REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 1996	3. REPORT TYPE AND DATES COVERED technical

4. TITLE AND SUBTITLE

Developing a New Medical Augmented Reality System

5. FUNDING NUMBERS

6. AUTHOR(S)

Frederick Morgan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

8. PERFORMING ORGANIZATION REPORT NUMBER

CMU-RI-TR-96-19

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release;
Distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Augmented reality is a technique for combining spplementary imagery such that it appears as part of the scene and can be used for guidance, training, and locational aids. In the medical domain, augmented reality can be used to combine medical imagery to the physician's view of a patient to help the physician establish a direct relation between the imagery and the patient. This project report will examine medical augmented reality systems for use in a surgical setting. Four areas will be examined: (1) applications for augmented reality system in medicine, (2) survey of basic technologies used for building augmented reality systems, (3) the development of a new augmented reality system, and (4) testing and validating an augmented reality system for clinical use. The goal of this project report is to develop a new design of a medical augmented reality system; the design will draw upon a number of different technologies in an attempt to build a more practical and capable system.

14. SUBJECT TERMS

15. NUMBER OF PAGES

32 pp.

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unlimited	unlimited	unlimited	unlimited

Developing a New Medical Augmented Reality System

Frederick Morgan
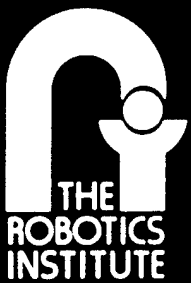
CMU-RI-TR-96-19

Carnegie Mellon University

The Robotics Institute

Technical Report

Developing a New Medical Augmented Reality System

Frederick Morgan

CMU-RI-TR-96-19

The Robotics Institute

Carnegie Mellon University

5000 Forbes Avenue

Pittsburgh, Pennsylvania 15213

15 May 1996

Acknowledgments

I would like to express my gratitude to:

Mike Blackwell, David Simon, Bob O'Toole,
Branko Jaramaz, Lori Gregor, Drew Shefman,
Takeo Kanade, Langham Gleason and
Debbie Morgan

Abstract

Augmented reality is a technique for combining supplemental imagery such that it appears as part of the scene and can be used for guidance, training, and locational aids. In the medical domain, augmented reality can be used to combine medical imagery to the physician's view of a patient to help the physician establish a direct relation between the imagery and the patient. This project report will examine medical augmented reality systems for use in a surgical setting. Four areas will be examined: (1) applications for augmented reality system in medicine, (2) survey of basic technologies used for building augmented reality systems including the current state of the art in medical augmented reality systems, (3) the development of a new augmented reality system and (4) testing and validating an augmented reality system for clinical use. The goal of this project report is to develop a new design of a medical augmented reality system; the design will draw upon a number of different technologies in an attempt to build a more practical and capable system.

Key Words: Augmented Reality, Computer Assisted Surgery, Registration, Tracking, Display Devices, Image Overlay, Surgical Guidance, Telemedicine.

Table of Contents

List of Figures

List of Tables

1 Augmented Reality

Augmented reality is a technique for combining supplemental imagery that appears to be part of the scene used for guidance, training, and locational aids to the user. The supplemental imagery can be as simple as arrows offering directional guidance to the user, or as complicated as a three dimensional model of the scene. This project report will examine a number of different techniques for building, testing and validating a medical augmented reality system for surgical use.

The medical domain can benefit greatly from augmented reality. Advanced imaging techniques, such as Magnetic Resonance Imaging (MRI), or Computerized Axial Tomography (CT or CAT), offer a non-invasive look inside a patient. These imaging techniques can be combined with augmented reality to give a physician the ability look beneath the surface of their patient's skin to the anatomical structures below. In addition, the imagery can be used to offer supplemental localization, guidance and training aides during surgery. Augmented reality has the potential of greatly reducing surgical times, while improving overall patient outcomes. While there are many used for augmented reality in medicine, this project report will focus on augmented reality system for use in a surgical setting.

1.1 What is Augmented Reality?

In the realm of medicine, augmented reality is a display technique that offers guidance and assistance to the physician in locating anatomical structures on the patient. Augmented reality combines supplemental imagery to the physicians view of the patient (Figure 1-1). Traditional methods used by physicians to view medical imagery include: films on light tables, imagery on a workstation monitor, etc. However, a major drawback with traditional approaches is the lack of a direct relationship between the patient and the imagery. When the imagery is used to locate anatomical structures during surgery, the surgeon views the structure in the imagery and then tries to locate it on the patient. To aid in the localization of the anatomical structures, the surgeon is often required to mentally transform the imagery to that of the same orientation and position as the patient.

Augmented reality provides to the physician a direct spatial relationship between imagery and patient. Through the use of augmented reality, a physician can view the imagery while at the same time viewing the patient. The imagery is overlaid on the patient to appear in the exact orientation and position as the corresponding anatomical structures. The ability to view the imagery in the correct position and orientation while viewing the patient allows the physician to visualize internal structures in proper position over the patient.

The direct relationship between the imagery and the patient can be beneficial to a surgeon. During a surgical procedure it is often necessary to locate structures on the patient that have been viewed in the imagery. Without the use of augmented reality, a physician may have a difficult time of locating various subcutaneous structures. For example, in neurosurgery, when a surgeon locates a brain lesion (tumor) in a CT or MRI scan, they might have a difficult time visualizing the exact location of the tumor in the patient before surgery. A substantial amount of time is spent localizing the tumor on the patient using a traditional method of viewing the imagery, such as films viewed on light tables [9]. Often, the lack of exact localization will cause a surgeon to make the initial incision on a patient's head significantly larger then the actual size required. With the use of augmented reality, the exact tumor location could be projected from the imagery directly over the patient, thus allowing the size of the incision and the time needed to plan the incision's location to be reduced.

1

(a)   (b)
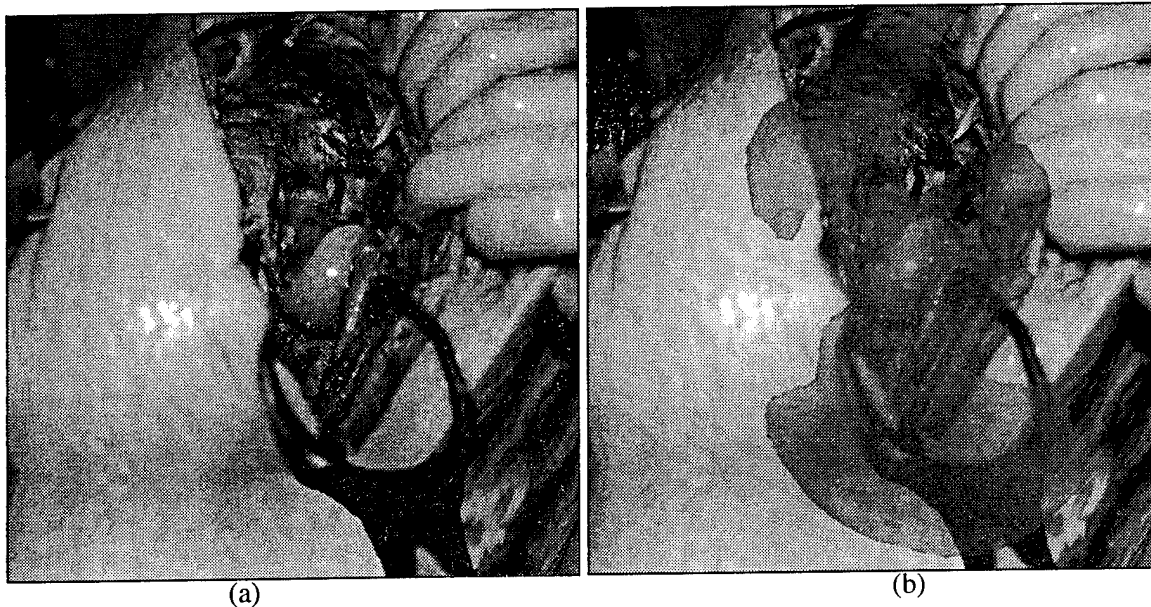
Figure 1-1: (a) a surgical exposure of a femur (leg bone) and hip. (b) The same surgical exposure this time, with anatomical reconstructions of the bones overlaid on the view of the exposure to simulate the view a surgeon would see through an augmented reality system.

1.2 Potential Uses for Augmented Reality in Medicine

Virtually any visual imaging modality can be used with augmented reality. Some of the more commonly used imaging techniques in addition to MRI and CT are: Magnetic Resonance Angiography (MRA), Digital Subtractive Angiography (DSA), and Ultrasound (US) [26]. All of these techniques produce multiple planar imagery slices that are spatially combined to create a volumetric data set. With the volumetric data set, it is possible to reconstruct three dimensional models of the anatomy being studied. In neurosurgery, for example, a tumor seen in a volumetric MRI scan can be reconstructed into a three dimensional model to allow the surgeon to visualize the entire tumor volume from any direction. By using a wide variety of visual imagery and 3D reconstructions, augmented reality can be applied to a number of different applications including surgical guidance, localization of anatomical structures, and surgical training.

Augmented reality offers a surgeon the ability to view pre-operative guidance cues to reach a desired surgical goal. Pre-operatively planning a procedure, followed by the execution of this plan in the operating room, allows for a better surgical outcome. For example, a surgeon can use imagery to pre-operatively define an approach to a lesion located deep within a patient's brain such that it avoids critical blood vessels and brain tissue. Intra-operatively, the surgeon can follow the pre-operative plan presented by an augmented reality system. Such an approach allows the surgeon to visualize and avoid critical structures identifiable in the imagery that are located along the path to the lesion, thus potentially reducing trauma to the critical structures.

In addition to pre-operative planning and intra-operative execution, augmented reality provides the means for a remote expert surgeon to offer guidance and consultations on a case (Figure 1-2) [13]. For example, a surgeon in a small rural town in central Pennsylvania would like the assistance of an expert surgeon in removing a lesion from a patient's brain. A remote expert surgeon in Pittsburgh is called and asked to assist. The remote expert views the imagery from the case and provides navigational aids to the rural surgeon to

help guide them to the correct surgical sight during surgery. The rural surgeon can follow these aids presented by an augmented reality system. Without the remote assistance of the expert surgeon, the patient would have to be flown to the expert's hospital for the case to be performed. However, if an expert can assist on the case remotely by using an augmented reality system, the cost of transporting the patient can be saved by performing the case in the patient's local hospital [2].
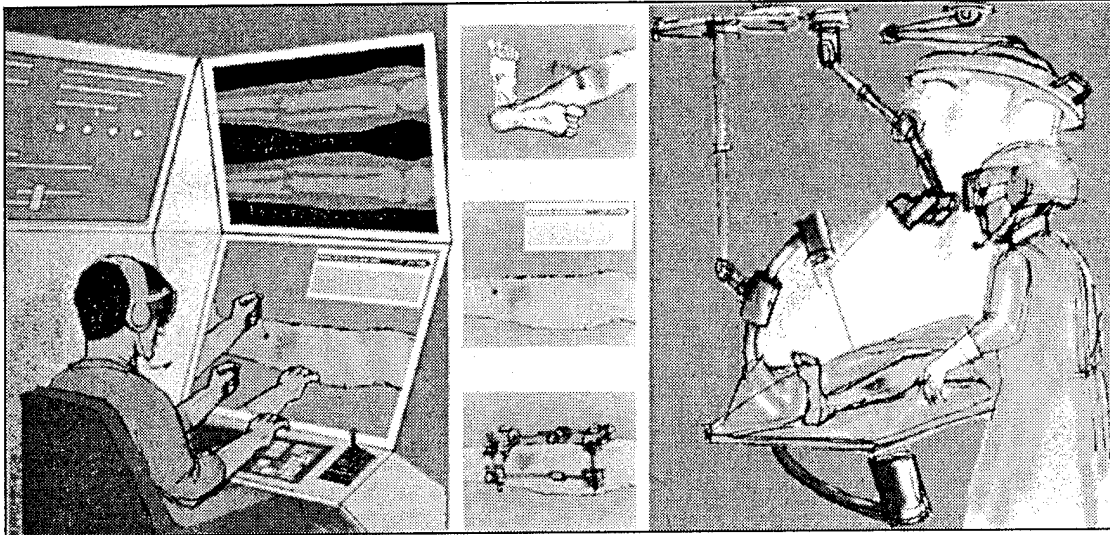
© 1996 Branko Jaramaz

Figure 1-2: An example of an expert surgeon (left) remotely assisting a surgeon on a case (right).

