2009

INTERFACE DESIGN FOR A VIRTUAL REALITY-ENHANCED IMAGE-GUIDED SURGERY PLATFORM USING SURGEON-CONTROLLED VIEWING TECHNIQUES

Jennifer Lo

INTERFACE DESIGN FOR A VIRTUAL REALITY-ENHANCED IMAGE-GUIDED SURGERY PLATFORM USING SURGEON-CONTROLLED VIEWING TECHNIQUES

(Spine title: Interface Design: VR-Enhanced Image-Guided Surgery)

(Thesis format: Monograph)

by

Jennifer Lo

Graduate Program in Biomedical Engineering

Submitted in partial fulfillment
of the requirements for the degree of
Master of Engineering Science

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Initiative has been taken to develop a VR-guided cardiac interface that will display and deliver information without affecting the surgeons' natural workflow while yielding better accuracy and task completion time than the existing setup. This paper discusses the design process, the development of comparable user interface prototypes as well as an evaluation methodology that can measure user performance and workload for each of the suggested display concepts.

User-based studies and expert recommendations are used in conjunction to establish design guidelines for our VR-guided surgical platform. As a result, a better understanding of autonomous view control, depth display, and use of virtual context, is attained. In addition, three proposed interfaces have been developed to allow a surgeon to control the view of the virtual environment intra-operatively. Comparative evaluation of the three implemented interface prototypes in a simulated surgical task scenario, revealed performance advantages for stereoscopic and monoscopic biplanar display conditions, as well as the differences between three types of control modalities. One particular interface prototype demonstrated significant improvement in task performance. Design recommendations are made for this interface as well as the others as we prepare for prospective development iterations.

Keywords: image-guided cardiac surgery; human-computer interaction; minimally invasive surgery; surgical interfaces; virtual reality; perception; information display; 3D image manipulation

Acknowledgments

To my supervisor, Dr. Terry Peters, thank you for your academic advise and personal care. Your dedication to the lab and its research is not only admirable, but truly inspiring.

To my co-supervisor, Dr. Roy Eagleson, and advisory committee, Dr. Stefan Everling and Dr. Gerard Guiraudon, thank you for your guidance and invaluable insight into the every aspect of this research - human factors, perception and clinical application.

To all members of the Peters' lab, thank you for your helpful consultations and countless coffee breaks, without which I may not have survived this research experience. John Moore and Chris Wedlake, I am indebted to you for all your technical expertise and constructive suggestions.

To Chris, thank you for your companionship and encouragement during many late nights of working at the lab. Your love and belief in me motivate me to reach my goals and never stop aiming higher.

To my family, especially Mom, Dad and Jackie, thank you for your patience, support and understanding of the stress I *occasionally* brought home. Thank you for letting me practice my research talks in front of you, and for the many pep talks.

To all the wonderful friends I have made throughout the years, thank you for keeping me sane, humble, and giving me "a life".

Lastly, I would like to thank the Imaging Laboratories at the Robarts Research Institute, the University of Western Ontario, the Ontario Graduate Scholarship program, the Ontario Graduate Scholarship program in Science and Technology, and my supervisor, for the financial support of this research.

Contents

List of Tables

List of Figures

List of Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
4D	Four-Dimensional
ANOVA	Analysis of Variance
ASD	Atrial Septal Defect
AR	Augmented Reality
CPB	Cardiopulmonary Bypass
CSTAR	Canadian Surgical Technology and Advanced Robotics
CT	Computed Tomography
DOF	Degrees of Freedom
FOV	Field of View
HCI	Human-Computer Interaction
HMD	Head Mounted Display
IGI	Image Guided Intervention
MI	Minimally Invasive
MIMVS	Minimally Invasive Mitral Valve Surgery
MRI	Magnetic Resonance Imaging
MS	Mitral (Valve) Stenosis
MT	Magnetic Tracking
MTS	Magnetic Tracking System
OR	Operating Room
PVA-C	Poly(vinyl) Alcohol Cryogel
RMS	Root Mean Square

TEE	Transesophageal Endoscopy
TLX	Task Load Index
UCI	Universal Cardiac Introducer
UH	University Hospital
UI	User Interface
US	Ultrasound
VASST	Virtual Augmentation and Simulation for Surgery and Therapy
VE	Virtual Environment
VR	Virtual Reality
VTK	Visualization Toolkit

Chapter 1

Introduction

1.1 Motivation

1.1.1 Seeing through the body

Minimally invasive (MI) interventions for therapy involve targeting organs inside the body without exposing the organ, as with traditional 'open' surgery methods. In the surgical sense, the term 'minimally invasive' refers to accessing the target organ through a small incision called a 'port' or through natural orifices in the body. As a result of operating through a small incision, direct vision to the patients' target anatomy is compromised. In order to visualize the surgical area and surrounding anatomy during intervention, medical imaging technology is necessary. MI procedures which involve the use of imaging are known as image-guided interventions (IGIs).

Applications of IGI exist in cancer treatment, biopsy, radiation therapy, ablation therapy, and surgery. The large number of medical specialties which have benefited from adapting image-guidance for surgical and therapeutic purposes include orthopedics, urology, cardiology, thoracics, and neurology.

IGI is based on the idea of using of three-dimensional (3D) spatial information to plan and/or guide a medical procedure. Using the imaging technologies described in the following section, the surgeon has an enhanced ability to 'conceive path, pose and orientation in six degrees of freedom' [1] without making any surgical incision.

Although the concept of 'seeing through' the patient is meant to alleviate the difficulty in navigating towards targets located *inside* the body, the augmentation of reality introduces a new, and sometimes unfamiliar, level of interaction for the surgeon. Whether the procedure is performed using laparoscopic or robotic techniques, the display and manipulation of medical imagery has a direct impact on the performance of the surgeon.

This introductory chapter outlines the motivation for using MI image-guided techniques to perform certain types of cardiac surgery, and describes the virtual reality (VR)-enhanced ultrasound (US)-guided cardiac surgery platform being developed at the lab for Virtual Augmentation and Simulation for Surgery and Therapy (VASST) within the Robarts Research Institute. Towards the end of this chapter, a brief overview of the current developments and challenges faced when designing surgeon-to-computer interaction technology is given to illustrate the significance of this thesis.

1.1.2 Minimally Invasive Cardiac Surgery

Compared to other surgical disciplines, cardiac surgery has been behind in its acceptance of MI methods. Up until the mid-90s, cardiac surgeons were still employing invasive techniques, such as the full median sternotomy, which is a large incision (6-10cm) along the sternum, as well as cardiopulmonary bypass (CPB) and associated aortic cross-clamping and cardioplegia [2]. Once the action of surgery on cardiac physiology was better understood, the inception of MI methods in 1995 [3] marked a new era in cardiac surgery.

When applied to cardiac interventions, the term 'minimally invasive' could either suggest a smaller incision, the use of a partial or minithoracotomy rather than a full sternotomy, or the avoidance of CPB. MI cardiac interventions can include procedures such as single [4] or two-vessel revascularization of the left anterior descending or diagonal coronary artery (as in coronary bypass grafting) [5]; valve repair [6] or replacement [7]; and atrial septal defect closure [8]. MI cardiac techniques can be conducted using robotic, laparoscopic or percutaneous techniques [2]. All of which require enhanced visualization and specialized instruments.

