged display frustums should be sheared or rotated (Figure A.8). Shearing the frustums keeps the image planes for the left and right eyes coplanar, thus eliminating vertical disparity or *dipvergence* [Rolland92] between the two images. At high convergence angles (i.e., for extreme close-up work), viewing such a stereo pair in our system would be akin to wall-eyed fusion of images specifically prepared for cross-eyed fusion.

On the other hand, rotating the display frustums with respect to the camera frustums, while introducing dipvergence between corresponding features in stereo images, presents to each eye the very same retinal image it would see if the display were capable of physically toeing in (as discussed above), thereby also stimulating the user's eyes to toe in.

To compare these two methods for display frustum geometry, we implemented an interactive control (slider) in the system's user interface. For a given virtual convergence setting, we can smoothly blend between sheared and rotated frustums by moving the slider. When that happens, the HMD user perceives a curious distortion of space, maybe similar to a dynamic prismatic distortion. We did not conduct a controlled user study to determine whether sheared or rotated frustums are preferable; we merely assembled an informal group of testers (members of our group and other researchers) and there was a definite preference towards the rotated frustums method overall. However, none of the testers found the sheared frustum images more difficult

to fuse than the rotated frustum ones, which is understandable given that sheared frustum stereo imagery has no dipvergence (as opposed to rotated frustum imagery). It is of course difficult to quantify the stereo perception experience without a carefully controlled study; for the time being we relied solely on users' preferences as guidance for further development.

A.3.4 Automating virtual convergence

Our goal was to achieve on-the-fly convergence changes under algorithmic control to allow users to work comfortably at different depths. We initially tested whether a human user could in fact tolerate dynamic virtual convergence changes at all. To this end, we implemented a user interface slider controlling convergence and assigned a human operator to continually adjust the slider while a user was viewing AR imagery in the VST-HMD.

The convergence slider operator was permanently watching the combined left-right (alternating at 60Hz) SVGA signal fed to the Glasstron HMD on a separate monitor. This signal appears like a blend between the left and right eye images, and any disparity between the images is immediately apparent. The operator was continually adjusting the convergence slider, attempting to minimize the visual disparity between the images (thereby maximizing stereo overlap). This means that if most of the image consists of objects located close to the HMD user's head, the convergence slider operator tended to verge the display frustums inward. With practice our operator became quite skilled; most test users had positive reactions, with one notable exception (a member of our team) reporting extreme discomfort.

Our next goal was to create a real-time algorithmic implementation capable of producing a numeric value for display frustum convergence for each frame in the AR system. We considered three distinct approaches for this:

(1) Image content based: this is the algorithmic version of the "manual" method described above. An attractive possibility would be to use a maximization of mutual information algorithm [Viola95]. An image-based method could run as a separate process and could be expected to perform relatively quickly since it need only optimize a single parameter. This method should be applied to the mixed reality output rather than the real world imagery to ensure that the user can see virtual objects that are likely to be of interest. Under some conditions, such as repeating patterns in the images, a mutual information method would fail by finding an "optimal" depth value with no rational basis in the mixed reality. Under most

conditions however, including color and intensity mismatches between the cameras, a mutual information algorithm would appropriately maximize the stereo overlap in the left and right eye images.

(2) Z-buffer based: this approach inspects values in the Z-buffer of each stereo image pair and (heuristically) determines a likely depth value to set the convergence to. [Ware98] gives an example for such a technique.

(3) Geometry based: this approach is similar to (2) but uses geometry data (models as opposed to pixel depths) to (again heuristically) compute a likely depth value to set the convergence to. In other words, it works on pre-rasterization geometry, whereas (2) uses post-rasterization geometry.

Approaches (1) and (2) both operate on finished images. Thus they cannot be used to set the convergence for the current frame but only to predict a convergence value for the next frame. Conversely, approach (3) can be used to immediately compute a convergence value (and thus the final viewing transformations for the left and right display frustums) for the current frame, before any geometry is rasterized. However, as we shall see, this does not automatically exclude (1) and (2) from consideration. Rather, approach (1) was eliminated on the grounds that it would require significant computational resources. We developed a hybrid of (2) and (3), characterized by inspection of only a small subset of all Z-buffer values, and aided by geometric models and tracking information for the user's head as well as for handheld objects.

The following steps describe our hybrid algorithm:

1. For each eye, the full augmented view (as described in A.3.2) is rendered into the frame buffer (after capturing video, reading trackers, etc.).

2. For each eye, inspect the z-buffer of the finished view along 3 horizontal scan lines, located at heights $h/3$, $h/2$, and $2h/3$ respectively, where h is the height of the image (Figure A.9). Find the average of the closest depths $z_{min} = (z_{min,l} + z_{min,r})/2$. Set the convergence distance z to z_{min} for now. This step is only performed if in the previous frame the convergence distance was virtually unchanged (we use a threshold of 0.01°). Otherwise z is left unchanged from the previous frame.

3. Using tracker information, determine if application-specific geometry (for example, the all-important ultrasound image in our medical application) is within the viewing frustum of either display. If so, set z to the distance of the ultrasound slice from the HMD.

4. Calculate the average value z_{avg} during the most recent n frames, not including the current frame since the above steps can only execute on a finished frame (steps 1-2) or at least on an already calculated display frustum (step 3).

5. Set the display frustums to point to a location at distance z_{avg} in front of the HMD. Calculate the appropriate transformations, taking into account the blending factor between sheared and rotated frustums (see A.3.3). Go to step 1.

The simple temporal filtering in step 4 is used to avoid sudden, rapid changes. It also adds a delay in virtual convergence update, which for $n = 10$ amounts to approximately 0.5 seconds at our current frame rate of about 20 Hz (a better implementation would vary n as a function of frame rate in order to keep the delay constant). Even though this update seems slower than the human visual system's rather quick vergence response to the *diplopia* (double vision) stimulus, we have not found it jarring or unpleasant.

The conditional update of z in Step 2 prevents most self-induced oscillations in convergence distance. Such oscillations can occur if the system continually switches between two (rarely more) different convergence settings, with the z-buffer calculated for one setting resulting in the other convergence setting being calculated for the next frame. Such a configuration may be encountered even when the user's head is perfectly still and none of the other tracked objects (such as handheld probe, pointers, needle, etc.) are moved.

A.4 Results

Figure A.10 contains simulated wide-angle stereo views from the point of view of an HMD wearer, illustrating the difference between converged and parallel operation.

The dynamic virtual convergence subsystem has been deployed within two different AR applications. Both use the same modified Sony Glasstron HMD and the hardware and software described in A.3. The first is an experimental AR system designed to aid physicians in performing minimally invasive procedures such as ultrasound-guided needle biopsies of the breast. This system and a number of recent experiments (Figure A.11) conducted with it are described in detail in [Rosenthal01]. Our principal medical collaborator used the system on numerous occasions, often for one hour or longer without interruption, while the dynamic virtual convergence algorithm was active. She did not report any discomfort while or after using the system. With her

help, we successfully conducted a series of experiments yielding quantitative evidence that AR-based guidance for the breast biopsy procedure is superior to the conventional guidance method in artificial phantoms [Rosenthal01]. Other physicians, the authors of this paper and other members of our team, as well as several lab visitors have all used this system, albeit for shorter periods of time, without discomfort (except for one individual previously mentioned, who experiences discomfort whenever the virtual convergence is changed dynamically).

The second AR application to use dynamic virtual convergence is a system for modeling real objects using AR (Figure A.12). The system and the results obtained with it were described in detail [Lee02]. Two of the authors of [Lee02] have used that system for sessions of one hour or longer, again without noticeable discomfort (immediate or delayed).

A.5 Conclusions

Other authors have previously noted the conflict introduced in VST-HMDs when the camera axes are not properly aligned with the displays. While we continue to believe that this is significant, our experience with this technique suggests that violating this constraint may be advantageous in systems requiring the operator to use stereoscopic vision at several distances.

Mathematical models such as those developed by [Takagi00] demonstrate the distortion of the visual world. These models do not demonstrate the volume of the visual world that is actually stereo-visible (i.e., visible to both eyes and within 1-2 degrees of center of stereo-fused content). Dynamically converging the cameras-whether they are real cameras as in [Matsunaga00] or virtual cameras (i.e., display frustums) pointed at video-textured polygons as in our case-makes a greater portion of the near field around the point of convergence stereoscopically visible at all times.

Most users have successfully used our AR system with dynamic virtual convergence to place biopsy and aspiration needles with high precision or to model objects with complex shapes. Our experience suggests that the distortion of the perceived visual world is not as severe as predicted by the mathematical models if the user's eyes converge at the distance selected by the system. (If they converge at a different distance, stereo overlap is reduced and increased spatial distortion and/or eye strain may be the result. We therefore believe that our largely positive experience with this technique is due to a well-functioning convergence depth estimation algorithm.)

Indeed, a substantial degree of perceived distortion is eliminated if one assumes that the operator has approximate knowledge of the distance to the point being converged on (experimental results in [Milgram92] support this statement). Given the intensive hand-eye coordination required for our applications, it seems reasonable to conjecture that our users' perception of their visual world may be rectified by other sources of information such as seeing their own hand. Indeed, the hand may act as a "visual aid" as defined by [Milgram92].

This type of adaptation is apparently well within the abilities of the human visual system as evidenced by the ease with which individuals adapt to new eyeglasses and to using binocular magnifying systems. On the other hand, while our approach has proved surprisingly unproblematic, we do not consider it superior to rigorous orthostereoscopy. We would therefore like to encourage HMD manufacturers to put more display pixels towards the wearer's nose in future designs.

A.6 Future Work

Our new technique reduces the accommodation-vergence conflict while introducing a disparity-vergence conflict. It may be useful to investigate whether smoothly blending between zero and full virtual convergence is useful. Also, should that a parameter to be set on a per user basis, per session basis, or dynamically?

Second, a thorough investigation of sheared vs. rotated frustums (should that be changed dynamically as well?), as well as a controlled user study for the entire system, with the goal of obtaining quantitative results, seem desirable.

Finally, we plan to use our technique on a parallax-free device. To this end, we have mounted a mirror-camera device on a Sony Glasstron HMD. This new orthoscopic device has recently been incorporated into our system and we plan to report on our experience with it in a future publication. (Of course, the term "orthoscopic" does not apply when virtual convergence is used.)