Another area in which augmented reality is being used in medicine is surgical training. For example, when a new student is learning a procedure, they often have a difficult time discriminating between areas that are safe to work in, and ones that should be avoided. Through the use of augmented reality, these regions can be identified. An expert surgeon can define safe and dangerous regions that are visually presented to the student during surgery with an augmented reality system. At the same time, the student's surgical instruments can be tracked to determine if they are located within the safe region. If the instruments are outside of the safe region, the system can notify the student of the potential problem. The ability to easily view the safe and dangerous regions during surgery can potentially aid the student in learning these boundaries faster than by traditional methods.

In all of the above augmented reality applications, it is critical that the overlaid images be accurately positioned in the scene. If the imagery being presented to the surgeon is not correctly positioned, inaccurate information will be conveyed. In the following chapters, the problem of providing accurate overlays to the surgeon will be addressed by examining a number of different technologies including: patient-imagery registration, display techniques and patient-surgeon tracking.

2 Augmented Reality Systems

All augmented reality systems have to solve some basic technological hurdles in order to be useful in a real world application. These technological hurdles include: image-patient registration, type of display devices and patient surgeon tracking. The success of the augmented reality systems relies on selecting the basic technologies that solve these hurdles in a way that make the system safe and easy to use. A number of different display techniques can be used when building augmented reality systems. A few of the state of the art medical augmented reality systems, each utilizing a different display technique, will be evaluated for ef-

fectiveness of conveying supplemental imagery to the user and their ease of use in a medical setting.

2.1 Basic Augmented Reality Technologies

2.1.1 Registration

The direct relation between imagery and the world[1] provided by an augmented reality system is lost if the overlaid imagery is not correctly aligned to the world. The process of aligning the overlaid imagery is called registration and is composed of several parts. First, the anatomical structure of interest is scanned, and a relation between the imagery and a 3-dimensional modal is created. A registration sensor is then used to collect data from the anatomical structure of interest to be used for the registration process (Figure 2-1). The registration process relates the data from the registration sensor to that of the model to find a transformation between the anatomy and the overlaid imagery. When the imagery is adjusted by this transformation, it becomes correctly aligned to the anatomy. Two of the most frequently used techniques to perform registration in the medical domain are fiducial based and shape (surface) based. These two techniques each have their own advantages and disadvantages.

Both the fiducial based and shape based registration methods work under very similar principles. Both techniques require points (X,Y, Z position) in the imagery and on the patient. The two point sets are matched to find the transformation that relates the point sets together. The matching process works by minimizing the distance between the point on the patient and a corresponding point in the imagery. One technique to find the minimal distance between these two data sets is to use a sum of squares algorithm Figure 2-1:. The sum of squares approach finds the unknown transformation, T, that minimizes the distance between the two data sets:

$$D = \sum_i [d_s (P_m, T \bullet P_p)]^2 \qquad (2.1)$$

where d_s is the distance between the patient point, P_p and the model point P_m [11].

Fiducial based and shape based registration differ from each other by the patients points (P_p) and model points (P_m) used in the registration process. For shape based registration, P_p can be any arbitrary point on the surface of the anatomical structure to be registered and P_m is the closest corresponding point on a three-dimensional reconstruction (produced from the medical imagery) of the anatomical structure. On the other hand, for fiducial based registration P_p is the location of a fiducial marker placed on the patient and P_m is the same fiducial marker located in the imagery data set. Fiducial based registration is constrained to the fiducial markers while shape based registration can utilize any arbitrary points on the patient.

Shape and fiducial based registration do, however, have some limitations. Shape based registration is very sensitive to the points collected. If points are collected on the patient such that the system is not constrained (limit the relative freedom of movement due to point ambiguity), and a limited number of points are used, the accuracy of registration will be low [23]. Fiducial based registration requires fiducial markers to be affixed to the patient prior to the imagery being collected. By requiring the fiducial markers to be present when the scan is performed, the initial diagnosis scan can not be used, instead another scan is required adding time and cost to the procedure. Once the scan has been completed, registration can be performed with as few as 3 markers if the fiducial markers remain stationary and are accurately localized in the imagery and on the patient.

1.World throughout this paper is referring to the direct scene of the patient as seen by the doctor, without any additional aides.
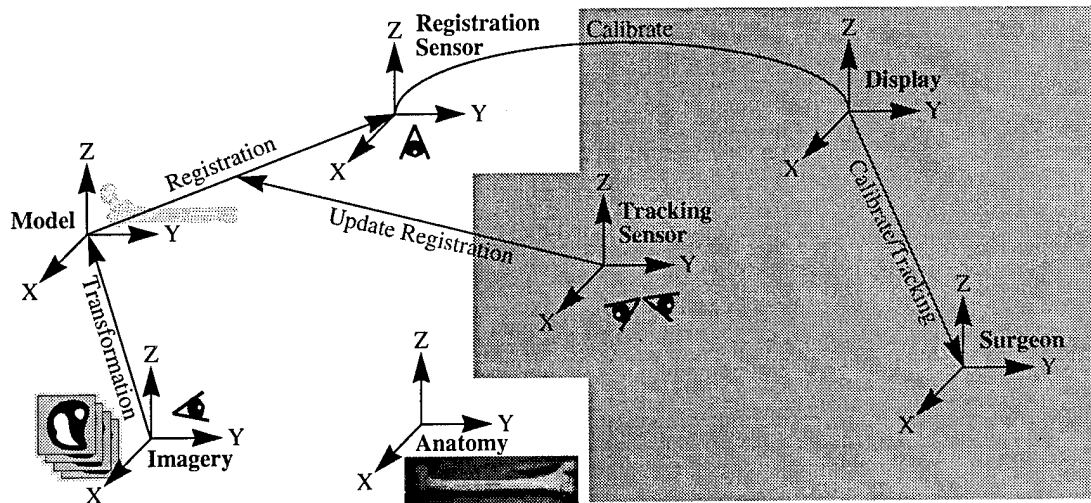
Figure 2-1: The Registration process. The anatomy is scanned producing the imagery in a coordinate system relative to the scanner. From the imagery a model is created along with a transformation that related the model coordinates to the imagery coordinate system. A registration sensor is then used to collected data from the anatomy. This data is used in the registration process. The registration process finds the relation between the model and the registration sensor to allow for the position and orientation of the overlay to be correct. The grayed out portion of the image is used for the display and tracking process of augmented reality.

Another method of performing registration, called physical templateing, is currently being investigated for use with augmented reality systems. Template based registration does not use any points collected from the imagery or the patient; instead, a template of the anatomical structure to be registered is created from the CT or MRI imagery [17]. The template is created by rapidly prototyping, via methods such as stereo lithography, the inverse of the 3D reconstruction made from the imagery (Figure 2-2). The registration process takes place by fitting the template onto the anatomical structure to be registered. Since the template is created directly from the imagery, the relation between the imagery and the patient is already known. All that is required to find the position and orientation of the patient in space is to measuring the location of the template in space. Template base registration only works however with rigid structures such as bone. Soft tissue structures can not be registered with template base registration. Template based registration has also not yet been implemented for use with augmented reality systems, but has been successfully used for other medical applications requiring accurate registration [17]. Template based registration has the potential for producing rapid registration in the operating room while requiring little work on the part of the surgeon to obtain the registration.

Figure 2-2: A template build from a 3D reconstruction for use in template based registration.

2.1.2 Display

The supplemental imagery or overlays are presented to the user by means of a display device. Two types of overlays can be used: planar (2-dimensional) or stereoscopic (3-dimensional or volumetric), each necessitating 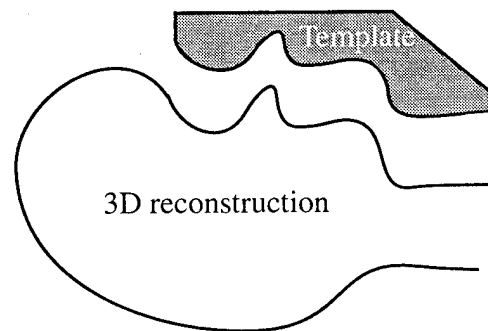a different display devices. Planar overlays use planar display systems such as computer monitors or televisions. Stereoscopic overlies, commonly used to enhance a 3-dimensional, require a dedicated display device in front of each eye to present a stereo pair to the user (see Section 2.2.1 on HMD/HUD and Section 2.2.3 on surgical microscopes) or a slightly more complicated single display device stereoscopic technique.

There are two methods of presenting 3-dimensional imagery from a single display: planar stereoscopic display devices incorporating spatially coded imagery to separate the left and right image, and volumetric displays presenting true 3D imagery in a three dimensional working volume. Planar stereoscopic systems present 2D specially created left and right image on the same display and use a filter to direct the correct image to each eye. If the two 2D images are created correctly, the brain will merge the images together to form a stereoscopic image. Some of the most common filtering methods for separating the images for each eye include active shutter glasses, polarized glasses (Figure 2-3), and color filter glasses[1]. The problem with these methods is the need for the users to wear special glasses to see the stereoscopic overlay. One single display stereoscopic system that does not require the use of special glasses uses a lenticular lens [8]. Lenticular lens systems spatially combine a stereoscopic pair into a single image of alternating strips of left and right images. A lenticular lens (a lens with evenly spaced "ripples" across it) is placed over the composite image and directs the left and right image strips to the correct eye.
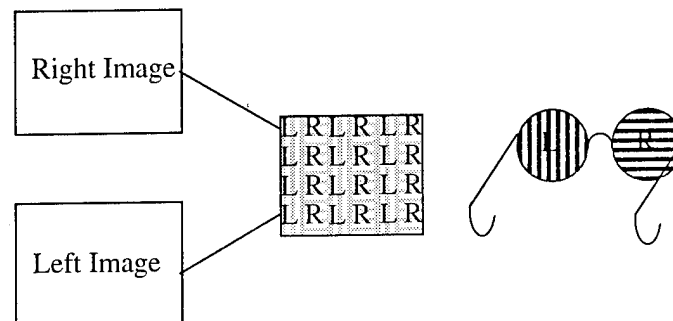
Figure 2-3: A polarization based planar stereoscopic display system. The left and right image are spatially combined to form one image. This composite image is projected through a polarizing filter with strips of polarizers rotated 90° from each other. The user wears a pair of polarizing glasses in which the left and right eye polarizers are rotated 90° from each other to separate the left and right image.

Volumetric displays produce a true 3D image without the need for special glasses to be worn by projecting the object into a 3-dimensional workspace. One such system works by projecting a voxel (3-dimensional picture element) onto a spinning helix. The 3D position of the voxel can be controlled by adjusting the time

1. All of these techniques are commonly used in the entertainment industry to present stereoscopic movies from a single screen.

at which the voxel is projected on the helix [24].