Within the past two decades, many MI cardiac techniques have achieved a performance level comparable to, and in some cases better than, traditional surgery [2]. While some techniques have resulted in lower hospital costs and reduced post-operative complications and surgical trauma, there are others that need clinical refinement. Due to their novelty, improvements in the areas of patient selection, procedure definition, instrument design and proper surgeon training are needed to determine the true benefits of MI cardiac surgery. Acceptance of any new surgical approach not only requires careful analysis of its advantages and disadvantages to patient outcome and its cost benefits, but also requires that surgeons' can adapt to the approach without compromising years of accustomed workflow and surgical training [9].

Two types of cardiac surgery are discussed in which novel VR-enhanced surgical techniques have been applied. The topic of this thesis is motivated by the following surgical procedures.

1.1.3 Mitral Valve and ASD Repair Surgery

1.1.3.1 Mitral Valve Stenosis

The mitral valve annulus is the flexible junction of fibrous and muscular tissue, joining the left atrium with the left ventricle, and anchors the hinges of the anterior and posterior mitral valve leaflets [10]. The function of the mitral valve is to control blood flow through the left side of the heart. Figure 1.1 shows the relevant structures of the heart.

Mitral valve stenosis (MS), or narrowing of the mitral valve orifice, is generally a result of rheumatic heart disease; however, it can also be caused by congenital mitral valve deformities, malignant carcinoid syndrome, neoplasm, left atrium thrombus, metabolic diseases and previous commissurotomy or implanted prosthesis [10]. Stenosis is almost always acquired before the age of 20 but does not clinically manifest until one to three decades later [10]. In its early stages, MS is a slow, continuously degenerate process that rapidly worsens once symptoms appear. If left untreated, MS can cause progressive heart failure in 60-70%, systemic embolism in 20-30%, pulmonary embolism in 10%, and infection in 1-5% [12]. As a result, MS is one of the most common reasons

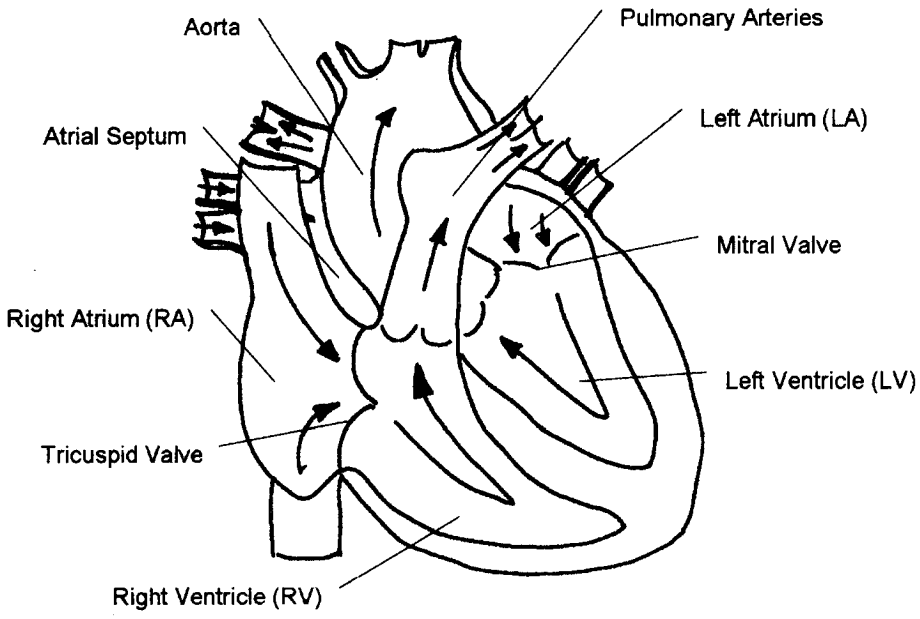

Figure 1.1: Diagram of the heart. Image adapted from the Lucile Packard Children's Hospital website [11].

to undergo valvular operation.

1.1.3.2 Atrial Septal Defect

Atrial septal defect (ASD) is a congenital heart disease most commonly caused by a patent foramen ovale [13]. An ASD of the patent foramen ovale type affects approximately 25% of the normal population and accounts for 5-10% of all coronary heart disease [13]. ASDs occur sporadically as a result of spontaneous genetic mutations; however, hereditary forms have been reported [13]. The amount of left-to-right atrial shunting caused by the ASD varies among affected individuals. Symptoms of this disorder, typically manifested in the third or fourth decade of adult life, involve atrial flutter or fibrillation, effort dyspnea, fatigue and in rare cases, heart failure [14]. For those with large annular dilation, the volume overload occurring in the right ventricle can cause tricuspid valve regurgitation [15]. To avoid the complications that can arise from the development of these sequelae, closing the defect upon diagnosis is preferable.

1.1.3.3 Minimal Incision Approaches

Two surgical options for treating MS are mitral valve comissurotomy (repair) or valve replacement. Deciding which of the two methods to perform depends on the intra-operative analysis of the entire valve pathology [16].

Previously, the assessment could only be performed intra-operatively on the open heart through a full median sternotomy. Due to evolving medical imaging technologies, surgeons are able to view the heart and plan the operation through a small 4-8cm incision usually made between the fourth intercostal space, known as a mini-thorocotomy [17]. With this approach, known as minimally invasive mitral valve surgery (MIMVS), a right anterolateral thorocotomy is the most common incision although left posterior thoracotomies are also possible [18]. Patients experience shorter hosiptal stays, faster resolution of postoperative pain and quicker return to normal activity [19]. MIMVS is evidenced to reduce surgical trauma for patients as well as hospital costs [9]. Cohn *et al.* concluded a reduction in overall hospital charges by 20% [20]. There is also a lower risk of perioperative mortality and wound complications in MIMVS compared

to sternotomy procedures [21]. Nevertheless, 'closed-chest' mitral valve procedures still require the use of CPB.

The typical operative closure of an ASD involves CPB and moderate hypothermia. In minimal incision procedures, a small right anterolateral thoracotomy or median sternotomy can be used to gain surgical access. The ASD can be closed using either primary suture closure or a patch implantation [15]. The surgical outcomes of minimal incision closures and full sternotomy methods both exhibit low (0-1%) mortality and morbidity rates [8] [22]. In addition, cross-clamp and CPB times are equivalent in both groups [8]. Ryan *et al.* [22], however, reports a significantly shorter hospital stay with minimal incision ASD closures. Given such comparable outcomes, the advantage of a shortened hospital stay and improved cosmesis, have popularized MI repairs over full sternotomy. Similar to MIMVS, these minimal incision ASD closures also require the use of CPB and cardioplegia.

1.2 Objective

1.2.1 Off-pump Beating Heart Techniques

CPB, known as the heart-lung machine, provides blood oxygenation and pumping to the body in the absence of functioning heart and lungs. The heart can then be arrested by infusing a cardioplegic solution and drained of its blood using aortic cross-clamping [23]. In doing this, physiological complications can arise such as the perforation of vessels or cardiac chambers, aortic dissection, incomplete de-airing, and systemic air embolism [24]. Prolonged CPB usage exposes the patient to risk of mortality (2-5%), stroke (2%), transfusion (30-90%), atrial fibrillation (30%) and neurocognitive dysfunction (50-75%) [25]. Introducing CPB to high-risk patients, such as those with medical comorbitities, can increase those risks [17].