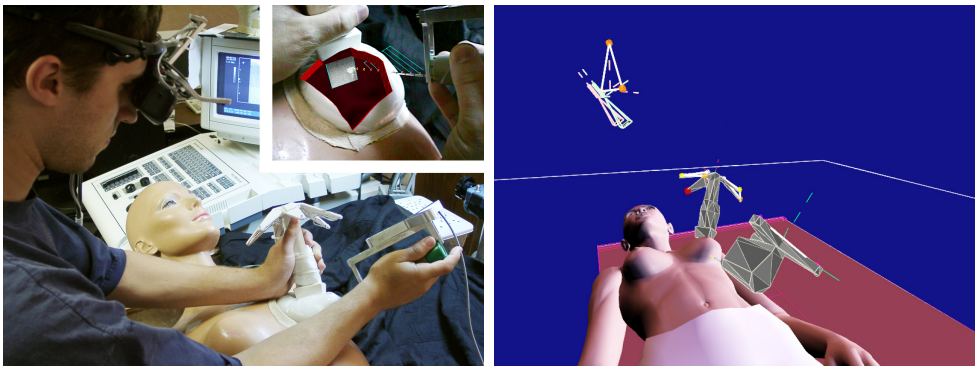


Figure A.1: Left: AR guidance system in use on a breat biopsy training phantom. The user holds the ultrasound probe (left hand) and a biopsy needle(right hand). Inset: typical HMD view shows synthetic opening into the patient and registered ultrasound image scanning a suspicious lesion. Right: The system displays a control view with dynamic avatars for the optically tracked VST-HMD, probe, and needle.

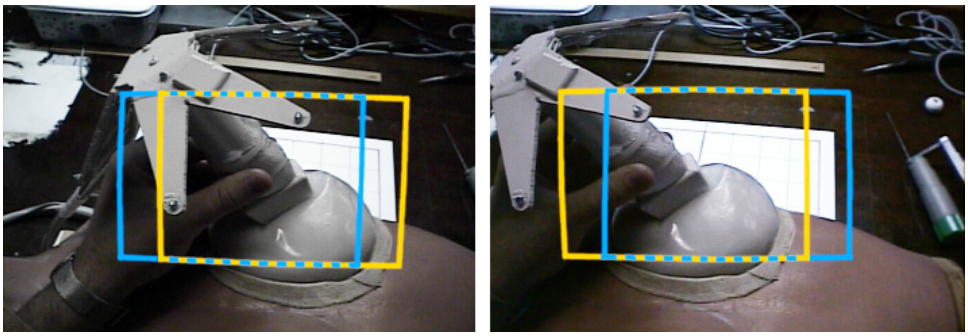


Figure A.2: Wide-angle stereo views (with barrel distortion) as acquired by the HMD cameras. The blue (virtual convergence off) and yellow (virtual convergence on) outlines show the re-projected parts of the video images corresponding to the HMD images in Figure A.3–curved because of distortion. (fuse all stereo pairs wall-eyed)

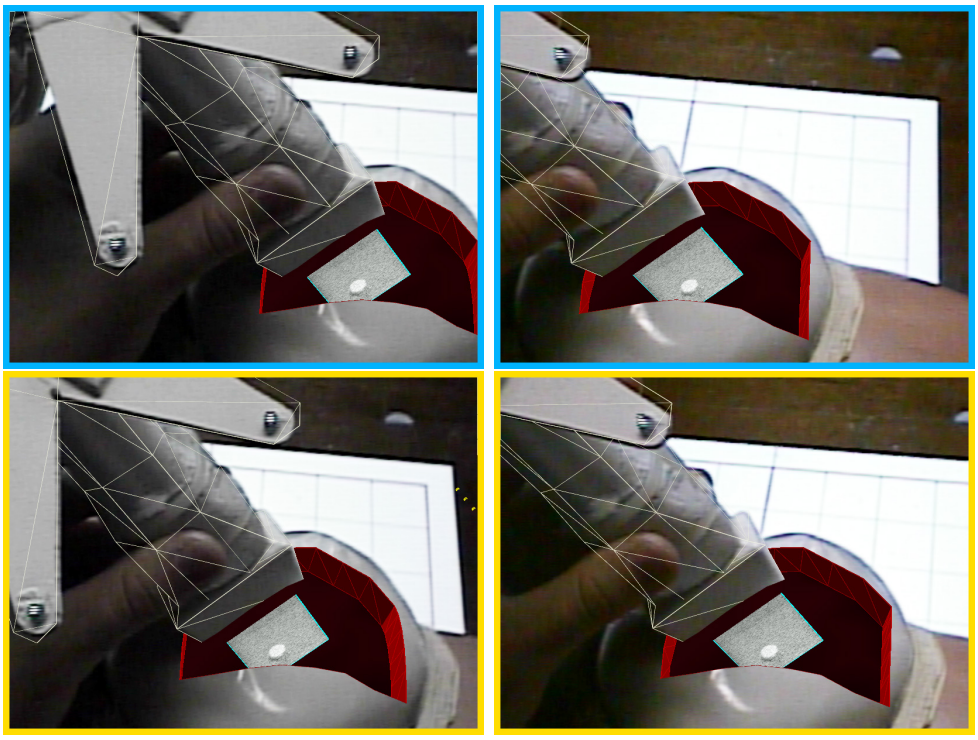


Figure A.3: Stereo images displayed in the HMD without (top) and with virtual convergence (bottom), all distortion-corrected.

Figure A.4: VST-HMD built from Sony unit. The frame holds cameras and IR tracking LEDs

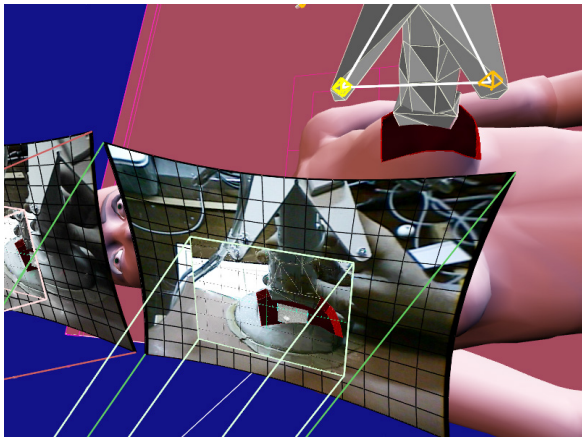


Figure A.5: A deformed polygonal grid removes the video texture distortion (exaggerated). Smaller display frustum has re-pprojected, distortion-corrected image shown in HMD.

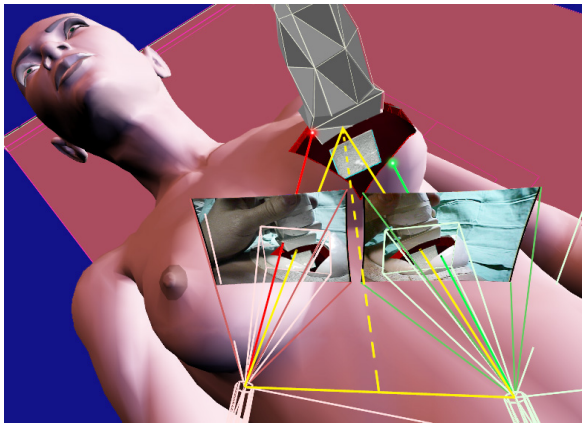


Figure A.6: Wide-FOV (camera) frustums and narrow, converged display frustums. The yellow isosceles triangle indicates display frustum convergence.

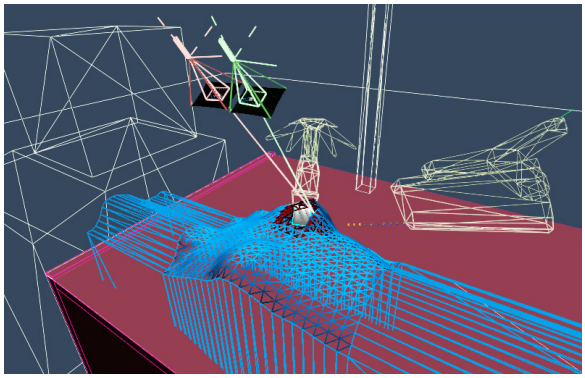


Figure A.7: Scene geometry known to the AR system.

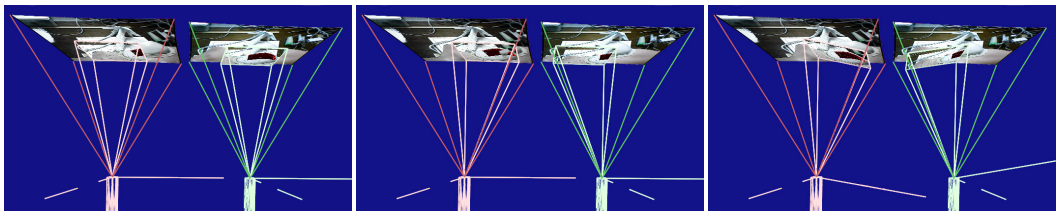


Figure A.8: Unconverged (left), sheared (center), and rotated (right) display frustums.

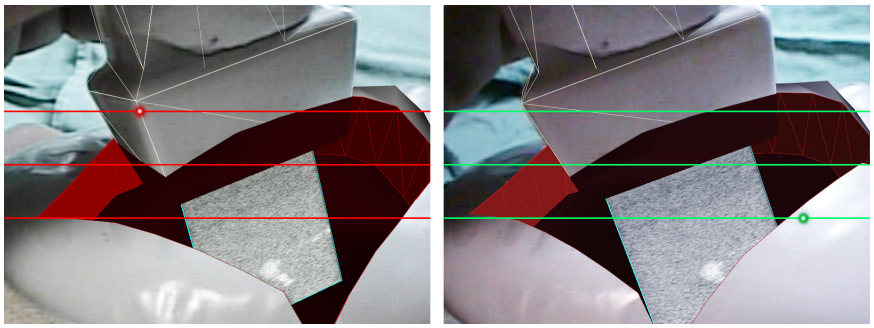


Figure A.9: Z-buffer inspection on 3 selected scan lines in each eye. The highlighted points mark the closest depth values found, correspondng to the red/green lines in Figure A.6.

Figure A.10: Simulated wide-angle stereo views through and "around" HMD (left and center). Top: Virtual convergence is off. Bottom: virtual convergence is on. Note how alignment between features in the display and below the display (for example, between nipples, which are vertically aligned in the top stereo pair) is lost with virtual convergence, illustrating the new disparity-vergence conflict.