2.1.3 Tracking

Once the imagery is registered to the patient and displayed, it is desirable to allow the patient to be re-positioned. For example, during the procedure, a patient is sometimes moved to obtain a better approach to the surgical site. To prevent having to re-register the patient to the imagery each time the patient is moved, patient tracking becomes necessary. In addition, it is desirable for the surgeon to be able to change head position (vantage point) while maintaining a registered overlaid image. If the surgeon's vantage point changes, the overlay must be updated to the new position to maintain the correct registration.

Tracking is used to maintain the correct registration between image and patient if the patient or the surgeon's vantage point has moved (Figure 2-4). Tracking measures the position of the patient and the surgeon's vantage point relative to a known coordinate system. If either the vantage point or the patient have moved from the time the initial registration was performed, a tracking sensor is used to update the registration between the imagery and the patient (Figure 2-5).

Five of the most common techniques for performing tracking are: optical (both infrared and visible), magnetic, ultrasonic, 3D optical range sensors and radio frequency. With the exception of the range sensor, markers must be attached to the object being tracked and need to be rigidly fixed so motion between them can not occur. The relation between the markers is known, and is used to obtain the position and orientation of the object being tracked. If the markers move relative to one another and the motion is not accounted for, tracking errors will be introduced.

World
Object

Correctly registered
world object and overlay
image

Observer's vantage point
translated to the right

Overlay
Image

Observer and world object
tracked to maintain correct
registration

Object in world has moved
Away from observer

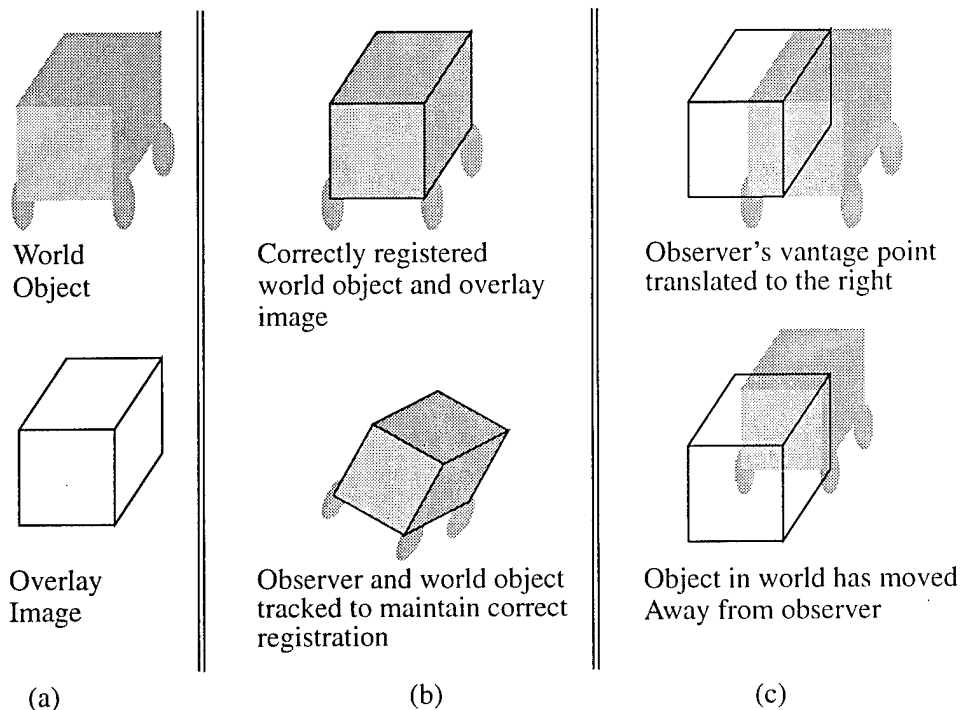(a)   (b)   (c)

Figure 2-4: (a) The world object and the overlay image. (b) The overlay object correctly registered to world object and tracked regardless of head or object motion. (c) The effects of not tracking the object or vantage point movements. Both types of motions will cause the overlaid imagery to no longer be registered to the world.

7

Optical trackers work by sensing the position of rigidly fixed markers using either infrared or visible light sensors. The relation between these markers is known, thus allowing the calculation of the position and orientation of the markers (a minimum of 3 markers is required to obtain the 6 degrees of freedom (DOF) of the system). Magnetic systems work by artificially creating three orthogonal magnetic fields. A receiver is placed in this magnetic field with three orthogonal pickup coils. By sensing the voltage across the three coils, it is possible to determine the position and orientation of the device. Ultrasonic sensors are very similar to optical ones, except that they use sound instead of light. Three microphones measure the time it takes for sound to travel from the transmitter (spark gap) to each of the microphones. The position of the transmitter is derived from the time required for the sound to reach each microphone via triangulation. A number of different range sensors are currently being developed. Range sensors return the distance to each point in the scene, a depth map, as opposed to the intensity as a regular camera does. The depth map is used with a surface based registration approach to find the position and orientation of the object. Radio frequency trackers are very similar to optical systems. The system locates the object in space by triangulating the transmitter's location relative to a number of receivers.
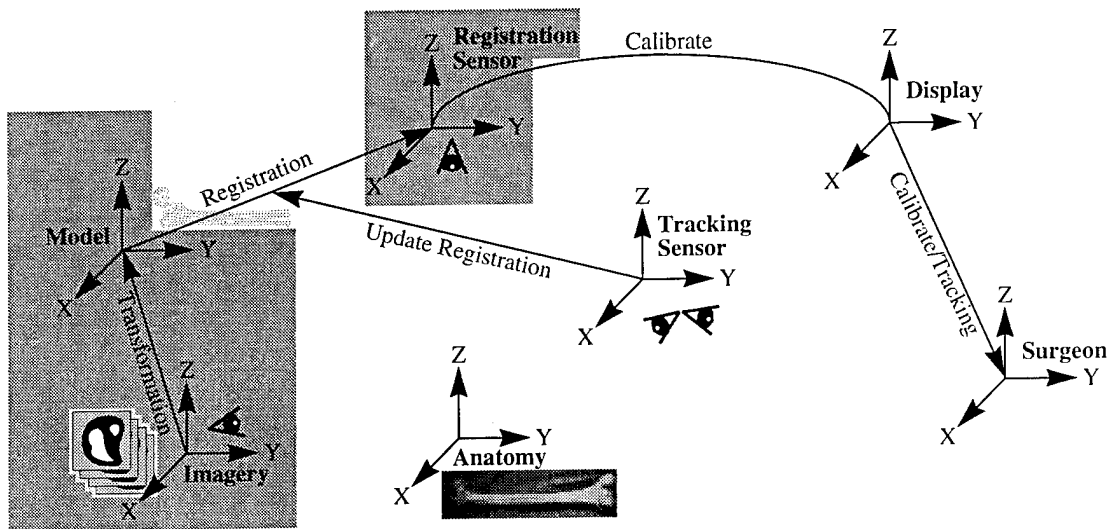
Figure 2-5: The tracking process. The registration process (grayed out area) from Section 2.1.1 found the initial position of the anatomy. The coordinate system the registration sensor used is calibrate to the display device to obtain the relation between the sensor and the display. Tracking is then used to measure the position of the surgeon's vantage point and the anatomy. Any change in position from the initial starting point is then used to update the registration between the overlays and the anatomy.

2.2 Types of Augmented Reality Systems

Augmented reality systems can be built using a number of different display techniques. Each techniques has its advantages over the other systems. Three of the most commonly discussed systems for use in medicine include heads up displays/head mounted displays, "Magic" windows, and microscope overlay systems.

2.2.1 Heads Up Displays/Head Mounted Displays

Head Mounted Display (HMD) and Head's Up Display[1] (HUD) consist of an display device mounted with-

---

1.HUD in this case refers to a head warn system and not a system that presents the imagery outside of a headset such as system found in some aircraft cockpit.

in a headset that presents imagery directly to the user's eyes. HMDs only allow the user to view the display device and occlude the direct view of the world. To enable the user of a HMD to see the world, a camera is used to present the world on the display device. HUD, on the other hand, allow the user to see the display device, plus a direct view of the world (Figure 2-6). To circumvent the potential problem of eliminating small anomalies in the world, due to the limited resolution of cameras and display devices, a HUD, is more desirable than a HMD for clinical use.

HUD/HMD have a number of problems that currently limit their successful use as medical augmented reality systems. First, the resolution available on HUD/HMD is poor. Most moderately priced HMD/HUD ($1K - $10K) are only capable of 640 X 480 pixels in gray scale. Radiologists and physicians will often complain about viewing imagery on low resolution displays for fear of missing small details of the imagery. Second, HMD/HUD have a limited field of view that limits the surgeon's ability to see the entire surgical site thus potentially by reducing the effectiveness of the system. Another problem with HUD/HMD due to the close proximity of the display to the user's eyes is so called "simulator sickness". Simulator sickness comes about from not updating the display fast enough when there is motion in the scene such as when the world object moves or the user's head moves thus changing the background view. This lack of display updating will produce a lag between the perceived motion and the actual motion, throwing off the user's sense of balance [16]. In addition, any perceived difference between the augmented environment and the actual environment can also cause simulator sickness such a perceived motion in the overlays relative to the world. Finally, current HUD/HMD are bulky and heavy, making them uncomfortable to wear over extended periods.
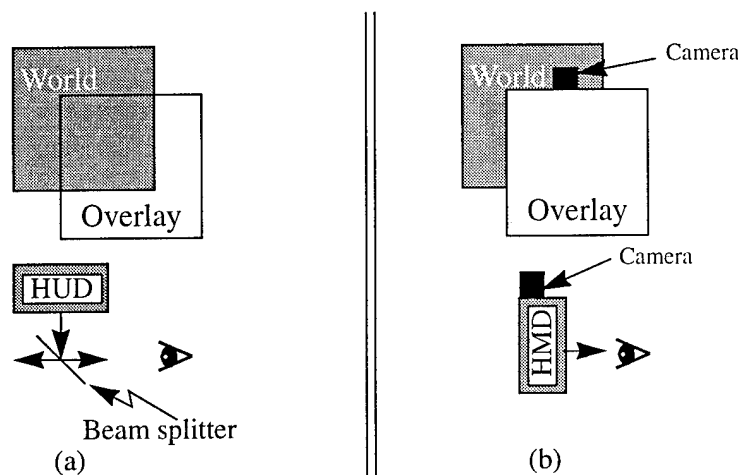
Figure 2-6: (a) Heads Up Display. A beam splitter is used to combine the display device with a view of the world. (b) Head Mounted Display. The display device occludes a direct view of the world. To view the world, a camera mounted on the HMD is used to present the world scene on the display device to the user.

2.2.2 "Magic" window and planar displays

"Magic" windows are movable displays placed over the patient used to add supplemental imagery to the world (Figure 2-7) [25]. The window is usually made out of an active matrix LCD panel that is lightweight and highly movable. The display devices occludes a direct view of the world, thus requiring a camera mounted to the back of the display to capture a view of the world. The view form the camera is combined with the overlaid imagery and presented on the LCD panel.

The magic window system does have a number of problems that must be overcome before it can be used in a medical setting. As with HMD, the magic window does not allow a direct viewing of the world. Instead a camera is used to capture the world scene that is then displayed on the LCD panel. Such an approach to display the world again limits its resolution, potentially eliminating small features. In addition, as the user's view point changes, the background scene must also change to be correct (Figure 2-8). This will complicate the device some by having to add a camera tracking and motion control system.