Off-pump, beating heart cardiac interventions are desirable as they eliminate the need for a CPB and cardioplegia, reducing the risk of such complications. In response to this need, the VASST laboratory has developed a VR-guided cardiac surgery system designed to allow a surgeon to replace a defective valve or repair an ASD through

off-pump beating heart techniques. However, because beating heart surgery occurs within a closed-chest, direct view to the heart is compromised.

As with other minimally-invasive procedures, vision can be provided indirectly via scopes inserted into the 'ports' or through natural orifices. For valvular surgery, modern methods include video-assistance with a thoracoscope [19] or endoscope [26] [27] as well as video-direction with an Aesop 3000 voice-activiated camera robot (Intuitive Surgical, Inc. CA). However, transesophageal echocardiography (TEE) is still the primary form of visualization for both valve and atrial septum surgery as it is a relatively inexpensive and ubiquitous tool in the operating room (OR).

The major drawback to TEE is that the two-dimensional (2D) US images of surgical tools and anatomical targets are extremely difficult to interpret even for experienced clinicians [28]. By fusing intra-operative US imaging with 3D models of patient anatomy, the VR-guided cardiac surgery system is designed to allow a surgeon to replace a defective heart valve or repair an ASD through off-pump MI techniques.

1.2.2 Virtual Augmentation

By definition, 'Augmented Reality' (AR) is an enhanced view of the physical environment through the use of superimposed virtual or simulated cues. The complement to 'Augmented Reality' is 'Augmented Virtuality', in which real images are superimposed onto a virtual environment [29]. The VR-enhanced MI cardiac surgery system is an example of 'Augmented Virtuality' since real-time TEE images are merged with VR models of the patient's heart. In other words, the system is augmenting 3D virtuality with real-time 2D images for surgical guidance.

Extending this visualization platform further, the VASST laboratory also incorporates imaging data from multiple modalities, such as 3D US, x-ray fluoroscopy, computed tomography (CT), magnetic resonance imaging (MRI) and endoscopic video. Using these high quality medical images, the VASST laboratory is exploring methods for generating 3D pre-operative patient models from volume and surface rendering techniques as well as creating a fused imaging environment to assist with surgical planning and intra-operative guidance.

The latest system developments include the optimization of volume rendering using graphics processing unit-based programming [30]; the use of gated four-dimensional (4D) US image acquisition to reconstruct a 3D US image [31]; the creation of dynamic surface models specific to patient data [32]; and the registration of 3D VR models to stereo-capable endoscopic video [33].

My thesis contributes to the progress of the cardiac intervention platform by focusing on the design and construction of a display and interaction modality suitable for clinical implementation. The ideal user interface (UI) should optimize user performance while minimizing the impact on surgeon workflow and workload. In an effort to achieve this end goal, an initial design strategy and understanding of design considerations through empirical user evaluation must be instituted. My research establishes a rigorous strategy to inform the design of such an interface, and it describes the development and evaluation of prototype solutions based on the resulting UI requirements.

1.3 Virtual Reality-Enhanced Image-Guidance

To understand the engineering behind the VR-enhanced image-guidance platform, the following system overview is provided.

1.3.1 Framework for Image-Guided Interventions

Medical imaging can be used pre-operatively, intra-operatively, or post-operatively for diagnoses, planning, guiding and assessing a procedure. Pre-operative images of the anatomy can be acquired using x-ray CT or MRI. Intra-operative surgical guidance is commonly achieved using an endoscope, laparoscope, US or C-Arm fluoroscopy.

The typical method of setting up an IGI involves the registering of the intra-operative information to pre-operative data. When implemented accurately, IGI can be advantageous for surgical planning, targeting and navigation in almost all MI interventions.

The general framework for any IGI can be described as follows [1]:

1. Gather pre-operative data including target information
2. Localize and track the position of the surgical tool
3. Register the physical volume location with the pre-operative image data
4. Display the position of the tool and target(s) in terms of the important anatomical structure visible in pre-operative image data
5. Account for differences between pre-operative data and intra-operative reality

1.3.2 Pre-and Intra-operative Imaging Data

The visualization environment presents medical data in the form of pre-operative models and intra-operative real-time images.

Pre-operative surface models can be reconstructed from MRI and/or x-ray CT images. For cardiac tissue, the superior soft tissue contrast and high resolution of MRI enable surgeons to 'see through' blood to decipher morphological landmarks and visualize heart function. As a result, this technology is optimal for the diagnosis of cardiac pathologies, selecting patients and assessing surgical outcomes [10]. Due to its ability to provide excellent soft tissue contrast, MRI data are optimal for segmenting cardiac images, extracting anatomical features of interest and effectively, generating 3D surface models. The segmented surface models and anatomical targets can be combined with intra-operative images for surgical planning and/or guidance.

In comparison, x-ray CT images produce high quality images of coronary arteries (angiograms) which are ideal for identifying and quantifying calcification. In cardiac surgery, CT can be used for identifying at-risk patients, evaluating cardiac function, assessing surgical outcome and planning the surgical approach - especially under re-operative circumstances [34].

Intra-operative cardiac images are typically obtained through US using TEE or endoscopic video. In TEE, the US tranducer is inserted into the esophagus to provide real-time, non-invasive imaging of the beating heart as well as to quantify blood flow

and overall cardiac performance [35]. For cardiac surgery, endoscopic video cameras can be inserted through the 'ports' on the patients' body to provide video-assistance.

The visualization platform, at the VASST lab, can combine pre- and intra-operative cardiac data as volumes, 3D surfaces or 2D slices to enhance vision of the heart anatomy.

1.3.3 Surgical Apparatus and Tracking

When applied to cardiac surgery, specialized surgical tools are designed to exploit the VR-enhanced image-guidance platform. Accessing the beating heart via mini-thorocotomy, a device known as the Universal Cardiac Introducer® (UCI) [28] can be attached to the left atrial appendage through this opening to seal off the surgical area and prevent blood loss (shown in Figure 1.2). The UCI is designed to provide safe port access into the chambers of a beating heart for up to four surgical tools. By suturing the insertion cuff of the UCI onto the appendage, an 'air-lock' mechanism is created that prevents blood from leaving the heart cavity. The other end of the cuff is then attached to the introductory chamber which can accommodate up to four instruments as shown in Figure 1.3. In valve or ASD repair procedure, these instruments are the valve or patch insertion tool, a pressure line, and a fastening tool depicted in Figure 1.4 [36].

Two methods used to track surgical tools are optical and electromagnetic tracking. While optical tracking has more reliable and stable accuracy, it requires an unobstructed line-of-sight between the tracked tool and camera sensor [37]. Magnetic tracking (MT) is more practical in situations where instruments, such as endoscopes, US transducers and catheters, must be tracked within the body. Unfortunately, MT is disturbed by the presence of ferro-magnetic objects in the environment.