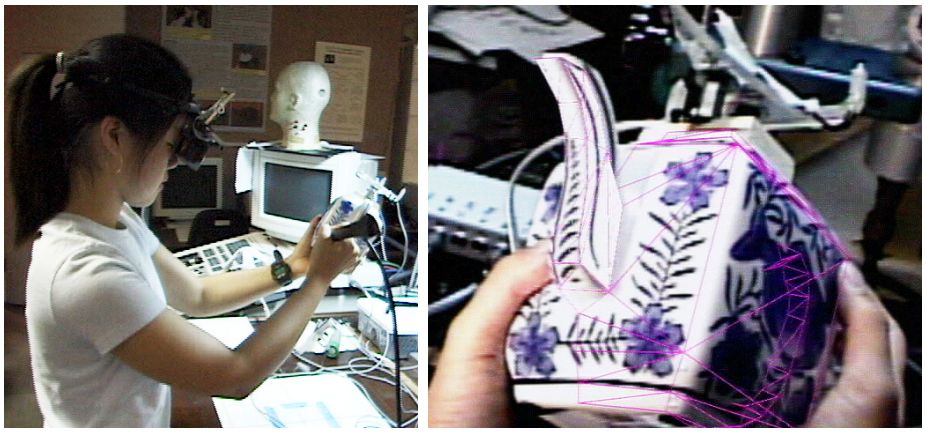


Figure A.11: Left: Lab view of AR ultrasound human subject experiment. Right: typical HMD view during the procedure.

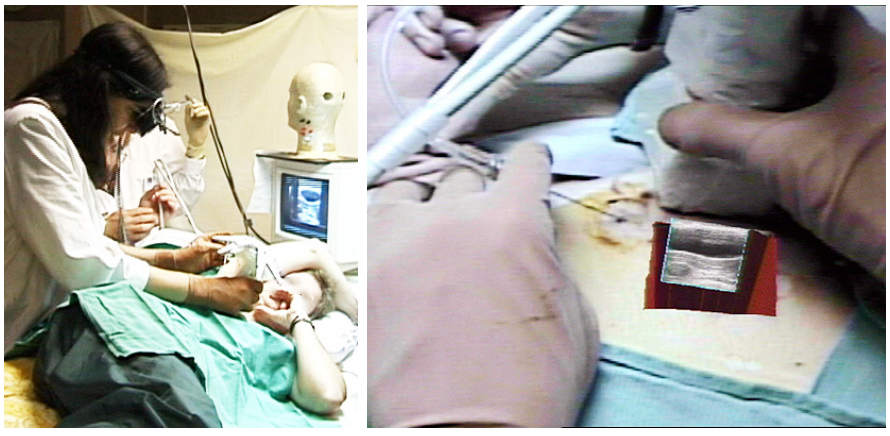


Figure A.12: Left: AR system used to model tracked real world objects with textures. Right: HMD while system is in use.

Appendix B

Simulation of depth extraction

B.1 Introduction

High speed depth extraction is an important component needed for many applications. In the context of medical applications, high speed depth extraction techniques may be used to capture data on the three-dimensional structure of the environment or as the source data for generating virtual imagery.

Simulation of depth extraction serves several purposes. The first is to predict what type of results we can reasonably hope to achieve with a particular depth extraction technology and configuration before a working prototype device is built. The second factor is to be able to test algorithms that visually correct errors in depth extraction or visual errors in rendering that occur systematically as a result of known characteristics of the depth extraction method.

Two simulation approaches are described here. The first is a simple environment for testing methods of correcting the visualization created from simulated depth extraction. The second is a simple simulation environment to model the characteristics of a complete depth extraction system with different design parameters (e.g., field of view, number of stripes).

B.2 A test rendering environment[1]

In the course of experiments conducted as part of the Medical Augmented Reality Visualization project, the following image was created (Figure B.1).

[1] **Some of this work was completed with Amy Henderson for a class project, the text here is excerpted from the final report on the project**

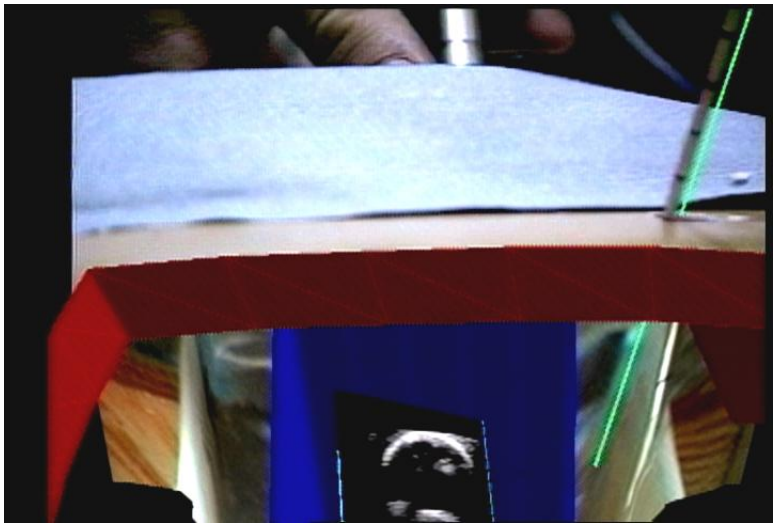


Figure B.1: Example image of a complex visualization combining simulated laparoscopically depth extracted surfaces, live intraoperative ultrasound, and a tracked instrument.

On the right side of the image a biopsy needle (3mm in diameter) is shown penetrating through the abdominal wall of the anatomical model. The view of objects outside the abdomen is primarily (except for the green line) what is acquired by the camera mounted on the head-mounted display of the user. The objects inside the abdomen are synthetically reconstructed. The red and blue polygons are computer generated and represent a "synthetic pit" through which the user may visualize the patient's internal organs. The other internal structures are reconstructed by projectively texture mapping the view from a tracked camera onto a pre-acquired polygon mesh representing the internal surface of the anatomical model. A green line is drawn by the computer to show the tracked position of the needle both inside and outside of the abdomen.

Outside of the abdomen, the tracked needle position and actual needle position coincide well. However, inside the abdomen, clearly something strange happens. The image of the needle is mapped onto the surface, but this produces incorrect-appearing results as seen in the above image. The green line and the mapped image of the needle do not coincide at all. The needle would have to be broken or severely bent for both its position inside and outside the abdomen to be correct. This is particularly a problem since the image of the needle is shaded, making it look more realistic than the green line, which more accurately represents the needle's true position.

Although it is not currently possible to extract depth inside the abdomen through

a laparoscope in real-time, the research group hopes to incorporate that technology into their system. We anticipate that an instrument may distort the depth extraction in the part of the image it covers. In both the current situation and (presumably) in the future, the instrument is tracked, and this information may be used to compensate for these irregularities.

Thus the motivation for this project.

B.2.1 Techniques

Depth extraction of an arbitrary scene can be simulated by reading back the z-buffer. The z-buffer is a component of many current graphics pipelines which resolves possible occlusions. An object in front of another one should fully block the overlapping areas and this is achieved by storing a depth value for each currently rendered pixel. Objects rendered later are compared on a pixel-by-pixel basis with the z-buffer and the color and z-values are replaced if the new object is closer than the one already rendered. It is not possible to read the z-buffer back into main memory on all hardware making use of this technique.

The z-buffer does not directly store the depth value, but instead stores a scaled value proportional, scaled plus a constant, to $1/z$. The numerical precision of the z-buffer is therefore very similar (depending on the number of significant figures) to depth from stereo and structured light triangulation techniques.

All of our correction methods rely on basic techniques used in several shadowing algorithms. In several shadowing methods, the scene is rendered/projected from the perspective of the light source(s). The information can be used in multiple different ways to create shadows. The projected position of the instrument (as seen in the image above) is essentially where the shadow of the instrument would be if the camera were a light source. Because of this observation, our correction techniques employ the same first step as used in shadowing algorithms.

In Figure B.3 the scene is greatly distorted by the range data acquired on the instrument. This is very disconcerting, and the protruding surface caused by the instrument may obstruct the user's view of the rest of the scene. Other distortions are visible in the scene and are caused by difficulties in re-rendering a range map as encountered in image based rendering.

In the Figure B.4 distortions akin to those seen in the current laparoscopic visualization system can be seen. The green line/cylinder in the image on the right represents the true position of the instrument. The gray "shadow" curving across the

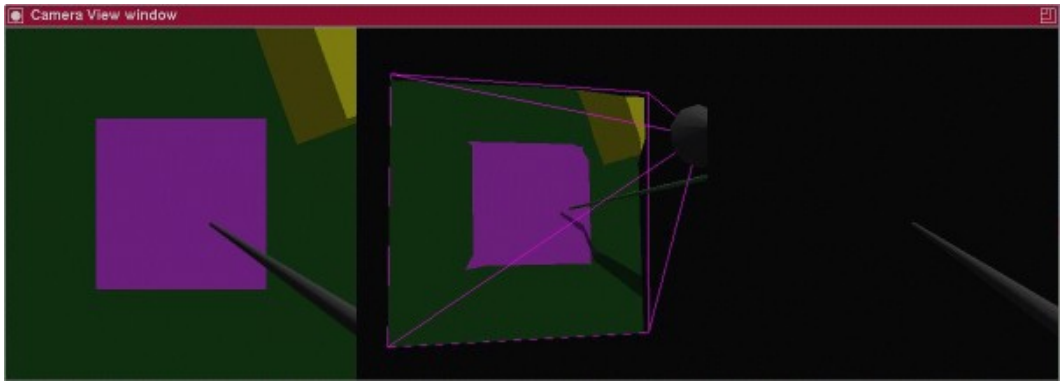


Figure B.2: The graphics window of our program (shrunk to 1/4 actual size) running on an SGI. The window is composed of three viewports. Starting on the left, the first viewport contains the camera's view of the scene. The middle view shows the off-axis view of the scene rendered using the depth and color information gathered from the camera's view. The view on the right shows only the tracked instrument rendered from the perspective of the camera. The images from this program shown later in this report will only show the camera and off-axis views.