© 1996 Branko Jaramaz
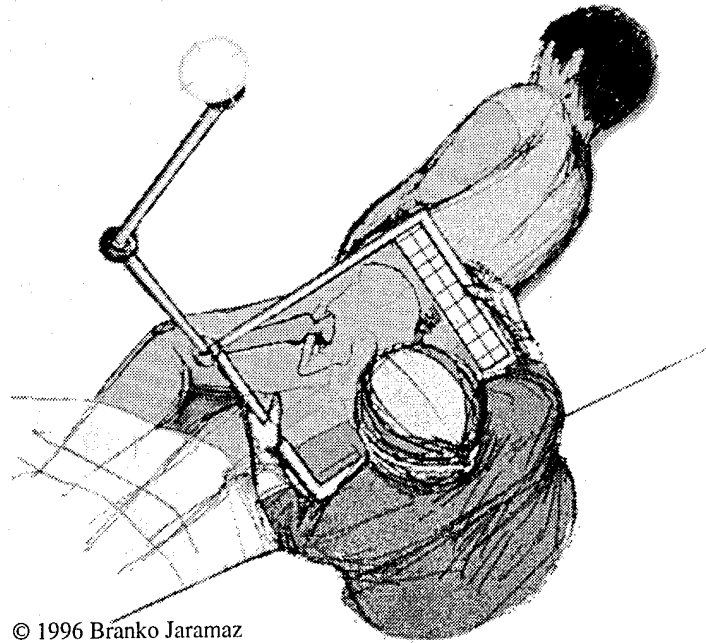
Figure 2-7: The magic window display. The surgeon can views both the patient by the aid of a camera and the supplemental overlay (the bone outlines)

Figure 2-8: The effect of the background when the head is moved relative to a fixed magic window. The background will drastically change as the head is moved through position 1 - 3 as shown in the mock magic window displays.

2.2.3 Surgical Microscope Augmented Reality System

The frequent use of surgical microscopes in the operating room make them ideal platforms for some augmented reality applications. Surgical microscope augmented reality systems combine the views of the patient and the overlaid ima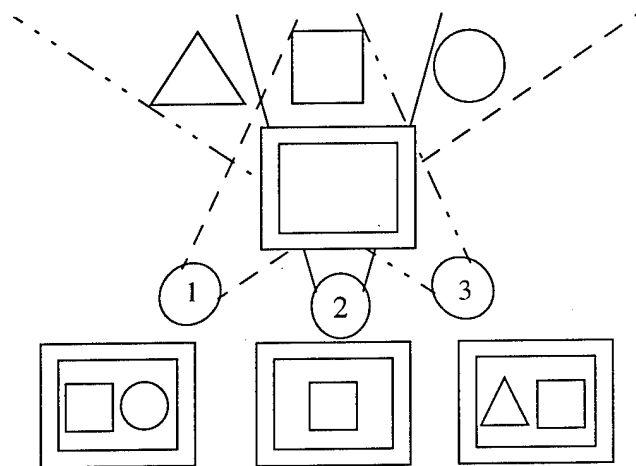gery by projecting the imagery onto a beam splitter located within the microscope's optical path (Figure 2-9) [20]. The surgeon then views the augmented environment by looking through their ocular (eye piece) to see the overlaid imagery and the patient. By presenting the overlaid imagery into the microscope's optical field, the resolution of the world is not limited. The tracking process to determine the surgeon's vantage point in order to maintain the correct registration is simplified by not having to track the surgeon's head; instead, the microscope itself is tracked. The surgeon's vantage point is the microscope's view of the world because the microscope constrains what the surgeon sees. Microscope based systems also allow for a second surgeon to view the augmented environment by the addition of a second ocular to the scope. Adding a second user to the system is sometimes important to allow surgeons and their assistants to have a common view of the supplemental imagery.
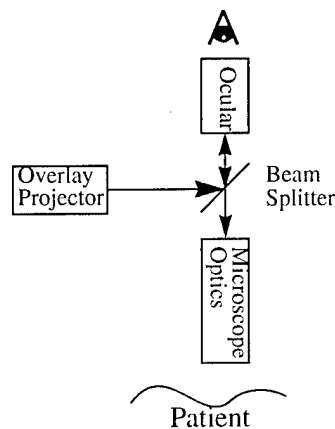
Figure 2-9: Schematic of an augmented reality surgical microscope.

Surgical microscope augmented reality systems do have a few drawbacks. The use of microscope augmented reality is limited to cases where a microscope is available. Surgical microscopes also present a small field of view (approximately 10cm X 10cm); for some surgical procedures, it may be desirable to have a much larger field of view. In addition, a surgeon's peripheral view of the surrounding operating room is limited by having to look through the ocular to see the augmented environment. Finally, an additional system is required to track any changes in the magnification of the microscope to maintain the correct scaling of the overlaid imagery. The tracking of changing magnification is easily solved, but does require addition mechanisms.

2.3 Current State of the Art Systems

In the early to mid 80's, Dr. David Roberts from Dartmouth medical center and Dr. Patrick Kelly from NYU medical center independently performed some of the earliest work in developing medical augmented reality systems [19,15]. Since then, a number of groups from around the world have begun developing medical augmented reality systems [7,16,14,5,18]. Most of these groups approach the problem of augmented reality slightly differently; however, they all use one of the above mentioned display techniques.

Some of the leading work in the field using HUD/HMD for augmented reality is currently being performed

by Henry Fuchs' group from the University of North Carolina [3]. Fuchs et al. have been building a system that enables a user to view reconstructed ultrasonic imagery of a fetus that appears to be inside the pregnant patient. The system collects ultrasonic imagery and in near real time updates a 3D volume rendering of the fetus as it moves. The volume rendering is registered to the patient by tracking the position of the ultrasound imager and where the user's head is located relative to the imager. Once the imagery is registered to the patient, it is presented to the user who is wearing a HMD. Along with the 3D reconstruction, a view of the patient is added to the display from a camera mounted on the HMD.

A number of groups have initially proposed a magic window type system but are currently only using a video monitor with a camera to capture the world scene [10,5,25]. Video monitor systems are similar to magic windows, but are mounted outside of the surgical field requiring the surgeon to look away from the surgical sight to view the augmented environment. Grimson et al. [10] have combined MRI reconstructions of tumors with a views of the patient's head to assist in localizing the site of a craniotomy (opening the skull to expose the brain for tumor removal). The system they have developed consists of the following parts: a camera looking down the sight lines of the surgeon (for collection of the world image), laser range finder (used for registration and tracking), and a monitor (for displaying the world and overlaid imagery). The group has devoted a significant amount of time to the development of automatic registration and patient tracking techniques. The 3D range finder is used to collect a three-dimensional depth map of the patient. The depth map is matched to a model of the patient created from the imagery by means of surface based registration. The current system has been used on at least 8 different neurological cases at Brigham and Women's Hospital in Boston, MA. While using the system, they have reported a great reduction in surgical time and an improved accuracy in the placement of the incision [9].

Davey et al. have proposed a similar system to that of Grimson's group [5]. The major difference between their work and that of Grimson's is in their application. Davey's group is using the system to determine if there are any changes in the anatomy between the time the MRI, CT or DSA was taken and the present. The registration process between the image and the patient is currently done by hand. They propose the use of a stereoscopic camera based system and surface based registration to obtain faster, more accurate image-patient registration, but have not yet implemented it. The Davey system is also integrated with an ISG viewing wand (ISG Technologies, Mississauga, Canada) to assist in intra-surgical localization and guidance. The ISG viewing wand is a commercial system that allows a surgeon to point to an anatomical feature on the patient and then view this specific feature in the imagery on a workstation. By combining the ISG viewing wand with their system, additional guidance and surgical cues can be obtained and an intuitive interface for the surgeon is provided to assist in the detection of anatomical changes between when the imagery was acquired and the present.

The surgical microscope is currently the most commonly explored technique for medical augmented reality systems. A number of groups are working on developing complete microscope systems [20, 7]. In addition, a commercial system has been developed [27]. Roberts' group from Dartmouth was one of the first to start developing and clinically using an augmented reality system based on a surgical microscope. The current system consists of a stereoscopic microscope, a sonic digitizer, a miniature CRT and beam splitter for image fusion. The group has been developing a low cost, accurate sonic digitizer to allow for easy tracking of the microscope and fiducial localization. The complete system has been used in approximately 20 cases with relatively good results [18].

Edwards et al. have been developing a system that is very similar to the Roberts surgical microscope [7]. Both systems currently only present simple contour outlines and guidance cues as overlaid imagery. However, through the use of depth cues Edwards' group has been developing techniques to make these simple overlays appear more realistic.

12

Carl Zeiss is the first company to introduce a product based on an augmented reality enhanced surgical microscope[27]; called the MKM system. As of January 1, 1996, 27 of these systems have been sold worldwide at a price of approximately $800K per system. The systems perform automatic registration between the patient and the imagery using fiducial marks located in the microscopic field of view. The current system does not allow for patient tracking; if the patient moves, the imagery must be re-registered to the patient. In addition, the overlaid imagery presented to the surgeon only consists of contours of the anatomical structures and simple guidance information.

2.4 Limitations of Current Systems

Despite the pioneering work performed by the groups mentioned above, all of the current systems have some limiting factors that will prevent wide-scale acceptance in the medical community:

- The use of fiducial based registration.
- No patient tracking system
- The quality and resolution of the overlay is low.
- Limited work on developing proper overlays.
- The surgeon must look away from the surgical sight to see the augmented environment.
- The real world scene is limited by a camera and display device.

The use of fiducial based registration is required by some of the above systems to obtain the registration between the patient and imagery. Fiducial based registration requires rigid fixation of markers to the patient. Any motion of the markers relative to each other will decrease the accuracy of the registration. Fiducial based registration also requires a scan to be performed once the markers have been affixed. The additional scan must be performed so that the markers can be located in the imagery. The fiducials must be accurately located in both the imagery and on the patient in the operating room to minimize registration error (Section 3.2.3). The accurate localization of fiducial marks can be difficult in a surgical setting if the markers are not placed in an easily assessable spot for the procedure.

With the exception of Fuchs' group, none of the above system have incorporated real-time patient tracking. With out patient tracking, the overlays must be re-registered to the patient whenever they are re-positioned. The re-positioning of a patient is commonly performed throughout the procedure to allow the surgeon to gain the best approach the surgical sight. With the lack of patient tracking, the above systems greatly limit the ability for the surgeon to re-position the patient so that they can obtain the best possible approach.

In all of the above augmented reality systems, the resolution of the overlays is very low (less then 640 x 480). In order to provide accurate navigation and legalization overlays to the surgeon, the resolution of the overlays must be increased. In addition to the limited resolution of the overlays, very little work has gone into developing and creating overlays that will provide the best localization and guidance information to the surgeon for the procedure at hand. It is a general conclusion in published work that the quality and types of the overlays should be improved to make augmented reality systems more acceptable; however, it is unclear what these improvements should be.

All of the monitor based systems require the surgeon to look away from the patient to view the imagery. If the surgeon can not view the augmented environment directly over the patient, having to look away can greatly diminish the significance of the direct relation created by the augmented reality system between the

imagery and the patient. Finally HMDs, magic windows and planar display systems all limit the resolution of the world by using a camera to capture the world scene. The limited resolution of the world is not acceptable to surgeons due to the potential of eliminating small but critical anatomical features.

3 The MRCAS Augmented Reality System

The Medical Robotics and Computer Assisted Surgery (MRCAS) group of Carnegie Mellon University is working on developing a new medical augmented reality system. The new augmented reality system is being designed in an attempt to overcome some of the limitations found with the current state of the art systems. Currently, a developmental and research prototype of the MRCAS system exists. However, a new clinical version of the system will be built and tested by the summer of 1996.