For the VR-guidance system setup, a magnetic tracking system, the Aurora® Northern Digital Inc.) is used to track the four instruments. This system identifies the position and orientation of tiny 6 Degrees-of-Freedom (DOF) sensors at a rate of 40Hz. By attaching the sensors to the surgical instruments, the tools can be tracked relative to a defined reference body. Aurora® generates a magnetic field, using alternating current, with cubic dimensions of 500mm x 500mm x 500mm. Within this cubic

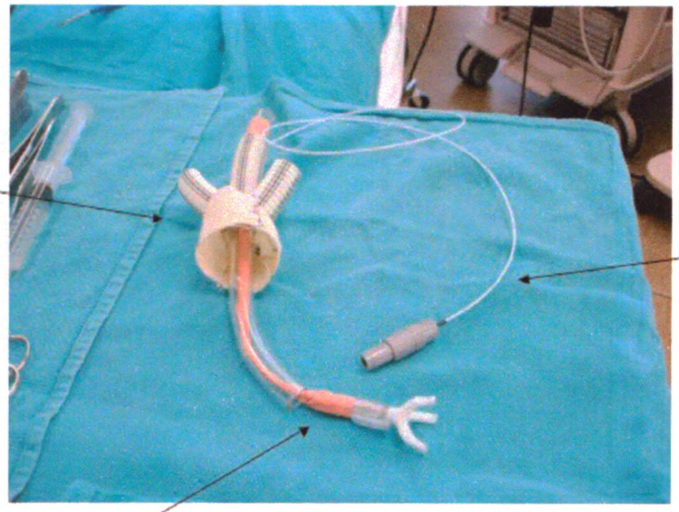

Figure 1.2: The Universal Cardiac Introducer® (UCI). The insertion cuff is not displayed here

---

volume, the Root Mean Square (RMS) positional error is < 0.6mm error and the RMS angular orientation error is < 0.4 degrees [38].

1.3.4 Image Registration

In mixed-reality environments, multiple imaging modalities can be merged together and much research has been spent on investigating the robustness of various image registration algorithms [39]. In addition to merging multi-modal images, the VR system can also incorporate computer generated images into the visualization space. To ensure a reliable VR-enhanced imaging environment, virtual models must be accurately registered to both the pre-operative and intra-operative data.

The two methods for registering the virtual models with physical anatomy are landmark or feature-based registration. Landmark registration works best for rigid bodies such as bone because fiducial points can be stabilized in place. Because the heart is a deformable organ, feature-based registration, is more suitable. To conduct

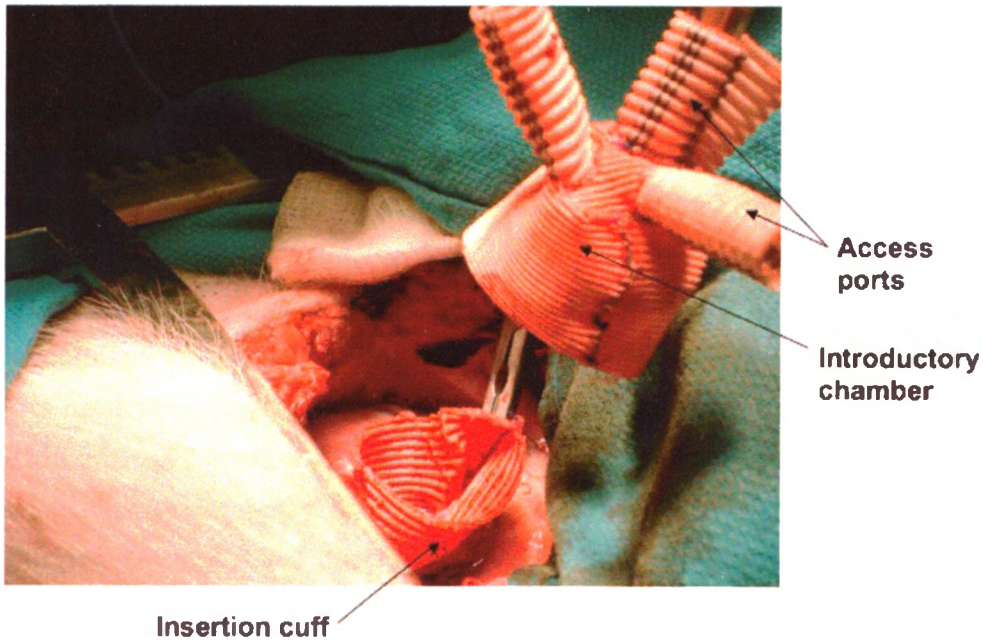

Figure 1.3: The process of attaching the introductory chamber end of the UCI to the insertion cuff (shown appended to the left atrium of the heart). This attachment creates the 'air-lock' mechanism.

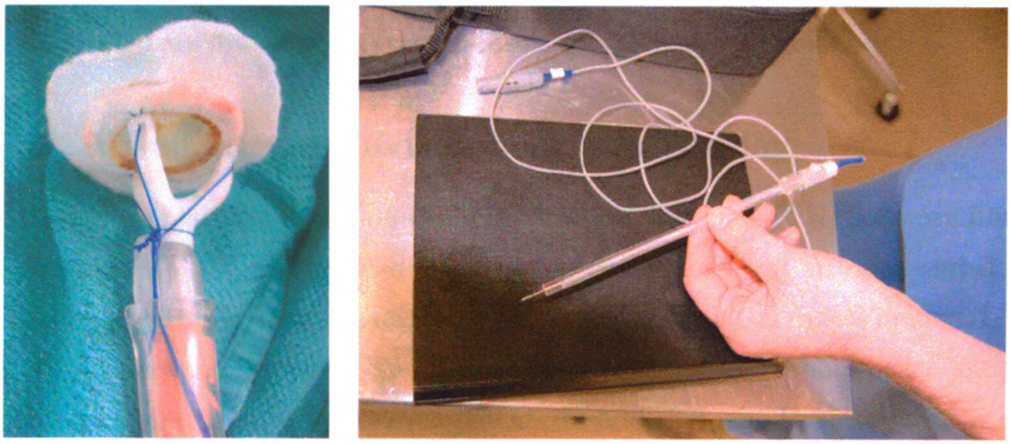

Figure 1.4: *Left*: The valve or patch insertion tool. *Right*: The fastening tool.

feature-based registration, the target valve annulus (aortic or mitral) is first identified in the pre-operative MRI volume. Second, the actual valve annulus is located in the intra-operative US images from a calibrated TEE (see Appendix A on details regarding TEE calibration). Physical reference points on the valve annulus are located using 2D US imaging. By identifying the location of the tracked tool tip, corresponding virtual points can be selected.

The physical and virtual annulus are then aligned using a registration algorithm that minimizes the centroid distances and tip of the valves' corresponding normal unit vectors until an optimal fit is attained [36]. Results for a typical registration produce an RMS error of 5.2mm, 4.1mm, and 7.3mm in each of the surrounding anatomy, left ventricle, left atrium, and right atrium and ventricle [40].

Surgical tools, tracked in real-time, can be also superimposed onto this hybrid environment. Because the magnetic field generator has an intrinsic coordinate system, a fixed reference body is needed to ensure tracking accuracy regardless of where the magnetic field is placed (see Appendix A). Tool positions and orientations must be defined by a coordinate system that is relative to a fixed reference body in the physical setting. If a physical representation of the imaged environment is available, the corresponding images can be loaded onto the software platform and calibrated to fit the tool coordinate system through a landmark registration using physical markers. The calibrated environment can then be used for navigational guidance.