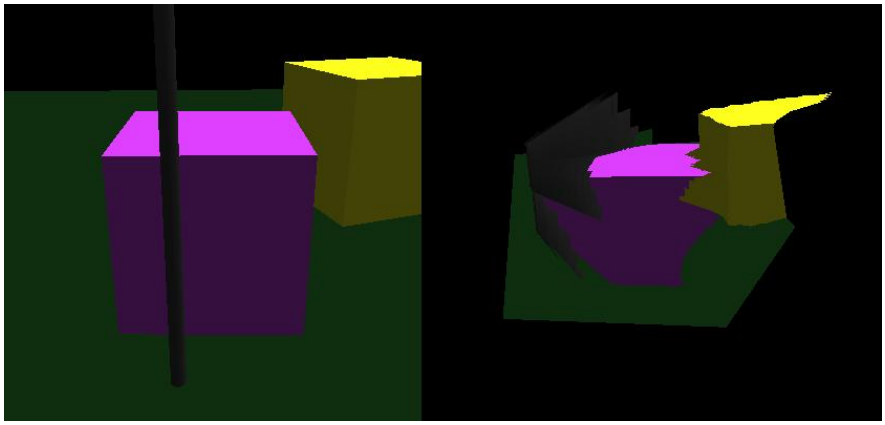


Figure B.3: Expected renderings of a depth extracted view of a scene with laparoscopic instruments

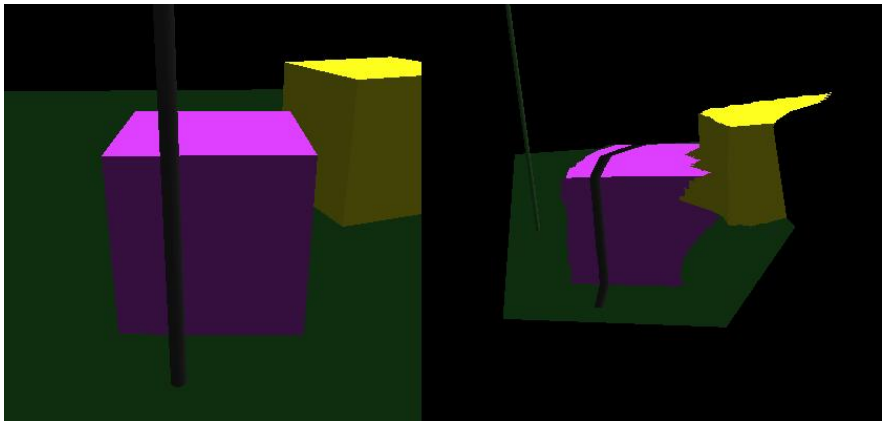


Figure B.4: A rerendering of a gold standard for depth extraction: a highly sampled z-buffer before laparoscopic instruments have been rendered over the surface, a texture map acquired after the instrument was rendered is mapped on the surface.

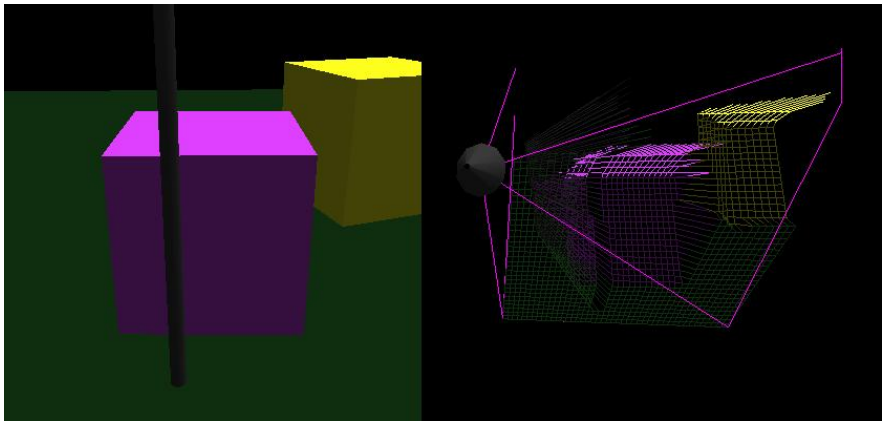


Figure B.5: Rerendering extracted depth with an instrument between the surface and the camera.

surface of the other objects is the projected version of the instrument. Because this "shadow" is shaded one might mistake the shadow for the actual instrument. Indeed this is the image of the actual image in the texture map applied to z-sampled surface.

B.2.1.1 Depth Correction

Figure B.5 demonstrates the problem of extracting depth from a scene containing a tracked instrument. On the right many polygons whose depth was acquired where the instrument can be seen by the camera can be seen to extrude away from the surface of the other objects in the scene.

Ideally one would like to get range maps more like the one in Figure B.6. This

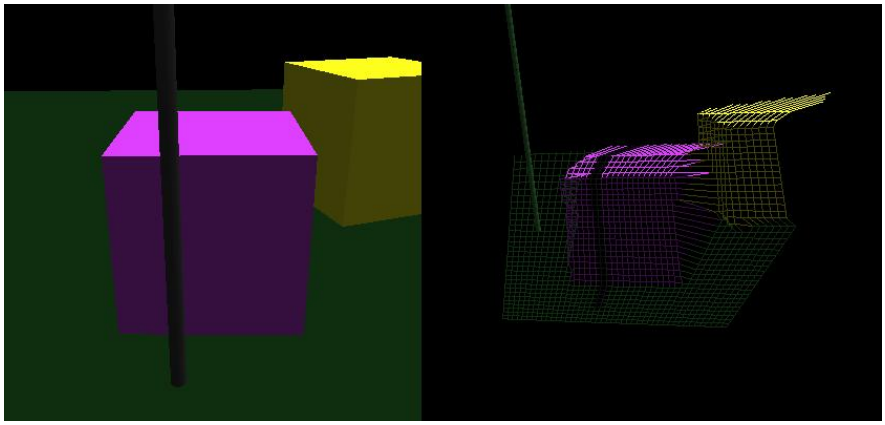


Figure B.6: Z-buffer captured before rendering instruments: a gold standard to compare z-buffer correction algorithms with.

range map, our "gold standard", was created by acquiring the z-buffer of the camera view before the tracked instrument (the gray cylinder) was drawn. Clearly, in the real world this approach is not possible and therefore correction techniques are needed.

Figure B.7 demonstrates our first approach to depth correction. We use the color buffer from the rendering of the instrument alone to determine which points in the range map are likely to be distorted by the instrument. These range values are then set to the mean range value found in the entire depth map. In many instances this approach is sufficient, in that it produces tolerable images. In the image below two artifacts created by this approach are visible. Where the instrument shadows the purple cuboid, the range map on the surface of the cuboid can be seen to be slightly depressed. Where the instrument is visible against the horizon and the green plane, the range map is obviously protruding from the true range values. Clearly there are many instances where this approach is insufficient.

Figure B.8 shows the results of a more intelligent approach to this problem. We again use the color buffer from the rendering of the instrument alone to determine which points on the range map are affected. We then traverse neighboring points on the range map until one or more points which are not shadowed by the instrument are reached. The new range value for this point is then an interpolation of the range values found. The results from this approach are usually visually indistinguishable from the "gold standard". The primary failure mode of this technique occurs when the geometry of the scene has rapidly changing range values.

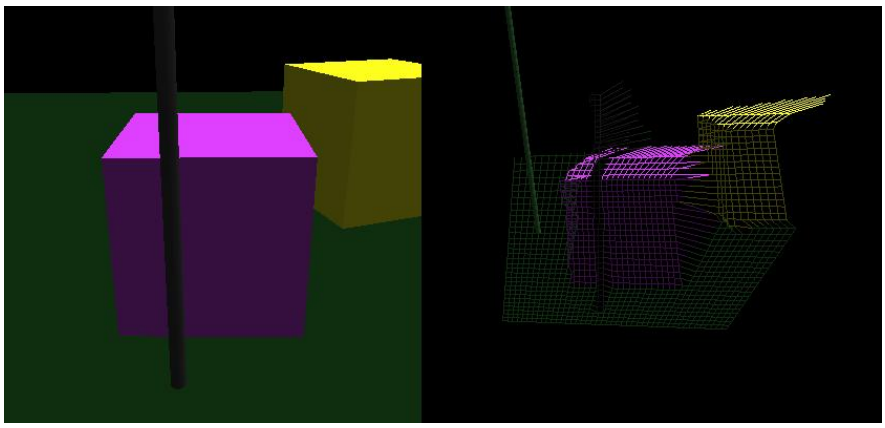


Figure B.7: Our first approach to correcting the range information.

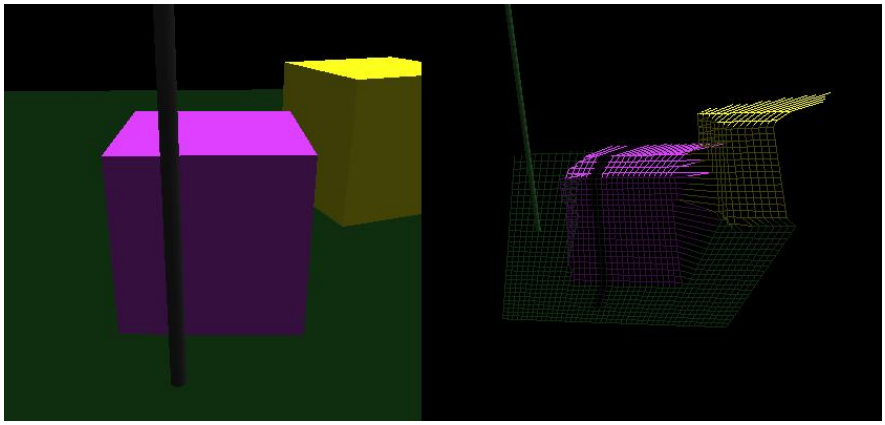


Figure B.8: Using the rendering of the tracked position of the instrument to identify areas of the range image that should be corrected

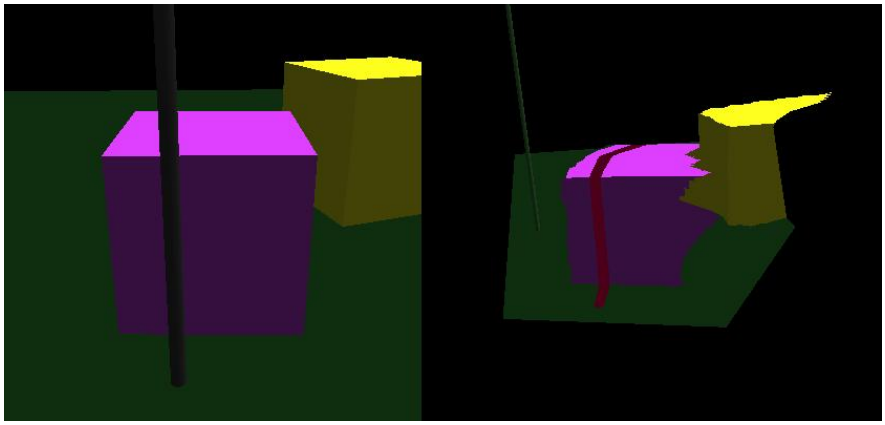


Figure B.9: A naive approach to correcting texture.