3.1 MRCAS Augmented Reality System - Overview

The basic concept of the MRCAS system parallels some preliminary work performed in the early 80's by Christopher Schmandt at Massachusetts Institute of Technology[22] and is similar to a microscope based augmented reality system. The user views the patient through a beam splitter (a half-silvered mirror) which is both transparent and reflective. Positioned above the beam-splitter is a display device (CRT monitor or video projector). The patient is seen directly through the beam-splitter, while a reflection of the video display appears to float within the workspace (Figure 3-1). The MRCAS system is capable of presenting both 2-dimensional and 3-dimensional overlays. For presenting stereoscopic imagery, currently a pair of liquid crystal shutter glasses are used. A 6 degree of freedom head tracking system is integrated with the overlay device allowing the user to change their vantage point, while maintaining the image/patient registration. In addition to head-tracking, a 6 degree of freedom patient tracker is integrated into the system to allow for patient motion.

An advantage of the MRCAS approach over some of the other systems is that since it does not limit the resolution of the world, the likelihood of eliminating small anatomical features in the world is removed. The MRCAS system also offers a larger field of view than a microscope based system and does not limit the surgeon's peripheral view of the operating room. In addition, the system can be built to take up considerably less room in an operating room than a microscope system. The first MRCAS augmented reality system was developed as a platform to testing different user interfaces, registration and tracking techniques.

A number of limitations with the current developmental prototype of MRCAS augmented reality system must be overcome before the system can be used in a clinical environment. First, the issue of registration between patient and imagery must be addressed. Second, the current system for presenting stereoscopic overlays limits the available display devices that can be used. Third, when the system is used to present stereo imagery, the available resolution on the display is not adequate. Fourth, the curved surface of the CRT display used limits the registration accuracy and distorts the overlay images. Finally, a new tracking system or configuration is needed for both head tracking and patient tracking to eliminate the need of a direct line of sight between the object being tracked and the tracker.
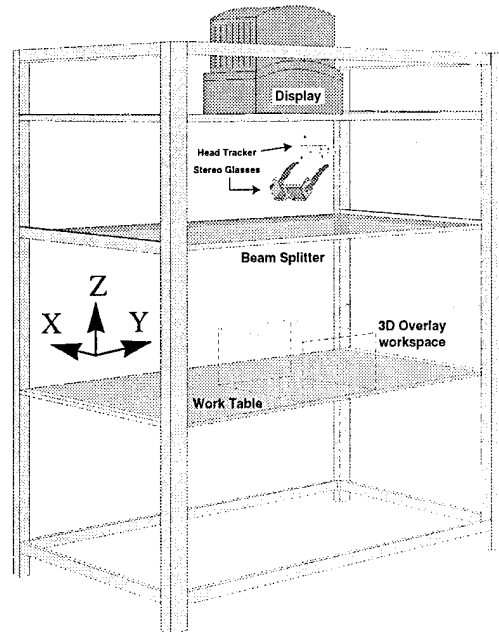
14

Figure 3-1: The MRCAS augmented reality prototype system. The Patient would be located on the work table.

3.2 Selecting Basic Technologies for the Clinical MRCAS System

The current version of the MRCAS augmented reality system was built as a developmental platform to try different overlay techniques and system configurations. The new clinical version of the MRCAS system requires solutions to be found for the above mention limitations. Several areas will be examined to resolve these issues including: stereoscopic display techniques, display devices, registration (fiducial bases, shape based, plus a new novel template based system) and finally patient/surgeon tracking.

3.2.1 Stereoscopic Displays

The MRCAS system currently uses active shutter glasses to produce the stereoscopic overlays. The active shutter glasses (CrystalEyes, Stereo Graphics Inc., San Rafael CA) run the display at twice its regular frame rate. The increased frame rate is used to sequentially display the correct stereo image to each eyes while attempting to eliminate image flicker. The user wears a pair of active shutter glasses synchronized with the monitor to occlude one of the eyes while the other eye views its part of the stereoscopic image (Figure 3-2).

For active shutter glasses to work without producing a flickering display, the display device used must be capable of running at the increased refresh rate. Most high definition monitors and video projectors are capable of running at the faster rate. However, a problem arises when liquid crystal displays (LCD) or plasma displays are used. These displays are not yet capable of the very fast refresh rate required because of the time it takes for each image displayed to "die" out. For example, if the display has just finished presenting the left image and is changing over to presenting the right image, the left image has some finite amount of time before it is no longer visible. This image persistence will produce a bleed-through between the left and right eyes, thus greatly diminishing the stereoscopic effect.
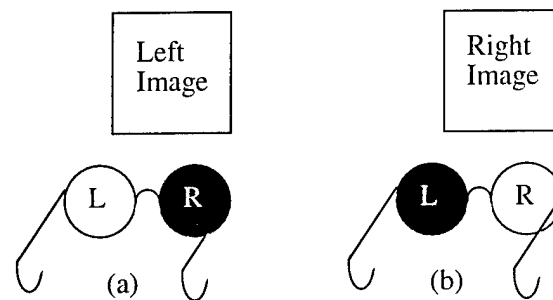
Figure 3-2: The Active shutter glasses. When the left image is displayed on the screen, the right eye shutter is blocked so only the left eye can see the screen (a). When the right image is displayed, the left eye is blocked so only the right eye can see the screen (b).

With the limitation of usable display devices introduced by using active shutter glasses for presenting stereoscopic overlays, a new approach is needed. A number of stereoscopic techniques including color filter glasses, polarization glasses, active shutter glasses and lenticular lens systems, as well as a volumetric display based on the spinning helix systems (Section 2.1.2), were all evaluated for use with the next generation of the MRCAS augmented reality system. The stereoscopic display technique for the medical augmented reality system should:

- Not impede the view of the world (i.e. darken, colorize, reduce the resolution of world, etc.).

- Not make the user sick or be bothersome to use (i.e. flicker, cause eye strain, or require heavy head wear).

- Be easy to integrate with the existing system.

- Present high resolution stereoscopic imagery.

- Be compatible with a wide range of display technologies (i.e. LCD, CRT, EL).

The different approaches for presenting stereoscopic overlays for use with the MRCAS Augmented reality system are compared in Table 3-1. After reviewing all of the different techniques against the above criteria, a polarization based stereoscopic system was selected for use with the next generation of the MRCAS augmented reality system for a number of reasons. First, because a passive polarizer is placed over the display device (Figure 2-3) and no special display timing is required, polarization based stereoscopic systems are compatible with a number of different display devices including LCD and CRT based systems. As will be seen in Section 3.2.2, LEDs are more desirable as a display device for use in the MRCAS system then a CRT based display due to the flat image they produce and the smaller physical volume. Second, polarization based stereoscopic systems require much lighter passive polarizing glasses to be worn than those required by active shutter glasses. The smaller, lighter weight glasses are much more comfortable to wear over extended periods of time.

The one problem with polarization based stereoscopic display systems that must be overcome, is that they darken the environment as a result of the polarizing filters. If a polarization system can not be found that does not darken the environment significantly, the active shutter glasses will be used with the next generation system thus limiting the types of display devices. Finally, the volumetric approaches while viable to use with the MRCAS system (Section 3.2.2), was not considered because it currently is an experimental systems, offering limited viewing angle, and low resolution displays.

Display Method	Darken Overlay?	Display in Color	Flicker?	Eye Strain?	Viewing Angle	Res of Overlay?	Work with different displays?
Active Shutter Glasses	20%	Yes	Yes - Some	Yes-Some	High	Half of Display Res	NO - Not LCD
Polarized Glasses	<= 40%	Yes	NO	Yes-Limited	High	Half of Display Res	Yes
Color Filter Glasses	20%	NO	NO	Yes-Some	High	Full Display Res.	Only Color Displays
Lenticular Lens	10%	Yes	NO	NO	Low	Dependent on lens geometry	Yes
DOD Volumetric Display	NO	Yes - Limited	Yes	NO	Limited	Poor	N/A

Table 3-1: Different ways of displaying stereoscopic imagery with MRCAS prototype augmented reality system

3.2.2 Displays Devices

The display device presents the overlay imagery to the user. To ensure a high quality overlay (high resolution, non-distorting, and easily viewed), a number of stringent design criteria should be considered when selecting a display for the MRCAS augmented reality system. The display device should:

- Produce a bright image.
- Be flat or project a flat image.
- Be compatible with one of the stereoscopic techniques discussed in section Section 3.2.1 if stereoscopic overlays will be used.
- Have a high resolution (at least 1024 x 768).
- Be portable.
- Be easy to integrate with the system.

The MRCAS system projects an image onto the beam splitter to produce the overlay imagery. To overcome the high ambient light levels present in the operating room (Approximately 50 - 100ft -lumens), the system must be capable of projecting a bright image onto the beam splitter. In the prototype system a standard 21" CRT monitor (Silicon Graphics, Mountain View CA, Model Number D-M21G) which produces 25ft - lumens at maximum intensity was used. The output from this monitor was bright enough to produce the overlaid imagery using an 80/20 beam splitter (80% transparent and 20% reflective to incident light) when the entire system was draped to reduce outside ambient light. However, for the operating room situation, the draping would have to be removed for the surgeon to have adequate access to the rest of the operating room, thus necessitating a display that is at least 50ft - lumens, which is significantly brighter than the 25ft - lumens produced with a standard computer monitor.

The use of a flat display (LCD) simplifies the process of presenting correctly registered stereoscopic overlays when compared to a curved display (CRT). The MRCAS system produces a virtual image that is the same shape as the display surface. Therefore, a flat display produces a flat virtual image and a curved dis-

17

play will produce a curved virtual image (Figure 3-3). The virtual image will be located a distance below the beam splitter equal to the distance the display is above the beam splitter[1]. The user views the virtual image, so the shape and position of the virtual image controls where a user will see the overlay. The problem with a curved virtual image is that it will cause the overlay to be distorted and be incorrectly placed (Figure 3-4). Both the distortion and the misplacement can be corrected for in the stereo prospective projection, but this greatly complicates the stereo projections because a geometric model of the display surface must be introduced. To reduce the rendering time for the overlaid imagery, the stereo prospective projections should be as simple as possible.
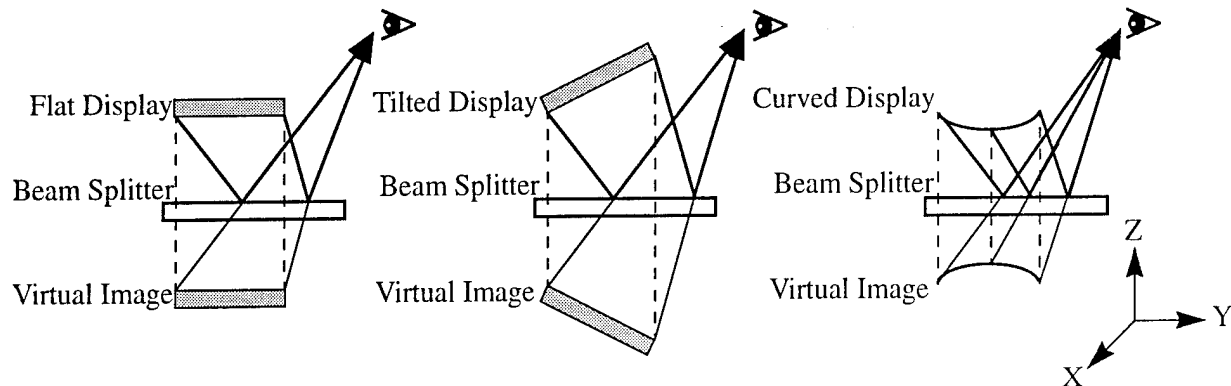

Figure 3-3: The effect of the screen geometry on the virtual image produced.