1.3.5 Visualization Environment and Display

A software platform based on the Visualization Toolkit (VTK) software library, the AtamaiViewer®, has enabled the ongoing development of an open-ended and module-based visualization system. This platform is portable across Windows, Linux, and OS-X operating systems. The user interface is coded in Python for the purpose of quick and simple prototype development. In addition, the Python script can be easily transposed to C++ classes for speed optimization. Over the past 6 years, this software platform has been utilized for various image-guided neurosurgical [41] [42], cardiac [36] [32] and abdominal applications [33].

The AtamaiViewer® is capable of rendering a stereoscopic, mixed reality visualization environment to include pre-operative MR or CT images, intra-operative US or endoscopic live feedback, tracked tools and computer generated surfaces or volumes. These components can be used in any combination and rendered using methods such as 3D clipping, colour and opacity adjustments, perspective manipulation, visualizing data along any oblique plane, etc.

VR environments can be immersive or non-immersive. An immersive environment is an egocentric VR setting in which users experience the VR environment first hand. Displays are placed in front of the users' eye as opposed to the monitor-based (non-immersive) setup.

In actual implementation, the final mixed reality environment is displayed to the surgeon on a 2D computer monitor positioned above the surgical operating table. The monitor is placed at 45° above the surgeon's horizontal line of sight and at a 15-30° vertical offset from the surgeon's forward stance. This monitor-based setup is non-immersive and non-stereoscopic (i.e. monoscopic). Although there are VR display alternatives such as commercially available head-mount displays and stereoscopic monitors, many are application-specific models developed by various research groups [43] [44] [45]. Design choices between factors, such as immersive vs. non-immersive, see-through vs. non-see-through and even stereoscopic vs. non-stereoscopic, must be made when considering the optimal method of display.

1.3.6 Accuracy and Validation Studies

Although various image-guidance systems have been attempted for soft tissue organs, there are still many obstacles to overcome. One difficulty lies in the motion and deformations of the organs. Research efforts are directed towards improving the accuracy of registration, which is generally assumed to be a rigid body. Currently, one way to address this challenge is through the use of predictive models to anticipate deformation [46].

Acceptance of image-guided technology depends on registration accuracy as well as tracking accuracy. State-of-the-art magnetic tracking systems such as the NDI

Aurora® (NDI, Waterloo, Canada) can guarantee positional error of 0.96±0.68mm at 50mm distance from the sensor, and 2.72 ±1.8mm at 150mm [37]. Optical tracking provided by the Polaris® system (NDI) can yield a tracking accuracy of better than 0.5mm for a pyramidal volume of $1.627m^3$ [47].

The VR-enhanced surgical guidance system was tested for its feasibility in both phantom and live animal studies [40]. In a simulated cardiac intervention using a plexiglass heart phantom, the performance accuracy with which an experienced surgeon could guide and secure a valve prosthesis to the correct target location using 2D US guidance versus hybrid US-VR guidance was investigated. While all attempts to dock and secure the valve prosthesis using 2D US guidance alone were unsuccessful, the surgeon performed with a 100% success rate when using the hybrid US-VR system. By complementing a 2D US image with surrounding virtual anatomy, the location of the valve could be identified and targeted with ease.

This VR-enhanced surgical system was then implemented in an actual OR environment using an *in vivo* intervention on porcine subjects. In a 'live' case, the use of real-time 2D US was imperative for verifying of final prosthesis placement as the targets within the heart shift with cardiac motion. Although implementation in the OR demonstrated system feasibility, observations during the *in vivo* trial uncovered challenges for future clinical integration. Among these were the determining of tracker-to-tool position range to attain the maximum tracking accuracy [40]; the handling of registration and target location shift errors caused by natural cardiac and breathing motions; and the designing of an optimal method for displaying the visualization environment. These latter issues are addressed in this thesis.

1.4 Design of Surgical Interfaces

For a modern surgeon living in a technologically driven age, there is an increasing amount of information to process and a number of transient media for which this information can be delivered. Surgical robotics, stereoscopic AR displays, interactive speech and gesture control systems are a few of the many examples of novel surgical interfaces. These developments, and their associated human factors issues, are

summarized below.

1.4.1 Surgical Robotics

At the present moment, the only existing FDA approved surgical teleoperator system is the daVinci® System (Intuitive Surgical Inc., Mountain View, CA). Since 1999, this surgical robot has been used to perform over 500 general surgery procedures, in addition to routine cardiac operations such as closed-chest endoscopic coronary bypass grafts and open heart mitral valve repairs and replacements, ASD repairs and tricuspid valve repairs [48].

The daVinci® System is designed to overcome human factor limitations faced by surgeons performing laparoscopic procedures. With the daVinci®, surgeons sit at a master console and control telemanipulators whose motions are coupled with computer-controlled endoscopic instruments. The endoscopic video image of the surgical field is projected to the surgeons' console and is registered spatially with the telemanipulators. Because the surgeon is viewing the scene in the same space as his or her hands, natural hand-eye coordination is enabled. In addition, the stereoscopic display provides depth perception which is lacking in conventional 2D endoscopes. The system restores the freedom of motion lost in traditional laparoscopy as tools can be manipulated with 7 DOF (3 for orientation, 3 for translation, 1 for grip). Furthermore, the design of the telemanipulators alleviate the reverse motion problem, known as the 'fulcrum effect' which is caused by the hinge formation at the tool's entry point, as well as the shear stress generated by the pivot point of endoscopic tools [49].

Despite these ergonomic enhancements, surgical robots are expensive, not suitable for all patients nor all procedures, require longer operative times discounting the time for setup and removal [50], and are incapable of providing haptic feedback, thus, overloading the surgeons' visual senses.

A comparison of robot-assisted and laparoscopic pyeloplasty demonstrated that the robotic approach had no significant clinical advantage over the laparoscopic methods, yet substantial time and monetary cost were added to the operation [50]. Due to robot depreciation, it is estimated that in addition to the robot consumables, a $2000

premium per case exists for daVinci® use based on 150 cases per year [50].

1.4.2 3D Augmented Reality Displays

While 2D visualization is a simple, straightforward method for displaying medical data, 3D and 4D visualization approaches are gaining popularity due to the rich information contained in volumetric medical datasets and their natural resemblance to physical reality. Ironically, whether image, volume or signal, 3D visual information is typically presented on a 2D computer monitor. Displaying 3D data introduces a whole new level of complexity, especially when considering the integration of AR and VR technology. Choosing the optimal display technique requires an understanding of the affected perceptual factors such as visual acuity, contrast sensitivity, field of view, and colour and depth perception [51].

The two major dichotomies of stereoscopic displays are video-based and optical. Video-based stereo displays use rendering engines to fuse virtual and real worlds and present them on stereoscopic monitors, whereas optical systems present computer generated images within the observer's view of the real environment. Video-based displays can be classified as either passive or auto-stereoscopic. Passive displays require the user to wear goggles whereas auto-stereoscopic displays rely on the viewpoint of the observer. Optical techniques can be head-mounted, hand-held, spatially aligned and separate from the user, such as with half-silvered mirrors [52].

Stereoscopic head-mounted displays (HMD) are a popular choice for presenting VR and AR information as they offer the possibility of a fully immersive simulated environment. Yet, the narrow visual field and immersive features of HMD make it a controversial method for assisting a mission-critical, collaborative task such as surgery.