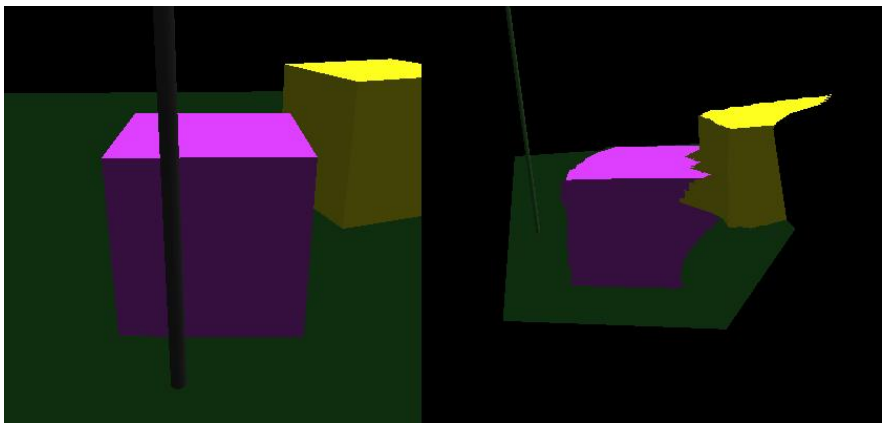


Figure B.10: Basic texture correction.

B.2.1.2 Color Correction

The problems observed in the current laparoscopic visualization system which were the primary motivation for this project are actually problems that occur from the reprojection of the image acquired from the laparoscopic camera (the texture). Approaches similar to those used to correct the range map may be used to correct the texture before it is reprojected. An image from our program demonstrating this problem can be seen in the first section as can an image from the laparoscopic visualization system.

Figure B.9 demonstrates a fairly naive approach to correcting the texture. Since we do not want the user to think they have real texture information from the texels covered by the instrument, we simply color all instrument texels a hideous reddish/pinkish color. This is obviously also a shadowing algorithm.

Figure B.10 shows a slightly more sophisticated method for correcting the texture

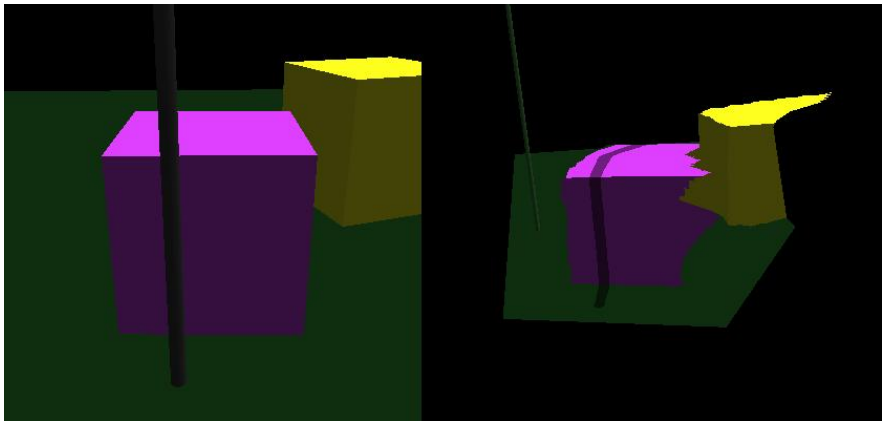


Figure B.11: Texture and depth correction simultaneously applied.

map. Similar to our approach to correction of the range map, we search for nearest neighbor texels which are not shadowed by the instrument. The colors from these texels can then be interpolated to calculate the corrected color for that texel. This works really well with untextured computer generated polygon models. The results with textured (or real-world) objects may not be so visually appealing. In the above image it is difficult (if not impossible) to determine which texels have been corrected. In some applications (such as in surgery) this is an undesirable property. We describe an additional technique in the following section.

B.2.1.3 Combining Corrections

In the future laparoscopic visualization system, both our range map and texture map corrections may be applied together to produce more understandable images in augmented reality. As noted above the user of such a system should be made aware of where significant "adjustments" have been made before rendering. To that end, the corrected texels and range values should somehow be indicated on the final image. It is also important that the marking not be so distracting that the user is has difficulty focusing on unaffected areas of the image. Our solution to this problem is shown in the image below.

We apply our best range and texture correcting techniques simultaneously for best results (Figure B.11. During color correction we darken all corrected colors slightly, thereby generating the appearance of a shadow. The shadow, in addition to the benefits listed above, may be useful to the user in keeping track of camera position.

Several other approaches were considered, but after seeing the results with this method we decided more complex methods would not likely produce better results.

B.2.2 Conclusions and Future Work

We created a demonstration system which runs without in real-time. This is a good sign because there is a chance that this could be incorporated into the laparoscopic visualization system of the future. It would be useful to extend our system to use a texture mapped scene to evaluate the effectiveness of our color correction methods in a more complex environment.

B.2.3 Reference List

1. Rendering antialiased shadows with depth maps William T. Reeves David H. Salesin Robert L. Cook ACM SIGGRAPH Computer Graphics Vol. 21, No. 4 (July 1987), Pages 283-291

2. Casting curved shadows on curved surfaces Lance Williams in: SIGGRAPH'98

3. Fast calculation of soft shadow textures using convolution Cyril Soler François X. Sillion in: SIGGRAPH '98.

4. A fast shadow algorithm for area light sources using backprojection; George Drettakis, and Eugene Fiume; Proceedings of the 21st annual conference on Computer Graphics, 1994, Page 223

5. Foley et al, Computer Graphics Principles and Practice 2nd edition in C. 1997. Chapter 16.4 pages 745-753

B.3 A simulator for selecting design parameters

This simulator was written as a Microsoft Excel spreadsheet. Excel was chosen for expediency, for ease of debugging, and for the integrated plotting.

B.3.1 Algorithm

A real depth extraction system works by projecting a pattern onto a surface. The image of the light on the surface is then projected onto pixels of the camera. The camera image is analyzed to identify the features (stripes) and then depth is calculated based on the position of the found stripes.

For simplicity within the simulation, it is easiest to begin with a model of the surface in the camera coordinates along a single scan line. If the assumptions of pinhole geometry parallel orientation for two-dimensional camaeras are satisfied than this model is sufficent to model most possible errors.

Because binary-encoded structured light may take a significant amount of time to fully encode all the stripes, movement of the surface, or the camera, may cause significant errors. It is therefore useful to have a model of the surface that can vary over time to model these errors. While not incorperated into this simulation, surface texture of the object should be modeled as well to formulate a complete model. For completeness surface texture is included in the equations although it has not been implemented.

The distance to the surface, from the plane of pinhole of the camera and projector, in camera pixel coordinates is defined to be $Z_s(i)$ and the intensity of the texture at that point is $T_s(i)$ where i is the index of the camera pixel. If there are p pixels then $(-p/2 < i < p/2)$, so that the pixel $i = 0$ corresponds to the ray at the center of the camera's view. The field of view of the camera is defined to be κ and the field of view of the projector is defined to be ρ. The number of stripes to be projected is 2^S where S is the number of stripe patterns to be projected. The separation between center of projection of camera and projector is defined to be d.

The first step is to project the coordinates of the projected pattern onto each camera pixel. The following equation describes the relationship $T_s(i)$, i, d, κ, and ρ:

$$T_s(i) = \frac{d}{(s_p x_p - 2^{S-1}) - s_c i} \tag{B.1}$$

where $s_c = \tan(0.5\kappa)/p$, $s_p = \tan(0.5\rho)/2^S$, and x_p is the coordinate of the stripe. The variables s_c and s_p are equivalent to $1/f$ in the usual model of a pinhole camera. This equation can easily solved for x_p:

$$x_p = 2^{S-1} + \frac{1}{f_p}\left(f_c i - \frac{d}{T_s(i)}\right) \tag{B.2}$$

The value x_p can then be used to look up or calculate the appropriate projected value. Errors from the projector, including blurring of the projected imager can be added at this point. The intensity of light projected to each camera pixel, the result of the look-up or calculation, is then defined to be L_i. The intensity seen by the camera is then:

$$I_i = L_i T_i \tag{B.3}$$

Effects of blurring in the camera can be simulated by convolving I_i with an appropriate kernal.

These equations can be applied to the entire sequence of stripe patterns and recovery methods for finding features, such as those described in Chapter 10. The result should be a list of found features, in the projected texture coordinates, x_p and their corresponding found positions in camera coordinates, X_P.

These can be converted back to a depth value, D_p, as follows:

$$D_p = \frac{d}{X_p s_c - (x_p + 0.5 - 2^{S-1}) s_p} \tag{B.4}$$

The a plot of the original depth values, $Z_s(i)$, versus i can then be compared to a plot of the pairs (D_p, X_p).

B.3.2 Additional features

The implemented model allows for motion of a foreground object against the background. It also models the shadow cast by the foreground object.

B.3.3 Results

These example figures demonstrate several patterns that can be seen with this simulation including the effect of spatial aliasing of the stripe patterns on the camera, distortions due to motion of the foreground object, and the effects of mismatched camera and projector fields of view.

Bibliography

[Akka92] Akka, R. Automatic software control of display parameters for stereoscopic graphics images. In *SPIE Volume 1669: Stereoscopic Displays and Applications III* (1992), pp. 31–37.

[Alterman97] Alterman, R. I., Kall, B., Beric, A., Sterio, D., and Kelly, P. J. Pallidal targeting with the COMPASS system. *Stereotactic Functional Neurosurgery 69*, 1-4 Pt 2 (1997), 69–72.

[Argotti02] Argotti, Y., Davis, L., Outters, V., and Rolland, J. Dynamic superimposition of synthetic objects on rigid and simple-deformable objects. *Computers and Graphics 26*, 6 (2002).

[Azuma92] Azuma, R. T. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments 6*, 4 (1992), 355–385.

[Baillot00] Baillot, Y., Rolland, J., Lin, K., and Wright, D. Automatic modeling of knee-joint motion for the virtual reality dynamic anatomy (VRDA) tool. *Presence: Teleoperators and Virtual Environments 9*, 3 (2000), 223–235.

[Bajura92] Bajura, M., Fuchs, H., and Ohbuchi, R. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. *Proceedings of SIGGRAPH'92 (Chicago, Il. July 26-31, 1992) In Computer Graphics 26*, 2 (July 1992), 203–210.

[Berger99] Berger, J. W., and Shin, D. S. Computer-vision-enabled agumented reality fundus biomicroscopy. *Ophthalmology 106*, 10 (October 1999), 1935–1941.

[Berger01] Berger, J. W., and Madjarov, B. Augmented reality fundus biomicroscopy: a working clinical prototype. *Archives of Opthamology 119* (December 2001), 1815–1818.

[Berguer01] Berguer, R., Smith, W., and Chung, Y. Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. *Surgical Endoscopy 15*, 10 (October 2001), 1204–1207.