The 21" CRT monitor used in the prototype systems had a center point 12.0 mm higher then the edges (Figure 3-5a). To keep the stereo projections simple for the prototype system, the overlays are rendered on a flat image plane located half way between the center and edge points, causing the actual overlay to be miss register by as much as 6.0 mm. In addition, the curved screen shifts the locations of the image seen as a function of the view point due to parallax (Figure 3-5a). Finally, large curved CRT displays will introduce refraction distortion from the thick front face plate (Figure 3-5b). The refraction in the glass and the display curvature tend to distort the image in opposite directions but do not cancel out. The problem with these distortions is that they are viewpoint dependent and non-linear making them more difficult to correct for in the stereo projections [6]. All of these effects drive for the use of a flat panel display for the next generation system.

Finally, the portability of the display should be considered when selecting a display device. It is possible to find a display that fulfills all of the above requirements, but is too big and awkward to easily integrate into the current system. The desired working volume of the new system is approximately $0.25m^3$ (1m X 0.5m X 0.5m). If the system is much larger than this, it will be difficult to use it in an operating room setting due to the confined nature of the operating room.

A number of displays were evaluated against the preceding criteria (Table 3-2). The display devices include: Liquid Crystal Displays (LCD) and projectors, plasma displays, Electroluminescent (EL) displays, high resolution CRT monitors and projectors, Super High Definition Display Television (SHDTV), and other novel displays including the Texas Instruments micro mirror display [21] and a diffraction grading display[1].

1.Because the virtual image exactly mirrors the display device and it location is a function of were the real image originates, volumetric display can also be used to present 3-demential overlay imagery.
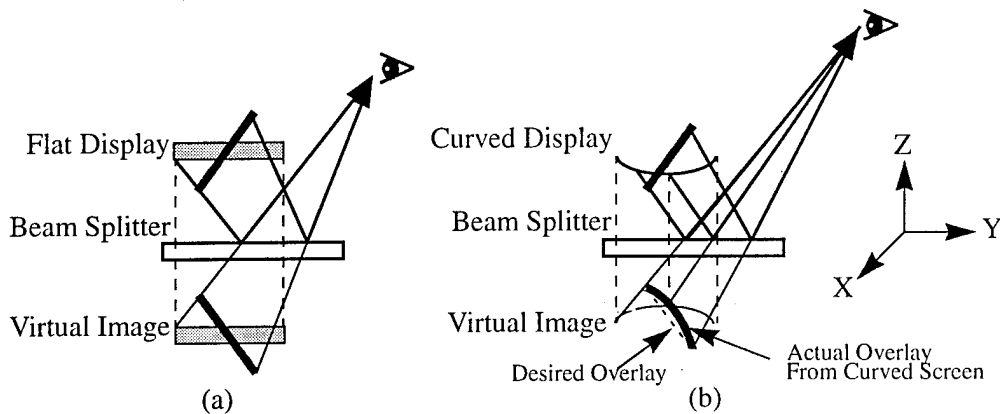
Figure 3-4: a) A flat screen will produce a flat un-distorted virtual image. b) A curved screen will produce a curved virtual image that has been distorted by the screen curvature making it no longer register to object in the Z direction.
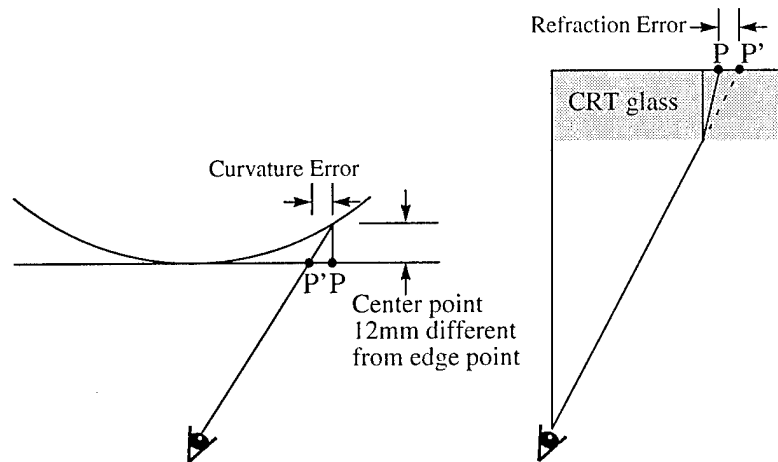

Figure 3-5: a) Curvature distortion. Point P is the pixel illuminated and point P' is the perceived pixel. This distortion will cause the image to appear reduced. b) Refraction distortion. Point P is the pixel illuminated and point P' is the perceived pixel illuminated. This distortion will cause the image to appear magnified. These two distortions have opposite effect but do not cancel out. The distortions are a function of the eye position [6].

If a polarized single source stereoscopic system can be found that does not significantly darken the environment (Section 3.2.1), then an LCD projection system will be used for the next generation of the MRCAS system. LCD projectors are capable of producing very bright imagery at high resolution from a flat screen and are available in a smaller physical volume than CRT based projection systems. If an adequate polarization based stereoscopic method can not be found or developed and the active shutter glasses must be used a high resolution CRT projection system will be used for the display device. The CRT projection system are capable of handling the increased frame rate required with active shutter glasses. Like the LCD projection systems, CRT projection system's produce bright, flat image and are available in high resolution. A limiting factor for the CRT based system is the size. CRT projection systems are generally much larger than a comparable resolution LCD projection system. This is why a LCD projector is deride for the clinical system.

Display	Active Light Source?	Bright?	Flat Image?	Compatible with Active Shutter and Polarization based stereoscopic systems?	High Resolution (Currently Available)?	Portable?	Easily Obtained?	Wide Viewing Angle?
LCD (Non-Backlit)	NO	NO	Yes	NO - Not active shutter	Up to 1280x1024	Yes	Yes	No - Max around 90°
LCD Projector	Yes	Depends on Light source	Yes	NO- Not active shutter	Up to 1280x1024	Yes	Yes	Yes
Plasma	Yes	Yes	Yes	NO- Not active shutter	640x480 +	Yes	Limited	Yes
EL	Yes	NO	Yes	NO- Not active shutter	Up to 640x480	Yes	Yes	Yes
High res Monitors	Yes	NO	NO	Yes	Up to 2048x 1280	NO	Yes	Yes
High res Projectors	Yes	Moderate	Yes (Can be Corrected)	Yes	Up to 2048 x 1280	Limited	Yes	Yes
SHDTV	Yes	NO	NO	Yes	2048 x 1280 +	NO	No	Yes
TI Micro Mirror	Yes	Depends on Light source	Yes	Unknown	640x480 +	Unknown	Not in full production	Initial Projections Limited
Diffraction Grading (Incident light)	NO	NO	Yes	Unknown	640 x 480 +	Unknown	Not in full production	Initial Projections Limited

Table 3-2: A comparison of different display device available for the MRCAS augmented reality system.

20

3.2.3 Registration

The three possible approaches for performing registration with the MRCAS augmented reality system are shape, fiducial and template based registration. A registration technique for a medical augmented reality system should meat the following criteria:

- Non-invasive to the patient as possible (i.e. would like to avoid placing additional markers or making incisions for registration).

- Does not require additional scans (i.e. the initial diagnosis scan can be used).

- Capable of an accuracy less than a millimeter.

- Fast, easy, and does not interfere with the case.

Table 3-3 compares shape, fiducial and template based registration for use in a neurosurgical version of the MRCAS augmented reality system. After reviewing the three possible approaches for registration, template based registration should be considered for use with the clinical version of the MRCAS augmented reality system. If templates can be designed and manufactured to hold tight tolerances, then it is believed template registration will offer comparable accuracy to the other approaches. In addition, the template approach potentially offers faster and easier intra-surgical registration than the other two approaches because the template is simply fitted to the anatomical structure to be registered and no points must be collected. The template is also a good place for markers to be affixed for patient tracking (Section 2.1.3) due to the tight coupling between the template and the patient. Finally, template based registration is potentially less dependent on the resolution and inter slice spacing[1] of the scan than the other approaches. Fiducial base registration requires high resolution scans around the fiducial marks to allow for accurate marker localization in the imagery. Shape based registration, however, requires a high resolution surface model wherever the user might collect points on the anatomical structure. The template based approach can be constrained so that the template will only make contact with the patient where the scan data is present. Such an approach does not place requirements on the scan resolution and spacing such as is required with the shape and fiducial based approach.

If it turns out that the required accuracy is not obtainable with template based registration, then shape based registration will examined further. Shape based is more desirable than fiducial based because of the freedom in point selection and not having to attach markers to the patient before the scan is performed. The addition of marks introduces additional trauma from the marks, and an extra scan as well as cost to the case. However, the reliability of the shape based approach is currently uncertain for use in medical domain. If it turns out that shape based registration does not have the required accuracy in the medical domain then the fiducial approach will have to be used because of the higher accuracy obtainable.

1.The inter slice spacing is the distance between the slices that compose the volumetric data set. For most CT systems this can be as small as 1mm between slices. Typical CT protocols for visualization currently use 5-10mm inter slice spacing.

Registration Method	Invasive to Patient?	Require Addition Scans Over Diagnosis Scan?	High Accuracy?	Fast/Easy to use intra-surgically?
Shape Based	No - Depending on Point Collection Method	Yet To Be Determined - Depends on resolution of initial scan	Yes - If the points used are well selected	More points required, but allows freedom in point used.
Fiducial Based	Some fiducial markers must be screwed into bone but other can be attached to skin	Yes - Must scan after fiducials are installed	Yes - If markers remain stationary. Not as accurate if skin markers are used due to possibility of motion.	Yes - Limited number of points to collect
Template based	No	Yet to Be Determined - Depends on resolution of initial scan	Yet to be Determined - believed template can be well constrained and good contact to bone can be achieved	Yes - Place mask on face.

Table 3-3: Registration methods for use with a neurosurgical version of the MRCAS system

3.2.4 Patient/Surgeon Tracking

The five most common techniques for performing tracking were evaluated for use with the MRCAS system: optical (both infrared and visible), magnetic, ultrasonic, 3D range sensors and radio frequency. With all of the different approaches, it is difficult to select one best approach to use. A tracking system for a medical augmented reality system must fulfill the following requirements:

- Not interfere or be affected by the equipment in the OR or its personnel.
- Be capable of producing an update rate of at least 30Hz for patient tracking and 30Hz for surgeon head tracking to maintain the usability of the system.
- Be accurate down to a few millimeters.
- Not be affected by environmental conditions.
- Be as non-invasive to the procedure as possible.

The five most commonly used tracking techniques mentioned in Section 3.2.4 are compared in Table 3-4 on page 27 for use with the clinical version of the MRCAS augmented reality system. The optical infrared approach was selected for use with the next generation system. This tracking system was chosen for its lack of interference with the OR environment, range, update rate and accuracy. However, the one problem with it is that it requires line of sight; that is, the tracker must have a direct view of the markers being tracked.

This can be a problem because it is sometimes difficult to ensure a direct line of sight between the patient/ surgeon and the sensor. The line of sight issue can be reduced with multiple receivers placed strategically around the room to allow for a larger working area before occlusion to the receiver occurs. Another possibility is to combine the magnetic tracking system with the infrared tracker. The advantage of the magnetic system is that it does not require a line of sight to the sensor. However, magnetic systems are very sensitive to interference from metallic objects located in the working volume. By combining the two, the magnetic tracker could be used for the few instances when the markers are occluded from the infrared tracker. This would allow for the higher accuracy of the infrared system to be used and still allow for some tracking when the line of sight can not be maintained.