When considering optical see-through HMDs, some studies have found improved target localization when projecting stereoscopic AR [45]. Others have demonstrated a notable error in depth estimation due to psychophysical factors introduced by transparent AR overlays [53] [44]. Due to hardware factors such as screen resolution and low colour contrast, prolonged use of the HMD can cause visual discomfort; thus, increasing visual and cognitive fatigue [54]. Users were more sensitive to these

hardware limitations when using optical see-through HMDs as the lack of visual detail is more apparent [55]. Physical comfort is also degraded, despite the ability to maintain a preferred downward frontal gaze. The added physical weight of the headpiece requires extra posture adjustments and increases muscle fatigue [56].

Despite the hardware constraints, there is still merit to evaluating HMDs as they are a unique method of stereoscopic display and the newer models on the market are much lighter, have better colour contrast and are more ergonomic.

1.4.3 Non-Contact Interaction

Human-computer interaction (HCI) modalities such as voice activation, gesture recognition and gaze control are ideal for use in a surgical setting where sterility and minimal hand-to-tool contact are not only desirable but essential for the task.

Voice recognition, a technique that involves speaker-dependant control of a procedure, allows a surgeon to issue pre-defined computer commands by speech. In clinical experiments conducted by Visarius *et al.* [57], a voice-activated system for computer assisted spinal surgery was evaluated for safety and accuracy. Voice templates of six surgeons, among which contained four different native tongues, were stored prior to surgery; all six could be recognized despite noise interference [57]. As a safety precaution, automatic shutdown of the system through the command "microphone" was used in case of accidentally issued commands. In addition, the choice of commands must be free of any phonetic similarities. Advantages for voice-activation in a surgical system are: (1) direct user control with improved learnability and (2) cost-effectivenss and elimination of communication error due to the removal of the system engineer as the intermediary between surgeon and computer [57]. Compared to the speech recognition, voice-recognition systems can acheive a recognition rate of 97% [58] whereas in speech control technology, an overall word recognition rate between 72% and 92.4% is acheived. However, compared to other forms of hands-free input, such as foot control, voice activation was more prone to errors and was less efficient to learn and operate [59].

Gestural control systems have also been developed and tested for usability in

medical visualization systems. According to a study by Wachs *et al.* [60], successful performance of 95-100% can be demonstrated. Grange *et al.* [61] found that the accuracy of a medical gesture-based system can be affected by ambiguous gestures, false triggers from detection of nearby fast moving objects, hand swapping and changes in lighting conditions. In a system that compared both speech and gesture commands for issuing a 'clutch' function, there was a noticeable delay between the time required to speak and recognize a voice command [62]. Under some circumstances, this delay may interfere with short term memory which can increase cognitive strain [62]. For graphic image manipulation, it was found that using both speech and gesture input in combination was more effective compared to using each on its own [63].

1.4.4 Medical Usability

In modern day healthcare, new surgical techniques are demanded with minimally invasive standards. Advances in surgical robotics, stereoscopic AR displays, and interactive speech and gesture control modalities demonstrate the ever-increasing pervasiveness of technology in the OR.

The development of robotics, computers, and VR in medicine are geared towards enhancing less invasive surgical techniques performed by surgeons. However, integration depends on a good understanding of the surgeon's space, task and working mentality.

Through careful user analysis, design and implementation, complex systems will become easier to use and actually accommodate the surgeon's needs, rather than having the surgeon adapt to the technology.

1.5 Thesis Outline

VR-enhanced US guided surgery is a novel approach designed to allow a surgeon to replace a defective heart valve or repair an ASD inside the beating heart via off-pump MI techniques. Components such as registration accuracy, efficient system performance and robust tracking technology are factors responsible for effective clinical integration,

but the final determinant of success is surgeon adoption.

The overall goal is to develop a VR-enhanced US-guided interface that will deliver information seamlessly into the surgeon's natural workflow and that will optimize the surgeon's ability to conduct his or her task.

My research began with the development of a design strategy that involves both qualitative and quantitative methods. Based on my initial results, potential interfaces were built and then evaluated to determine what needs to be done in the next design iteration.

The second and third chapters of this thesis describe the design strategy employed to refine the understanding of the design space, to optimize the choice of design factors, and ultimately to specify an appropriate UI design. Task-centered, domain-centered and user-based studies were conducted. Through *user task analysis* and *participatory design*, the infinite number of combinations and variations of visualization techniques were distilled to specific tasks and clinician requirements. The effect of design parameters on user accuracy and timing are examined through a controlled *user-based* experiment. The combined result of these aforementioned design strategies is a conclusive list of surgeon requirements which can serve as a guideline for developing an effective UI.

The major advantage offered by virtual environments (VE) is the ability to examine data from any pose. To fully exploit the visual benefits of a virtually enhanced surgical platform, the viewer must be able to rotate, pan, and tilt the 3D data as naturally and easily as possible. Currently, surgeons must communicate camera view adjustments to a technician when conducting VR-enhanced US-guided surgery. Like most image-guided procedures, inadequate navigation of the camera by the technician can impair the surgeon's visualization, cause frustration and compromise patient safety [43]. Three user interfaces have been conceived to address these problems. The fourth chapter describes the prototype development of these interfaces, each allowing surgeons to directly manipulate the VE intra-operatively.

The fifth chapter presents a user-centered and comparative evaluation of the efficacy of these three interface prototypes. Quantitative and qualitative results conclude that the optimal surgeon-controlled interface among the three proposed methods is the one

that requires minimal integration of new equipment and makes use of multiple planar views.

The methodology established by this study can be followed for future UI design within the VASST engineering lab. The final chapter discusses the contributions of the presented work, the limitations of the hardware and experimental methodology, and the directions for future work in this area of study.

Chapter 2

Non-Empirical Design Strategies

'The purpose of visualization in image-guided interventions (IGI) is to faithfully represent the patient and surgical environment, and to accurately guide the surgeon to navigate toward, and localize the treatment target during an intervention'

- David Holmes III

Image-Guided Interventions: Technology and Applications [64]

2.1 The Design Methodology

By nature, IGI visualization is interactive, and so an intuitive UI is required to facilitate surgeon performance.

As mentioned in Chapter 1, using US imaging alone for surgical navigation is not as effective as using a VR-US hybrid visualization environment. Merging 2D US imaging with 3D VR models requires accurate registration between images and models, unification of different coordinate systems and correct fusion of real-time and pre-operative data. The feasibility of this VR-US hybrid environment has been validated in an *in vivo* porcine cardiac procedure [40]. Yet, the major challenge for successful transition from laboratory to OR still remains: *How should the VR-US information be presented in order to be relevant, interactive and coherent for the*

surgeon?

A novel approach to IGI, VR-enhanced cardiac surgery alters the users' visualization space, thus affecting conventional tool-to-task interaction and surgical workflow. As an emerging technology, VR-assisted surgical guidance systems have few established design guidelines or interaction metaphors to direct the development of the UI.

Traditional usability engineering strategies such as Royce's waterfall model [65], the iterative spiral model [66], and Hix and Hartson's star life cycle model [67] are respectively, too rigid for novel techniques, geared for software design or assume the existence of preliminary designs or standard heuristics [68]. The waterfall model assumes each design phase must be completed before moving to the next [65]. This is inadequate for most products as designers may not be fully aware of the implementation difficulties by the end of each phase. Using prototypes to crossover between planning and implementation can allow for quick and efficient design refinement. In the spiral model, analysis and engineering occur in cycles - between establishing design goals and validating with the end user [66]. This model is easy to follow for software development; however, this VR-enhnaced surgical interface involves choosing hardware components, making it difficult to revise. The life cycle model is based on prototyping and user evaluation [67]. However, for systems containing new and unfamiliar interactive components, controlled user studies that evaluate these components must be conducted, rather than evaluating the prototypes as a whole.