[Berquer02] Berquer, R., Smith, W., and Davis, S. An ergonomic study of the optimum operating table height for laparoscopic surgery. *Surgical Endoscopy 16*, 3 (March 2002), 416–421.

[Besl88] Besl, P. Active optical range imaging sensors. *Machine Vision and Applications 1* (1988).

[Bingham00] Bingham, J., McKie, L. D., McLoughlin, J., and Diamond, T. Biliary complications associated with laparoscopic cholecystectomy – an analysis of common misconceptions. *Ulster Medical Journal 69*, 2 (November 2000), 106–111.

[Bitner76] Bitner, J. R., Erlich, G., and Reingold, E. M. Efficent generation of the binary reflected gray code and its applications. *CACM 19*, 9 (1976).

[Boyer87] Boyer, K. L., and Kak, A. Color-encoded structured light for rapid active ranging. *Trans. PAMI 9*, 1 (1987).

[Breedveld00] Breedveld, P., Stassen, H., Meijer, D., and Jakimowicz, J. Observation in laparoscopic surgery: overview of impeding effects and supporting aids. *Journal of laparoendoscopic and advanced surgical techniques 10*, 5 (October 2000), 231–241.

[Bridgewater99] Bridgewater, F. H. G., and Mouton, W. G. Rationale and intended use for the veress needle: a translation of the orginal descriptive article. *Surgical Laparoscopy and Endoscopy 9* (1999), 241–243.

[Briel00] Briel, J., Plaisier, P., Meijer, W., and Lange, J. Is it necessary to lift the abdominal wall when preparing a pneumoperitoneum? a randomized study. *Surgical Endoscopy 14*, 9 (September 2000), 862–864.

[Bush74] Bush, R. B., Leonhardt, H., Bush, I. M., and Landes, R. R. Dr. Bozzini's Lichtleiter a translation of his original article (1806). *Urology 3*, 1 (January 1974), 119–123.

[Calvert00] Calvert, N. W., Troy, G. P., and Johnson, A. G. Laparoscopic cholecystectomy: a good buy? a cost comparision with small incision (mini) cholecystectomy. *European Journal of Surgery 166* (2000), 782–786.

[Calvete00] Calvete, J., Sabater, L., Camps, B., Verdu, A., Gomez-Portilla, A., Martin, J., Torrico, M., Flor, B., Cassinello, N., and Lledo, S. Bile duct injury during laparoscopic cholecystectomy: myth or reality of the learning curve? *Surgical Endoscopy 14*, 7 (July 2000), 608–611.

[Carrihill85] Carrihill, B., and Hummel, R. Experiments with the intensity ratio depth sensor. *Computer Vision, Graphics, and Image Processing 32* (1985).

[Caspi96] Caspi, D., and Kiryari, N. Range imaging with adaptive color structured light. *Trans. PAMI 20*, 5 (1996).

[Catarci01] Catarci, M., Carlini, M., Gentileschi, P., and Santoro, E. Major and minor injuries during the creation of pneumoperitoneum. a multicenter study on 12,919 cases. *Surgical Endoscopy 15*, 6 (June 2001), 566–569.

[Chandler01] Chandler, J., Corson, S., and Way, L. Three spectra of laparoscopic entry access injuries. *Journal of the American College of Surgeons 192*, 4 (April 2001), 478–491.

[Chiuso02] Chiuso, A., Favaro, P., Jin, H., and Soatto, S. Structure from motion causally integrated over time. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 4 (April 2002).

[COMPASS] Compass international, inc. Corporate Website. http://www.compass.com.

[CompMo] Computer Motion, Inc. Corporate Website. http://www.computermotion.com.

[Corbitt91] Corbitt Jr., J. Laparoscopic cholecystectomy: laser versus electrosurgery. *Surgical Laparoscopy and Endoscopy 1*, 2 (June 1991), 85–88.

[CR90] CR, C. V., Meena, A., Petro, A., Haick, A., and Koury, A. Electrocautery is superior to laser for laparoscopic cholecystectomy. *American Journal of Surgery 160*, 5 (Novemeber 1990), 457.

[Cundiff01] Cundiff, G., Weidner, A., and Visco, A. Effectiveness of laparoscopic cadaveric dissection in enhancing resident comprehension of pelvic anatomy. *Journal of the American College of Surgeons 192*, 4 (April 2001), 492–497.

[Curet00] Curet, M. J. Special problems in laparoscopic surgery. previous abdominal surgery, obesity, and pregnancy. *Surgical Clinics of North America 80*, 4 (August 2000), 1093–1110.

[Däuber02] Däuber, S., Hoppe, H., Kreimpien, R., Hassfeld, S., Brief, J., and Wörn, H. Intraoperative guidance of pre-planned bone deformantions with a surface scanning system. *Studies in Health Technology and Informatics 85* (2002), 110–115.

[Davies96] Davies, C., and Nixon, M. Sensing surface discontinuities via coloured spots. *Proc. IWISP* (1996).

[Derossis98] Derossis, A., Freid, G., Abrahamowicz, M., Sigman, H., Barker, J., and Meakins, J. Development of a model for training and evaluation of laparoscopic skils. *American Journal of Surgery 175*, 6 (June 1998), 482–487.

[Derossis99] Derossis, A., Antoniuk, M., and Fried, G. Evaluation of laparoscopic skills: a 2-year follow-up during residency trainingr. *Canadian Journal of Surgery 42*, 4 (August 1999), 293–296.

[Drascic96] Drascic, D. Perceptual issues in augmented reality. In *SPIE Volume 2653: Stereoscopic Displays and Virtual Reality Systems III* (1996), pp. 123–124.

[Durrani95] Durrani, A. F., and Preminger, G. M. Three-dimensional video imaging for endoscopic surgery. *Computers in Biology and Medicine 25*, 2 (1995), 237–247.

[Edwards95] Edwards, P. J., Hawkes, D. J., Hill, D. L., Jewell, D., Spink, R., Strong, A., and Gleeson, M. Augmentation of reality using an operating microscope for otolaryngology and neurosurgical guidance. *Journal of Image Guided Surgery 1*, 3 (1995), 172–178.

[Edwards99] Edwards, P. J., King, A. P., Hawkes, D. J., Fleig, O., Maurer Jr., C. R., Hill, D. L., Fenlon, M. R., de Cunha, D. A., Gaston, R. P., Chandra, S., Mannss, J., Strong, A. J., Gleeson, M. J., and Cox, T. C. Stereo augmented reality in the surgical microscope. *Studies in Health Technology and Informatics 62* (1999), 102–108.

[Edwards00] Edwards, P. J., King, A. P., Maurer Jr., C. R., de Cunha, D. A., Hawkes, D. J., Hill, D. L., Gaston, R. P., Jusczyzck, A., Strong, A. J., Chandler, C. L., and Gleeson, M. J. Design and evaluation of a system for micro-scope-assisted guided interventions (MAGI). *IEEE Transactions on Medical Imaging 19*, 11 (November 2000), 1082–1093.

[Emam02] Emam, T., Hanna, G., and Cuschieri, A. Ergonomic principles of task alignment, visual display, and direction of execution of laparoscopic bowel suturing. *Surgical Endoscopy 16*, 2 (February 2002), 267–271.

[Eyal01] Eyal, R., and Tendick, F. Spatial ability and learning the use of an angled laparoscope in a virtual environment. *Studies in Health Technology and Informatics 81* (2001), 146–152.

[Fitzpatrick01] Fitzpatrick, C., Kolesari, G., and Brasel, K. Teaching anatomy with surgeons' tools: use of the laparoscope in clinical anatomy. *Clinical Anatomy 14*, 5 (September 2001), 349–353.

[Fleig01] Fleig, O. J., Devernay, F., Scarabin, J.-M., and Jannin, P. Surface reconstruction of the surgical field from stereoscopic microscope views in neurosurgery. In *Computer Assisted Radiology and Surgery* (2001), H. Lembke, M. Vannier, K. Inamura, A. Farman, and K. Doi, Eds., Elsevier, pp. 259–264.

[Fuchs98] Fuchs, H., Livingston, M. A., Raskar, R., Colucci, D., Keller, K., State, A., Crawford, J. R., Rademacher, P., Drake, S. H., and Meyer, A. A. Augmented reality visualization for laparoscopic surgery. In *First International Conference on Medical Image Computing and Computer-Assisted Intervention* (1998).

[Fuchs99] Fuchs, H., and Ackerman, J. Displays for augmented reality: Historical remarks and future prospects. In *Mixed Reality – Merging Real and Virtual Worlds* (1999), pp. 31–40.

[Galizia01] Galizia, G., Prizio, G., Lieto, E., Castellano, P., Pelosio, L., Imperatore, V., Ferrara, A., and Pignatelli, C. Hemodynamic and pulmonary changes during open, carbon dioxide pneumoperitoneum and abdominal wall-lifting cholecystectomy. a prospective, randomized study. *Surgical Endoscopy 15*, 5 (May 2001), 477–483.

[Gallagher99] Gallagher, A., McClure, N., McGuigan, J., Crothers, I., and Browning, J. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer in vritual reality (MIST VR). *Endoscopy 31*, 4 (May 1999), 3100–3103.

[Gao01] Gao, H. H. C., Biocca, F., and Rolland, J. P. An ultra-light and compact design and implementation of head-mounted projective displays. *VR* (2001), 175–182.

[Gartner96] Gartner, H., Lehle, P., and Tiziani, H. New, highly efficent, binary codes for structured light methods. *SPIE 2599* (1996).

[Gering01] Gering, D., Nabavi, A., Kikinis, R., Hata, N., O'Donnell, L., Grimson, W., Jolesz, F., Black, P., and Wells III,

W. An integrated visualization system for surgical planning and guidance using image fusion and an open MR. *Journal Magnetic Resonance Imaging 13*, 6 (June 2001), 967–975.

[Greenberg01] Greenberg, D., Peiser, J. G., Peterburg, Y., and Pliskin, J. S. Reimbursement policies, incentives and disincentives to perfrom laparoscopic surgery in israel. *Health Policy 56* (2001), 49–63.

[Grimson99] Grimson, W., Kikinis, R., Jolesz, F., and Black, P. Image-guided surgery. *Scientific American 280*, 6 (June 1999), 62–69.

[Hajri00] Hajri, A., Mutter, D., Wack, S., Bastien, C., Gury, J. F., Marescaux, J., and Aprahamian, M. Dual effect of laparoscopy on cell-mediated immunity. *Eur Surg Res 32*, 5 (2000), 261–266.

[Hall-Holt01] Hall-Holt, O., and Rusinkiewicz, S. Stripe boundary codes for real-time structured-light range scanning of moving objects. *Proceedings of ICCV* (2001).