3.3 The Clinical MRCAS Augmented Reality System

The new clinical version of the MRCAS augmented reality system will be smaller and easier to integrate into the operating room. The system will consist of an LCD projector mounted above the surgical field. The projector will present the overlay on a beam splitter mounted on an arm that is easy to move in and out of the surgical field. The surgeon will wear passive polarizing glasses to view the stereoscopic overlays. The registration between the patient and the imagery will take place by using template base registration. Finally, the surgeon, patient and entire system will be tracked using an infrared optical tracker. A magnetic tracker will also be integrated into the system to allow for surgeon and patient tracking to occur even if they are occluded from the optical tracker. The entire system is tracked, including the display and beam splitter so that the relation between the display and the beam splitter can be known so that the position of the virtual image can be accurately determined. In addition, since it is possible to determine where the virtual image will be if the display and beam splitter are tracked, the display and the beam splitter do not have to be held in a constant relation to each other, thus allowing for the structure of the system to be less rigid.

Tracking Method	Affected by OR equipment?	Interfere with OR Equipment?	Line of sight tracking?	Affected by OR environment conditions?	Accurate to better then a few millimeters?	Update rate of at least 1 Hz?	Non-Invasive to procedure?	Readably available?
Optical (Infrared)	NO	NO	Yes	NO	Yes	Yes	Yes	Yes
Optical (Visible)	NO	NO	Yes	Yes - Lighting	Depends on range	Yes	Yes	Yes
Magnetic	Yes - ferrous materials	Has potential to	NO	NO	Yes - However sensitive to working volume	Yes	Yes	Yes
Ultrasonic	Yes - Sounds by equipment	NO	Yes	Yes - temperature variations	If temperature compensated	Yes	Yes	Yes
Range sensor	NO	NO	Yes	Yes - Lighting	Depends on range	Some systems	NO - Structure must be exposed	Limited
Radio Frequency	Has potential to be	Has potential to	NO	NO	Yes	Yes	Yes	NO

Table 3-4: A comparison of the possible tracking systems for using with the medical augmented reality systems

4 Validating an Augmented Reality System

The success of an augmented reality system relies on all of the different technologies working together. Before an augmented reality system can be placed into a clinical setting, the complete system made up of individual technologies must be evaluated. A number of different evaluation methods are currently being developed and applied to the MRCAS system including: the utility of a 3D system relative to a 2D one, determining if 3D reconstructions or the planar cross section slices from the volumetric data set offer better guidance and localization and measuring the end to end accuracy of the entire system.

4.1 2D Vs. 3D Performance Experiment

A good test of the validity of a 3D augmented reality system involves the following question: does a 3D approach to augmented reality offer better guidance and navigational aids than a 2D system? The answer to this question is a function of the procedure being performed with the augmented reality system. Some procedures can benefit from a system that provides stereoscopic overlays, while others require monocular overlays. To evaluate the best approach for each class of procedure, a number of different tasks must be performed using a 2D and 3D system. One such evaluation task applied to the MRCAS system measures the speed and accuracy with which a user can align a pin in 3D space. The task requires the positioning of a pin in 5 degrees of freedom (the rotation about the long axis of the is of no interest). The clinical equivalent of such a task is the positioning of a surgical drill for insertion of external fixation pins [4].

To run the performance experiments, a 2D augmented reality system was created. The 2D system consists of three cameras mounted orthogonally to each other (top, left and frontal views). The user was presented with the three views plus the overlay information on three different monitors (Figure 4-1). The overlays presented to the users were lines drawn in each of the three views corresponding to the desired position of the pin in 3D space. The user had to align the pin with the line presented in each of the three monitors. The time required to align the pin as well as the position and orientation of the pin was recorded using an Optotrak 3D position system (Northern Digital Inc. Waterloo Canada).

The MRCAS augmented reality system described in Section 3.1 was used as the 3D system (Figure 4-2). For the pin alignment task, a 3D model of the pin was presented to the user as a stereo overlay. The user had to position and align the pin to that presented by the overlay in 3D space. When the user was satisfied that the pin was aligned to the overlay, the time required to position the pin as well as the final position of the pin was recorded using the Optotrak system.

A number of interesting trends were apparent from these experiments. When using the 3D MRCAS augmented reality system, the placement of the pin on average was 13.4 seconds faster than with the 2D system with an ANOVA statistical significance of p=0.01. Despite the dramatic improvement in speed obtained by the 3D system, the 2D system's position accuracy (rotational and translational error) was 2.2mm (p=0.02) and 4.7° (p=0.01) better than the 3D MRCAS system. The results of the two systems are compared in
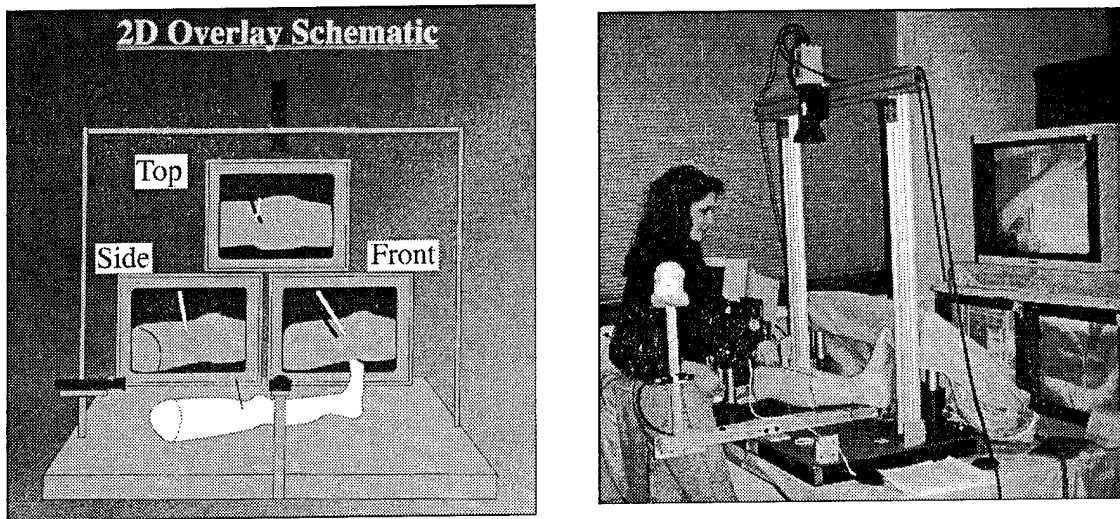
Figure 4-3.

Figure 4-1: A schematic and actual picture of the 2D augmented reality system used to in the performance experiments. The white lines in the schematic correspond to the desired position of the pin (black) for each camera position.
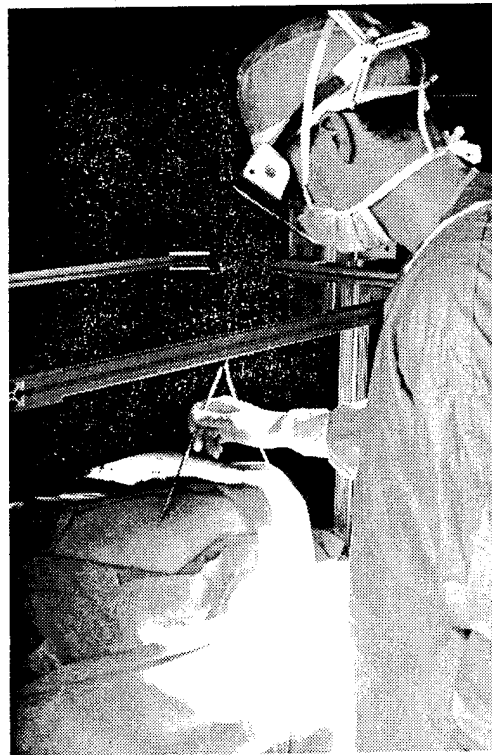

Figure 4-2: The MRCAS augmented reality system being used for a pin alignment test.
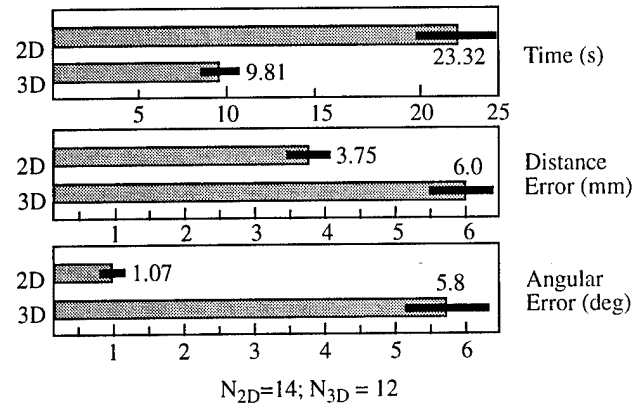
$N_{2D}=14; N_{3D} = 12$

Figure 4-3: Summary of the results between the 2D and 3D augmented reality system.

The significant positioning accuracy advantage obtained by the 2D augmented reality system can be explained by examining the display accuracies of the two systems. The 2D overlay system has a physical accuracy of 0.5mm/pixel over the entire system (camera, lens and monitors). The accuracy was obtained by measuring the optical work space in which the experiment was performed and then dividing by the number of pixels on the monitor. On the other hand, the 3D overlay system had a physical accuracy of 0.7mm/pixel in the XY plane (calculated with a similar technique as the 2D system) but the Z error (vertically in and out of the workspace) was as high as 4.0mm/pixel. Some of the errors in the 3D overlay system can be attributed to not correcting for a curved screen, head tracking position error, and system calibration errors.

Calibration errors in the position of the screen can also introduce errors. In addition, not knowing exactly where the virtual image is located (caused from not knowing the exact orientation and position of the beam splitter and monitor) will introduce errors in the Z direction (Figure 3-3). One way to circumvent this problem is to track the position of the monitor and the beam splitter. The virtual image will be the shape as the monitor, except it will be flipped 180° at a distance below the beam splitter equal to the distance that the monitor is above the beam splitter (Figure 3-2).

The final major source of error in the 3D overlay system came from head tracking. The stereo perspective projections, used to determine what pixel to illuminate to make the overlay appear in the correct position require knowing where each of the user's eyes are located (Figure 4-4). When building the head tracking system, an assumption was made as to the distance between the user's eyes (Inter-Pupillary Distance, IPD). A single IPD was used for the system to eliminate having to configure the system for each user's IPD. The problem with this assumption is that each user of the system will have a slightly different spacing between their eyes ranging from 50mm up to 75mm with the mean IPD equal to 64mm [12]. If the correct IPD is not used, the stereo projection will not present the overlay correctly to the users (Figure 4-5). If an IPD is used that is too small, the overlay will appear too be closer to the user. However, if the IPD is too large, the overlay will appear to be too far away. The only way to correct for this problem is to configure the system for each user's IPD.