The domain of surgical interfaces involves hardware, as well as software components. In addition, the use of VR in surgical guidance for beating heart surgery lacks a design predicate. To avoid relying on the developer's best guesses and on small incremental improvements, as would be the case if any of the aforementioned models were applied, the fundamental techniques taught in the field of HCI should be applied as part of a structured methodology.

Gabbard *et al.* [68] proposed a usability engineering methodology which incorporated an empirical user-based study along with traditional task and domain-centered

analysis. I adapted Gabbard's methodology (shown in Figure 2.1) to inform the design of the VR-enhanced surgical interface. Since it was important to incorporate domain expertise during the early stage of development, my model employed a technique known as participatory design to gather an 'expert evaluation' of the concepts. The resulting part-cyclic and part-sequential design process is illustrated in Figure 2.2.

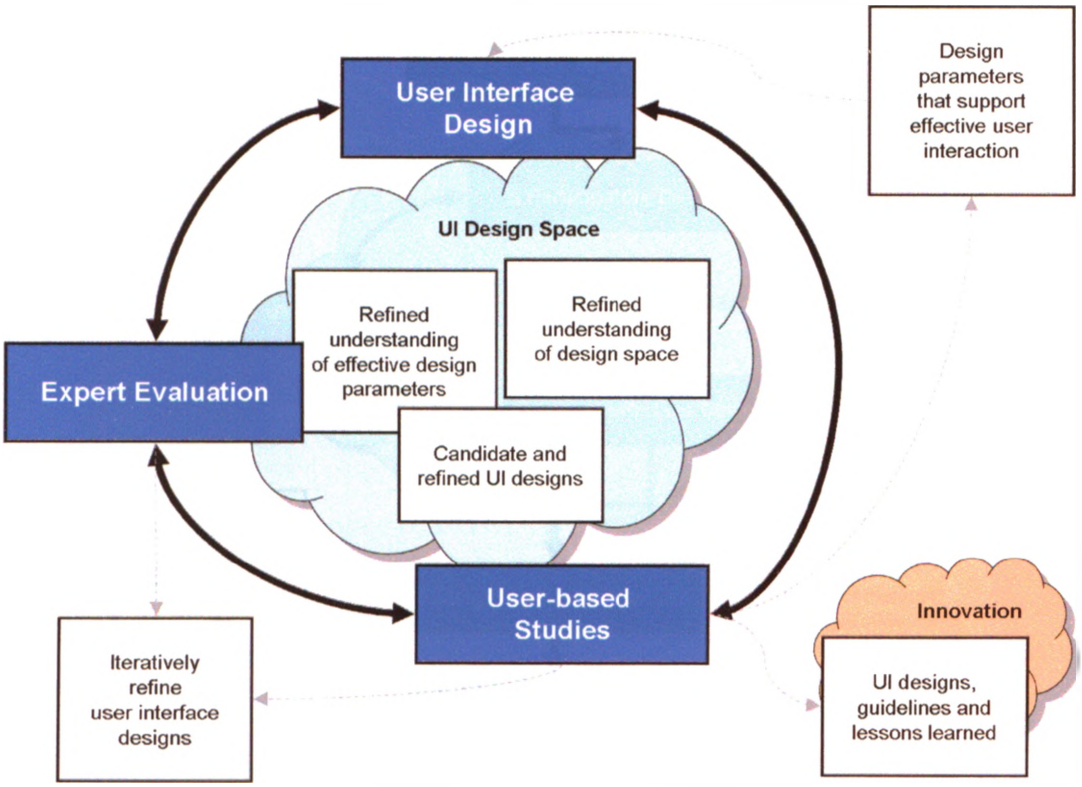

Figure 2.1: Gabbard's design cycle for emerging technologies. Adapted from Gabbard *et al.* [68].

The steps (1) and (2), classified as non-empirical methods, are presented in this chapter. Chapter 3 describes the empirical evaluations embodied in step (3). Chapter 4 presents the initial user interface designs as part of step (4). Chapter 5 discusses the empirical findings in steps (5) and (6).

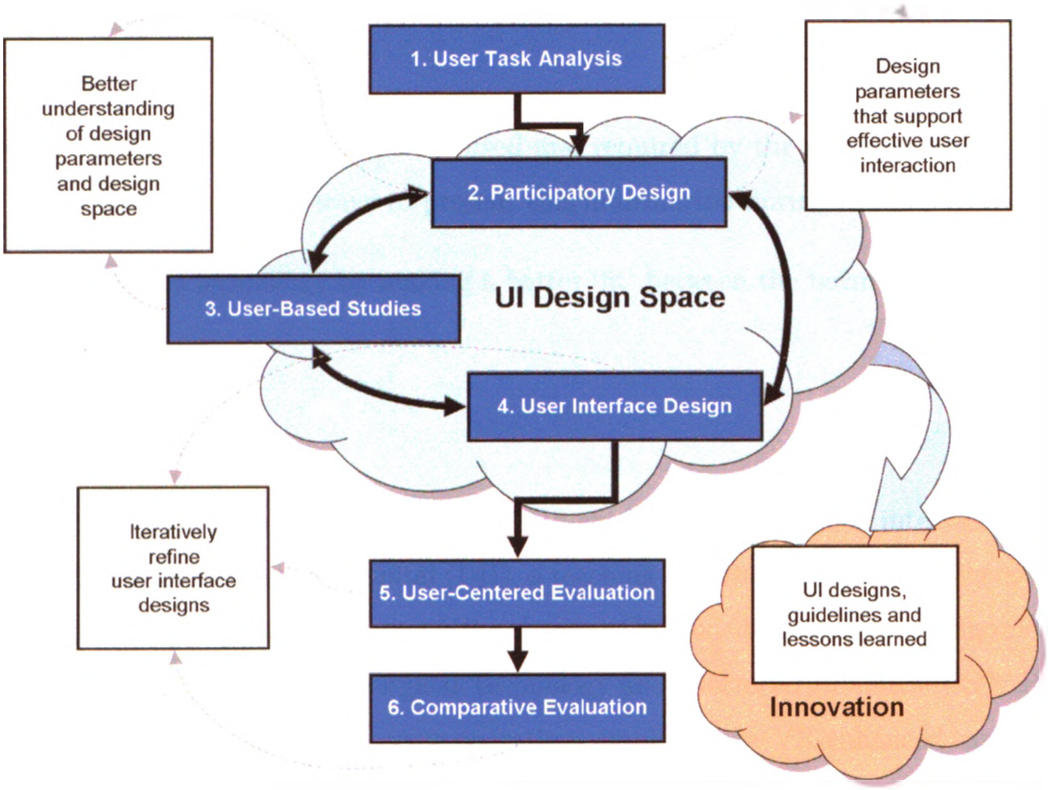

Figure 2.2: Model for my design strategy.