[Hanna01] Hanna, G., Elamass, M., and Cuschieri, A. Ergonomics of hand-assisted laparoscopic surgery. *Seminars in Laparoscopic Surgery 8*, 2 (June 2001), 92–95.

[Hasson74] Hasson, H. Open laparoscopy: a report of 150 cases. *Journal of Reproductive Medicine 12*, 6 (June 1974), 234–238.

[Herron99] Herron, D. M., Lantis II, J. C., Mayket, J., Basu, C., and Schwaitzberg, S. D. The 3-D monitor and the head-mounted display: a quantitative evaluation of advanced laparoscopic viewing technologies. *Surgical Enoscopy 13* (1999), 751–755.

[Holloway95] Holloway, R. L. *Registration Errors in Augmented Reality Systems*. PhD thesis, University of North Carolina, 1995.

[Hopkins00] Hopkins, M. The myths of laparoscopic surgery. *American Journal of Obstetrics and Gynecology 183*, 1 (July 2000), 1–5.

[Hoppe01a] Hoppe, H., Brief, J., Däuber, S., Hassfeld, S., Raczkowsky, J., and Wörn, H. *Projector-based augmented reality in cranio-maxillo-facial surgery*. Elsevier Press, Berlin, 2001, p. 1153.

[Hoppe01b] Hoppe, H., Dauber, S., Raczkowsky, J., Worn, H., and Moctezuma, J. L. Intraoperative visualization of surgical planning data using video projectors. *Studies in health technology and informatics 81* (2001), 206–208.

[Horn99] Horn, E., and Hiryati, N. Toward optimal structured light patterns. *Image and Vision Computing 17* (1999).

[Hornbeck88] Hornbeck, L. J., and Nelson, W. E. Bistable deformable mirror device. *OSA Technical Digest: Spatial Light Modulators and Applications 8* (1988).

[Hua00] Hua, H., Girardot, A., Gao, C., and Rolland, J. P. Engineering of head-mounted projective displays. *Applied Optics 39*, 22 (2000), 3814–3824.

[Intuitive] Intuitive Surgical, Inc. Corporate Website. http://www.intuitivesurgical.com.

[Jannin00] Jannin, P., Fleig, O. J., Seigneuret, E., Morandi, X., Raimbault, M., and Scarabin, J.-M. Multimodal and multifunctional neuronavigation. In *Computer Assisted Radiology and Surgery* (2000), pp. 167–172.

[Johnston96] Johnston, R., Bhoyrul, S., Way, L., Satava, R., McGovern, K., Fletcher, J. D., Rangel, S., and Loftin, R. B. Assessing a virtual reality surgical skills simulator. In *Health Care in the Information Age: Future Tools for Transforming Medicine* (1996), S. Weghorst, H. Sieburg, and K. Morgan, Eds., IOS Press, pp. 608–617.

[Jonsson00] Jonsson, B., and Zethraeus, N. Costs and benefits of laparoscopic surgery – a review of the literature. *European Journal of Surgery Supplement 585* (2000), 48–56.

[Kalawsky93] Kalawsky, R. *The science of virtual reality and virtual environments*. Addison-Wesley, 1993.

[Kanbara00] Kanbara, M., Okuma, T., Takemura, H., and Yokoya, N. A stereoscopic video see-through augmented reality system based on real-time vision-based registration. In *Proceedings of Virtual Reality 2000* (March 2000), pp. 255–262.

[Kappert00] Kappert, U., Schneider, J., Cichon, R., Gulielmos, V., Schade, I., Nicolai, J., and Schueler, S. Closed chest totally endoscopic coronary artery bypass surgery: fantasy or reality? *Current Cardiology Reports 2*, 6 (November 2000), 558–563.

[Kelly00] Kelly, P. J. Stereotactic surgery: what is past is prologue. *Neurosurgery 46*, 1 (January 2000), 16–27.

[Knowlton75] Knowlton, K. Virtual pushbuttons as a means of person-machine interaction. In *Proceedings of the Conference on Computer Graphics, Pattern Recognition and Data Structure* (May 1975), pp. 14–16.

[Knowlton77] Knowlton, K. Computer displays optically superimposed on input devices. *Bell Technical Journal 56*, 3 (1977), 367–83.

[Kurian01] Kurian, M., Patterson, E., Andrei, V., and Edye, M. Hand-assisted laparoscopic surgery: and emerging technique. *Surgical Endoscopy 11*, 15 (November 2001), 1277–1281.

[Lecuru01] Lecuru, F., Leonard, F., Philippe Jais, J., Rizk, E., Robin, F., and Taurelle, R. Laparoscopy in patients with prior surgery: results of the blind approach. *JSLS 5*, 1 (January-March 2001), 13–16.

[Lee02] Lee, J., Hirota, G., and State, A. Modeling real objects using video see-through augmented reality. *Presence: Teleoperators and Virtual Environments 11*, 2 (2002), 144–157.

[Litynski96] Litynski, G. S. *Highlights In The History Of Laparoscopy.* Barbara Bernert Verlag, Frankfurt/Main, Germany, 1996.

[Liu00] Liu, D. C., Meyers, M. O., Hill, C. B., and Loe Jr., W. A. Laparoscopic splenectomy in children with hematological disorders: preliminary experience at the children's hospital of new orleans. *American Surgery 66*, 12 (December 2000), 1168–1170.

[Long01] Long, K., Bannon, M., Zietlow, S., Helgeson, E., Harmsen, W., Smith, C., Ilstrup, D., Baerga-Varela, Y., and Sarr, M. A prospective randomized comparision of laparoscopic appendectomy with open appendectomy: Clinical and economic analyses. *Surgery 129*, 4 (April 2001), 390–400.

[Lukban00] Lukban, J., Jaeger, J., Hammond, K., LoBraico, D., Gordon, A., and Graebe, R. Gasless versus conventional laparoscopy. *New Jersey Medicine 97*, 5 (May 2000), 29–34.

[Marescaux01] Marescaux, J., Gagner, M., Rubino, F., Mutter, D., Vix, M., Butner, S., and Smith, M. Transatlantic robot-assisted telesurgery. *Nature 413* (September 27 2001), 379–380.

[Matsunaga00] Matsunaga, K., Yamamoto, T., Shidoji, K., and Matsuki, Y. The effect of the ratio difference of overlapped areas of stereoscopic images on each eye in teleoperation. In *SPIE Volume 3957, Stereoscopic Displays and Virtual Reality Systems VII* (2000), pp. 236–243.

[Mayol97] Mayol, J., Garcia-Aguilar, J., Ortiz-Oshiro, E., and De-Diego Carmona, J. A. Risks of the minimal access approach for laparoscopic surgery: multivariate analysis of morbidity related to umbilical trocar insertion. *World Journal of Surgery 21*, 5 (June 1997), 529–533.

[Melzer95] Melzer, A., Riek, S., Roth, K., and Buess, G. Endoscopically controlled trocar and cannula insertion. *Endoscopic Surgery and Allied Technologies 3*, 1 (February 1995), 63–68.

[Merhoff00] Merhoff, A. M., Merhoff, G. C., and Morris E. Franlkin, J. Laparoscopic versus open appendectomy. *American Journal of Surgery 179* (2000), 375–378.

[Milgram92] Milgram, P., and Krüger, M. Adaptation effects due to online changes in camera configuration. In *SPIE Volume 1669, Stereoscopic Displays and Applications III* (1992), pp. 122–134.

[Milgram99] Milgram, P., and Colquhoun Jr., H. A taxonomy of real and virtual world display integration. In *Mixed Reality – Merging Real and Virtual Worlds* (1999), pp. 5–30.

[Minkes00] Minkes, R. K., Lagzdins, M., and Langer, J. C. Laparoscopic versus open splenectomy in children. *Pediatric Surgery 35*, 5 (May 2000), 6999–701.

[Mueller99] Mueller, M., Camartin, C., Dreher, E., and Hänggi, W. Three-dimensional laparoscopy: Gadget or progress? a randomized trial on the efficacy of three-dimensional laparoscopy. *Surgical Endoscopy 13* (1999), 469–472.

[Nayar96] Nayar, S. K., Watanabe, M., and Noguchi, M. Real-time focus range sensor. *Trans. PAMI 18*, 12 (1996).

[Nguyen01] Nguyen, N., Goldman, C., Rosenquist, C., Arango, A., Cole, C., Lee, S., and Wolfe, B. Laparoscopic versus open gastric bypass: a randomized study of outcomes, quality of life, and costs. *Annals of Surgery 234*, 3 (September 2001), 279–91.

[Nicholson01] Nicholson, T., and Tiruchelvam, V. Comparision of laparoscopic-assisted appendectomy with intracorporal laparoscopic appendectomy and open appendectomy. *Journal of the Soceity of Laparoendoscopic Surgeons 5*, 1 (Jan-Mar 2001), 47–51.

[Nishiguchi01] Nishiguchi, K., Okuda, J., Toyoda, M., Tanaka, K., and Tanigawa, N. Comparative evaluation of surgical stress of laparoscopic and open surgeries for colorectal carcinoma. *Dis Colon Rectum 44*, 2 (February 2001), 223–230.

[O'Boyle01] O'Boyle, C., Royston, C., and Sedman, P. Cost-utility analysis of open versus laparoscopic groin hernia repair: results from a multicentre randomized clinical trial. *British Journal of Surgery 88* (2001), 653–661.

[Ohshima98] Ohshima, K. Ar2hockey: A case study of collaborative augmented reality. In *Proceedings of VRAIS'98* (1998), pp. 268–295.

[O'Malley01] O'Malley, C., and Cunnningham, A. Physiologic changes during laparoscopy. *Anesthesiol Clin North America 19*, 1 (2001), 1–19.

[Park01] Park, A., Schwartz, R., Witzke, D., Roth, J., Mastrangelo, M., Birch, D., Jennings, C., Lee, E., and Hoskins, J. A pilot study of new approaches to teaching anatomy and pathology. *Surgical Endoscopy 15*, 3 (March 2001), 245–250.

[Passmore00] Passmore, P., Read, O., Nielsen, C., Torkington, J., and Darzi, A. Effects of perspective and stereo on depth judgements in virtual reality laparoscopy simulation. In *Medicine Meets Virtual Reality 2000* (2000), J. Westwood et al., Eds., IOS Press, pp. 243–245.

[Polymeneas01] Polymeneas, G., Theodosopoulos, T., Stamatiadis, A., and Kourias, E. A comparative study of postoperative adhesion formation after laparoscopic vs open cholecystectomy. *Surgical Endoscopy 15*, 1 (June 2001), 41–43.