It is believed that once all of the above problems have been corrected in the new clinical version of the MR-CAS system, the 3D approach will offer the same advantage in speed while increasing the overall accuracy of the system to a level comparable to the 2D system.
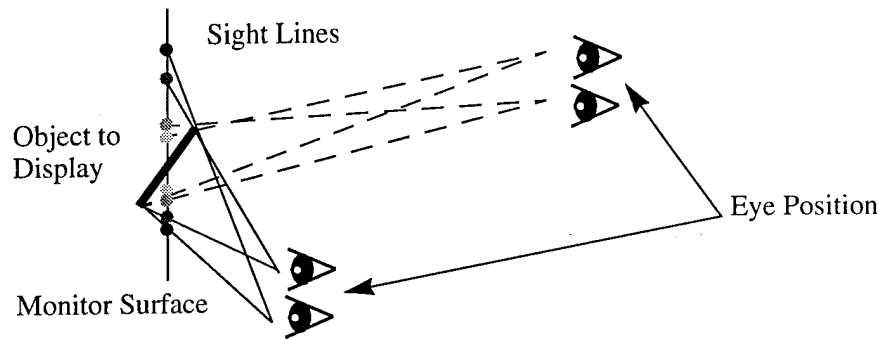
Figure 4-4: The stereoscopic projections are used to determine what pixels to illuminate as a function of the object to display and the user's head position. The lines that intersect the monitor are the pixels that are illuminated for each eye position.
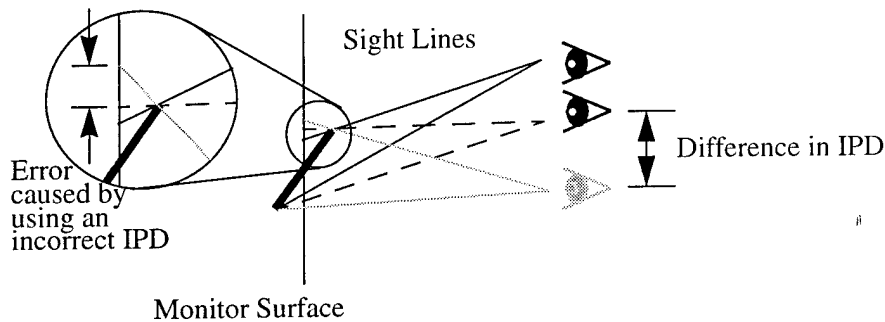

Figure 4-5: The effect of using the wrong IPD in the stereo projections.

4.2 3D Reconstruction Vs. Slices in Volumetric Data

In the medical domain, surgeons more often view slices taken through a volumetric data set, then 3D reconstructions of the imagery. It is still an open question whether or not slices from the volumetric imagery are the best way to present overlay information with an augmented reality system, or if 3D reconstructions should be used instead. It is possible to present both types of overlays; however, the question remains as to which one will provide the best guidance and be of the most use to the surgeons.

To determine the best imagery for each application, a number of tests must be performed. One such test currently being proposed will present both standard slices taken from CT and a 3D reconstruction from the same data set. When the using the slice imagery, the system will present an interface to the user to allow the user to position the slice anywhere in the volumetric data. The user will be asked to perform a number of mock surgical procedures using the slice and reconstruction overlays. The ability of the user to accurately perform the task with minimal error as well as the time required will be recorded and compared between the two approaches. The interface that offers the best accuracy to the user will be selected for the task at hand.

28

4.3 Accuracy Test

Knowing the overall accuracy of the system is of the utmost importance. If a surgeon is to use the system for navigational or localization, the system should be above a required accuracy threshold [23]. The accuracy threshold should be set such that errors below this threshold are still acceptable for safe operation. Accuracy is determined by the position of the overlaid guidance information (due to registration errors, accuracy in the systems, 3D reconstruction errors, etc.), Each procedure that uses the augmented reality system can potentially have its own accuracy threshold depending on the accuracy necessary so that injury will not occur. If the accuracy threshold is set much higher than necessary for the case to be performed safely, the cost of the system will be unnecessarily raised by requiring higher accuracy trackers, better displays, and higher resolution medical scans.

4.4 Cadaver Study

Once a system has gone through all of the above tests, the final test that should be performed with the system before it is used in a clinical setting is a cadaver test. Cadavers allow for the most clinically realistic possible test before the system is used on a patient. With a cadaver, an actual clinical procedure that will be performed with the augmented reality system can be tested. A cadaver study is important to help identify a number of complications that can develop when trying to work with real tissues in an OR as opposed to working in a controlled lab environment. For example, the process of building 3D reconstructions from real tissue is more difficult then from simple phantoms. Discontinuities in the boundary between the tissues can greatly complicate the 3D reconstruction process. By performing the actual clinical case on a cadaver the hope is that these potential problems will materials before an actual patient case is performed.

5 Future Work and Conclusions

5.1 Future Work

With a number of errors identified in the current system, it will be possible to design and build the next generation of the MRCAS system such that the errors are reduced as much as possible. The current goal is to build the new system for practical use in a clinical environment. The medical procedures currently being targeted will be lesion localization to assist in the incision planing and removal of the tumor. When the new system is completed, the same set of experiments must be run to determine if the major sources of error were found and eliminated. The hope is to have a system running and fully validated on cadavers by the end of the summer of 1996. After this time, a number of human trials will be performed to determine the specific clinical acceptance of the systems.

One of the intrinsic problems with all augmented reality systems is that once the initial incision has been made, it is possible that internal shifts of the soft tissue will cause the imagery to no longer correctly represent the actual tissue geometry. If the imagery used by the augmented reality system is not updated in any way, the guidance information will be inaccurate if a soft tissue shift has occurred. One way to circumvent this problem is to constantly update the imagery. New magnetic imaging systems are able to accomplish this. The incorporative MRI scanners allow for imagery to be collected at around 1Hz while the surgeon is operating. The imagery collected during the surgery can then be used to constantly re-build the accurate representations of the soft tissue. The problem with this approach, however, is that the augmented reality system must be compatible with a high magnetic field (on the order of 0.5T up to 1.5T). The hope is to develop a new MRI compatible augmented reality system of some type that will draw from the experiences learned with the clinical experiment performed in 1996. The new MRI compatible augmented reality system is planned to be operational by the end of 1997.

5.2 Conclusions

Augmented reality is a technique for combining supplemental imagery to a scene such that it appears as part of the scene and can be used for guidance, training, and locational aids. In the medical domain, augmented reality uses visual medical imagery to offer a physician a direct relation between the imagery and the patient. The many different applications of augmented reality systems in medicine are just beginning to be realized. Before medical augmented reality systems become widely accepted, solutions must be found for some basic technological hurdles, including patient-image registration, patient tracking, stereo display techniques and display devices. Some solutions to these hurdles have been offered by groups currently building medical augmented reality systems; however, none of these systems have offered complete solutions to all of these problems. The MRCAS augmented reality system was developed as an experimental prototype system to test different solutions to these hurdles in order to overcome some limitations with the current systems. Before any augmented reality system can be used in a clinical situation, the accuracy and effectiveness of the system's supplemental imagery should be found. If systems can not be designed that offer intuitive and accurate supplemental imagery to the user, they will be of little use. When solutions can be found to these problem areas and the areas can be brought together to work as a single system, great advances can be made to medical augmented reality systems.

Bibliography

[1] R.B. Apte and D.M. Bloom. Deformable grating light valves for high resolution displays. In *Hilton Head Conference*, pages 1–6, 1994.

[2] Julian Bailes. Neurolink cost savings, 1994.

[3] Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *Computer Graphics*, 26(2):203–210, July 1992.

[4] M. Blackwell, R.V. O'Toole, F. Morgan, and L. Gregor. Performance and accuracy experiments with 3d and 2d image overlay systems. In *Second Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, pages 312–317. Wiley, November 1995.

[5] Bruce Davey, Patrice Munger, Roch Comeau, Laura Pisani, Andre Oliver, and Terry Peters. Applying stereoscopic visualisation to image-guided neuosurgery. In *First Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, pages 264–271, September 1994.

[6] Michael Deering. High resolution virtual reality. In *Computer Graphics*, volume 26, pages 195–201, July 1992.

[7] P.J. Edwards, D.J. Hawkes, D.L.G. Hill, D. Jewell, R. Spink, A. Strong, and M. Gleeson. Augmented reality in the stereo operating microscope for otolaryngology and neurosurgical guidance. In *Second Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, pages 8–15. Wiley, November 1995.

[8] A.N. Goldsmith. Three dimensional television system. *US Patent 3,74,921*, 1972.

[9] W.E.L. Grimson, G.J. Ettinger, S.J. White, P.L. Gleason, T. Lozano-Perez, W.M. Wells III, and R. Kikinis. Evaluating and validating an automated registration system for enhanced reality visualization in surgery. In *First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, pages 3–12. Springer, April 1995.

[10] W.E.L. Grimson, T. Lozano-Perez, W.M. Wells III, G.J. Ettinger, S.J. White, and R. Kikinis. Automated registration for enhanced reality visualization in surgery. In *First Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, pages 82–89, September 1994.

[11] Stephane Lavallee, Philippe Cinquin, Richard Szeliski, Olivier Peria, Ali Hamadeh, Guillaume Champleboux, and Jocelyne Troccaz. Building a hybrid patient's model for augmented reality in surgery: A registration problem. *Computers in Biology and Medicine*, 25(2):149–164, March 1995.

[12] Lenny Lipton. *The Crystal Eyes Handbook*. Stereo Graphics, 1991.

[13] R.V. O'Toole, M.K. Blackwell, F.M. Morgan, L. Gregor, D. Shefman, B. Jaramaz, A.M. DiGioia, and T. Kanade. Image overlay for surgical enhancement and telemedicine. In *Interactive Technology and the New Paradigm for Healthcare*, volume 18, pages 271–273. IOS Press, January 1995.

[14] Diana Phillips. Enhancing brain surgery with 3d models and live video. *Computer Graphics World*, pages 12–13, January 1994.

[15] Kelly PJ, Alker GJ, and Goerss SJ. Computerized-assisted steretactic laser microsurgery for the treatment of intracranial neoplasms. *Neurosurgery*, 10:324–331, 1982.

[16] Tim Poston, Luis Serra, Wayne Lawton, and Beng Choon Chua. Interactive tube finding on a virtual workbench. In *Second Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, pages 119–123, November 1995.

[17] K. Radermacher, H.W Staudte, and G. Rau. Computer assisted orthopedic surgery by means of individual templates - aspects and analysis of potential applications. In *First Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, volume 1, pages 42 – 48, September 1994.

[18] David W. Roberts, Eric M. Friets, John W. Strohbehn, and Taku Nakajima. *The Sonic Digitizing Microscope*, pages 105–111. American Association of Neurological Surgeons, 1993.

[19] David W. Roberts, John W. Strohbehn, John F. Hatch, William Murray, and Hans Kettenberger. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *Journal of Neurosurgery*, 65:545–549, October 1986.

[20] D.W. Roberts, J.W. Strohbehm, E.M. Friets, J. Kettenberger, and A. Hartov. The stereotactic operating microscope: Accuracy refinement and clinical experience. *Acta Neurochirurgica, Suppl.*, 46:112–114, 1989.

[21] J.B. Sampsell. The digital micromirror device and its application to projection displays. In *Transducers*, pages 24–27, 1993.

[22] Christopher Schmandt. Spatial input/display correspondence in a stereoscopic computer graphics work station. *Computer Graphics*, 17(3):253–261, July 1983.

[23] D. Simon, R.V. O'Toole, M. Blackwell, F. Morgan, A.M. DiGioia, and T. Kanade. Accuracy validation in image-guided orthopedic surgery. In *Second Annual Incarnation Symposium on Medical Robotics and Computer Assisted Surgery*, pages 185 – 192. Wiley, November 1995.

[24] Parviz Soltan, John Trias, Weldon Dahlke, Mark Lasher, and Malvyn McDonald. Laser-based 3d columetric display system. In *Interactive Technology and the New Paradigm for Healthcare*, volume 18, pages 349–358. IOS Press, January 1995.

[25] Michihiro Uenehara and Takeo Kanade. Vision-based object registration for real-time image overlay. In *First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, pages 13–22, April 1995.

[26] S. Webb, editor. *The Physics of Medical Imaging*. Adam Hilger, 1988.

[27] Zeiss. Stereotactic systems. *Zeiss Sales Brochure*, 1995.