[Posdamer82] Posdamer, J. L., and Altschuler, M. D. Surface measurement by space-encoded projected beam systems. *Computer Graphics and Image Processing 18* (1982).

[Poulin01] Poulin, E. C., Schlachta, C. M., Seshadri, P. A., Cadeddu, M. O., Gregoire, R., and Mamazza, J. Septic complications of elective laparoscopic colorectal resection. *Surgical Endoscopy 15*, 2 (February 2001), 203–208.

[Proesmans96] Proesmans, M., Van Gool, L., and Oosterlinck, A. One-shot active 3d shape acquisition. *Proc. ICPR* (1996).

[Raskar98] Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., and Fuchs, H. The office of the future : A unified approach to image-based modeling and spatially immersive displays. In *SIGGraph '98 Conference Proceedings* (1998), ACM SIGGraph, pp. 179–188.

[Rathert74] Rathert, P., Lutzeyer, W., and Goddwin, W. E. Philipp bozzini (1773-1809) and the Lichtleiter. *Urology 3*, 1 (January 1974), 113–188.

[Rioux94] Rioux, M. Digital 3-d imaging theory and applications. *SPIE 2350* (1994).

[Robinett92] Robinett, W., and Rolland, J. P. A computational model for the stereoscopic optics of a head-mounted display. *Presence: Teleoperators and Virtual Environments 1*, 1 (Winter 1992), 45–62.

[Rolland] Rolland, J., Davis, L., Hamza-Lup, F. G., Norfleet, J., Imielinska, C., and Kerner, K. F. *Merging augmented reality and anatomically correct 3D models in the development of a training tool for endotracheal intubation.* http://www.cs.ucf.edu/ fhamza/papers/AugmentedReality_Hamza-Lup.pdf.

[Rolland92] Rolland, J., and Gibson, W. Towards quantifying depth and size perception in virtual envirnments. *Presence: Teleoperators and Virtual Environments 4*, 1 (Winter 1992), 24–49.

[Rolland97] Rolland, J. P., Wright, D. L., and Kancherla, A. R. Towards a novel augmented-reality tool to visualizae dynamic 3-D anatomy. *Studies in health technology and informatics 39* (1997), 337–348.

[Rosenthal01] Rosenthal, M., State, A., Lee, J., Hirota, G., Ackerman, J., Keller, K., Pisano, E. D., Jiroutek, M., Muller, K., and Fuchs, H. Augmented reality guidance for needle biopsies: A randomized, controlled trial in phantoms. *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (2001).

[Rosser Jr.98] Rosser Jr., J., Rosser, L., and Savalgi, R. Objective evaluation of a laparoscpoic surgical skill program for residents and senior surgeons. *Archives of Surgery 133*, 8 (August 1998), 911–912.

[Rosser97] Rosser, J., Rosser, L., and Savalgi, R. Skill acquisition and assessment for laparoscopic surgery. *Archives of Surgery 132*, 2 (February 1997), 200–204.

[Rousu98] Rousu, J. S., Kohls, P. E., Kall, B., and Kelly, P. J. Computer-assisted image-guided surgery using the Regulus Navigator. *Studies in Health Technology and Informatics 50* (1998), 103–109.

[Rusinkiewiczi01] Rusinkiewiczi, S. *Real-time Acquisition and Rendering of Large 3D Models*. PhD thesis, Stanford University, August 2001.

[Santala99] Santala, M., Jarvela, I., and Kauppila, A. Transfundal insertion of a veress needle in laparoscopy of obese subjects: a practical alternative. *Human Reproduction 14*, 9 (September 1999), 2277–2278.

[Sato87] Sato, K., and Inokuchi, S. Range-imaging system utilizing nematic liquid crystal mask. *Proc. ICCB* (1987).

[Sato98] Sato, Y., Nakamoto, M., Tamaki, Y., Sasama, T., Sakita, I., Nakajima, Y., Monden, M., and Tamura, S. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging 5*, 17 (October 1998), 681–693.

[Schafer01] Schafer, M., Lauper, M., and Krahenbuhl, L. Trocar and Veress needle injuries during laparoscopy. *Surgical Endoscopy 15*, 3 (March 2001), 275–280.

[Scott01] Scott, D., Young, W., Tesfay, S., Frawley, W., Rege, R., and Jones, D. Laparoscopic skills training. *The American Journal of Surgery 82* (2001), 137–142.

[Semm95] Semm, K. The history of endoscopy. In *Laparoscopic surgery: an atlas for general surgeons*, G. C. Vitale, J. S. Sanfilippo, and J. Perissat, Eds. J.B. Lippincott Company, Philadelphia, 1995, pp. 3–11.

[Shennib98] Shennib, H., Bastawisy, A., Mack, M., and Moll, F. Computer-assisted telemanipulation: an enabling technology for endoscopic coronary artery bypass. *Annals of Thoracic Surgery 66*, 3 (September 1998), 1060–1063.

[Smith01] Smith, C., Farrell, T., McNatt, S., and Metreveli, R. Assessing laparoscopic manipulative skills. *The American Journal of Surgery 181* (2001), 547–550.

[Spaner97] Spaner, S. J., and Warnock, G. L. A brief history of endoscopy, laparoscopy, and laparoscopic surgery. *Journal of Laparoendoscpic & Advanced Surgical Techniques 7*, 6 (December 1997), 369–373.

[State96a] State, A. Technologies for augmented-reality systems: realizing ultrasound-guided needle biopsies. In *SIGGraph '96 Conference Proceedings* (1996), ACM SIGGraph, pp. 439–446.

[State96b] State, A., Livingston, M. A., Hirota, G., Garrett, W. F., Whitton, M. C., Fuchs, H., and Pisano, E. D. Technologies for augmented-reality systems: realizing ultrasound-guided needle biopsies. In *Proceedings of SIGGRAPH'96 (New Orleans, LA, August 4-9, 1996). In Computer Graphics Proceedings, Annual Conference Series* (1996), ACM SIGGRAPH, pp. 439–446.

[State01] State, A., Ackerman, J., Hirota, G., Lee, J., and Fuchs, H. Dynamic virtual convergence for video see-through head-mounted displays: Maintaining maximum stereo overlap throughout a close-range work space. *Proceedings of the International Symposium on Augmented Reality (ISAR)* (2001), 137–146.

[Stetten00] Stetten, G. D., Chib, V. S., and Tamburo, R. J. Tomographic reflection to merge ultrasound images with direct vision. *IEEE Proceedings of the Applied Imagery Pattern Recognition* (2000), 200–205.

[Stocchi00] Stocchi, L., Nelson, H., Young-Fadok, T., Larson, D., and Ilstrup, D. Safety and advantages of laparoscopic vs. open colectomy in the elderly: matched-control study. *Diseases of the colon and rectum 43*, 3 (March 2000), 326–332.

[Sutherland65] Sutherland, I. The ultimate display. In *Proceedings of IFIP Congress vol. 2* (1965), pp. 506–508.

[Sutherland68] Sutherland, I. A head-mounted three dimensional display. In *Proceedings, Fall Joint Computer Conference, AFIPS Conference Proceedings vol 33* (1968), pp. 757–764.

[Takagi00] Takagi, A., Yamazaki, S., Saito, Y., and Taniguchi, N. Development of a stereo video see-through hmd for ar systems. In *Proceedings of the International Symposium on Augmented Reality (ISAR)* (2000), pp. 69–77.

[Tang98] Tang, S.-L., Kwah, C.-K., Teo, M.-Y., Sing, N. W., and Ling, K.-V. Augmented reality systems for medical applications. *IEEE Engineering In Medicine and Biology* (May/June 1998), 49–58.

[Tuech00] Tuech, J. J., Pessaux, P., Rouge, C., Regenet, N., Bergamaschi, R., and Arnaud, J. P. Laparoscopic vs open colectomy for sigmoid diverticulitis: a prospective comparative study in the elderly. *Surgical Endoscopy 14*, 11 (November 2000), 1031–1033.

[Underwood97] UNDERWOOD, I. Liquid crystal over silicon spatial light modulators - principles, practice and prospects. In *OSA Trends in Optics and Photonics* (Washington DC, 1997), G. Burdge and S. C. Esener, Eds., vol. 14, Spatial Light Modulators, Optical Society of America, pp. 76–88.

[Underwood00] UNDERWOOD, I. A review of microdisplay technologies. In *Digest of SID@EID (The Annual Conference of the UK and Ireland Chapter of the SID)* (London, November 2000).

[van Veelen01] VAN VEELEN, M., MEIJER, D., GOOSSENS, R., AND SNIJDERS, C. New ergonomic design criteria for handles of laparoscopic dissection forceps. *Journal of Laparoendoscopic and Advanced Surgical Techniques 11*, 1 (February 2001), 17–26.

[Veress38] VERESS, J. Neues Instument zur Ausfuhrung von Brust-oder Bauch-punktinonen und Pneumothoraxbehandlung. *Disch Med Wochenshr 41* (1938), 1480–1481.

[Viola95] VIOLA, P., AND WELLS, W. Alignment by maximization of mutual information. In *International Conference on Computer Vision* (Boston, MA, 1995).

[Voorhorst97] VOORHORST, F., OVERBEEKE, C., AND SMETS, G. Spatial perception during laparoscopy: implementing action-perception coupling. *Studies in Health Technology and Informatics 39* (1997), 379–386.

[Voorhorst99] VOORHORST, F., MEIJER, D., AND OVERBEEKE, C. Head-controlled laparoscopy: experiment, prototype, and preliminary results. *Journal of Laparoendoscopic and Advanced Surgical Techiques, Part A 9*, 5 (October 1999), 379–388.

[Ware98] WARE, C., GOBRECT, C., AND PATON, M. Dynamic adjustment of stereo display parameters. *IEEE Transactions on Systems, Man, and Cybernetics 28*, 1 (1998), 56–65.

[Watson95] WATSON, B. A., AND HODGES, L. F. Using texture maps to correct for optical distortion in head-mounted displays. In *Proceedings of the Virtual Reality Annual Symposium '95* (1995), IEEE Computer Society Press, pp. 172–178.

[Weidenbach00] WEIDENBACH, M., WICK, C., PIEPER, S., QUAST, K., FOX, T., GRUNST, G., AND REDEL, D. Augmented reality simulator for training in two-dimensional echocardiography. *Computrtd and Biomedical Research 33*, 1 (February 2000), 11–22.

[Welch01] Welch, G., Bisphop, G., Vicci, L., Keller, K., and Colucci, D. High-performance widea-area optical tracking: The HiBall tracking system. *Presence: Teleoperators and Virtual Environments 10*, 1 (2001), 1–21.