2.2 User Task Analysis

Hierarchical task decomposition, a human factors approach aimed at providing an objective framework for analyzing task and function in a complex system [69] is applied to:

- Design and evaluate the effectiveness of new visualization and manipulation tools
- Understand what information is used and required by the surgical team, and identify the optimum ways to present this information during the procedure
- Improve the technology by making a better 'fit' between the technology and the surgeon

2.2.1 Process

Three repeated observational studies of US-guided patient cardiac interventions were conducted at University Hospital (UH), a teaching hospital within The University of Western Ontario. In addition, two VR-enhanced US-guided porcine valve surgeries were observed at the Canadian Surgical Technology and Advanced Robotics (CSTAR) centre within UH. These acute porcine operations employed the VR-enhanced surgical system for ASD repairs. Observations made in both the US-guided human and VR-US-guided porcine trials form the basis of the hierarchical framework shown in Figure 2.3.

2.2.2 Key Findings

The resulting structure is designed to capture both the high-level goal-directed user requirements for VR-enhanced surgery, as well as the low-level task-to-tool interactions that occur during VR-enhanced valve replacement. These tasks are either performed

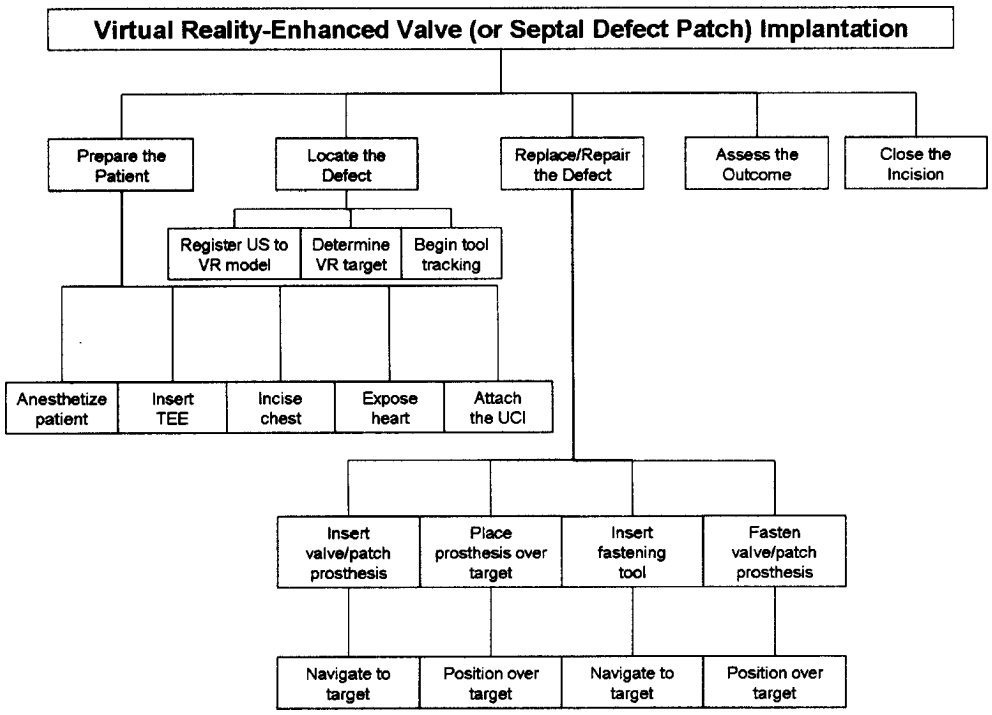

Figure 2.3: Hierarchical task decomposition of virtual reality-enhanced valve (or septal defect patch) implantation.

in sequence, or iterated in parallel. In any case, the bottom-level of the tree signifies the visible accomplishment of it's parent task or goal. The surgical step of interest is the 'Replace/Repair the Defect' during VR surgical guidance. This step can be decomposed into two distinct phases, 'Navigation' and 'Positioning'.

Navigation requires the surgeon to estimate the location and depth of the target and bring the tool, inserted into the heart chamber, towards the target. In most image-guided surgeries, the lack of direct vision and separation of visual and motor information compromises natural hand-eye coordination and increases cognitive load as surgeons are forced to mentally align the disparate reference frames for perception and action [43].

Positioning requires the surgeon to place the tool onto the target with as much accuracy as possible. Naturally, this task requires fine motor skills and precision. In an image-guided MI procedure, the task also requires adaptation to limited vision, the 'fulcrum effect' (where the intended motion is reversed due to the pivoting of laparoscopic instruments around a fixed entry point) and shear stress generated at the pivot point [49]. Surgeons must rely on US imaging for accurate tool-to-target positioning in addition to the VR-environment due to potential misregistration or inability to model a dynamically moving heart.

The hierarchical structure in user task analysis is useful for defining the tasks and subtasks involved in achieving the overall user goal; however, it is incapable of elucidating the human cognitive processes involved in performing each surgical task for specific clinical outcomes. Understanding the users' context can help to expose the human factors underlying each task.

In the next step of my design method, I focus on a domain-centered analysis of the existing system. Perceptual and cognitive limitations as well as potential solutions are discussed with the main users of the technology, the surgeons.

2.3 Participatory Design

As a collaborative process between computer engineers and domain experts, participatory design has been successful in the design of usable VR technology [70]. Three basic aspects of this design strategy are [71]:

1. Defining a goal in which the end-users' quality of life can be improved;
2. A collaborative orientation between system developers and end-users; and
3. An iterative process

2.3.1 Process

Defining a goal

The purpose of developing VR-enhanced surgical guidance technology is to preserve clinician workflow while integrating a new method for improving surgical guidance for conventional image-guided procedures. Visualization of a live organ is typically achieved through 2D US fan images which are difficult to interpret even for experienced surgeons [28]. By providing VR representations of the surrounding anatomy, the surgeon's ability to navigate towards and position a tool onto a target may improve.

Collaborative orientation

A collaborative effort is fostered by the unique working environment at the Robarts Research Institute where engineers, clinicians and other medical professionals interact on a regular basis. As an example, the medical outcome and clinical implementation of the US-VR guidance system has been tested in several acute porcine studies where the engineers are exposed to the interaction and visualization difficulties faced by the surgeons.

Iterative process

To initiate an iterative process, a group brainstorming session was conducted. Four cardiac surgeons and eight members of the engineering and development team participated in this conceptual design meeting. The intent was to identify what information needs to be displayed, and when and how the resulting data should be presented. I utilized storyboarding and scenario development to analyze the current workflow and envision how the technology can be best 'fitted' to the working environment. Scenario development, combined with storyboarding has the ability to resolve detailed specifications for building prototypes [70]. Screen mockups and animations depicting the UI and interaction (shown in Figures 2.4 and 2.5) were created to establish a mutual understanding of the envisioned concepts.

Design concepts

The design enhancements, shown in Table 2.1, can be grouped into two categories: (1) autonomous view control and (2) visual overlays. The first involves the idea of allowing surgeons to control and interact with the 3D display of medical data using specifically designed hardware and software components. The second employs graphical overlay techniques to provide visual feedback cues to augment target distance visualization. Autonomous view control addresses human factor issues concerning mental rotation of 3D imagery, accommodating to camera view adjustments and zooming and panning of data required during the task. Visual overlays explore the use of opacity, shading, and colour. Optical factors such as contrast, acuity, and response to illusions are of main concern. For the scope of this thesis, I first examined the design enhancements in category (1) automonous view control. Future work will address the concepts in (2).

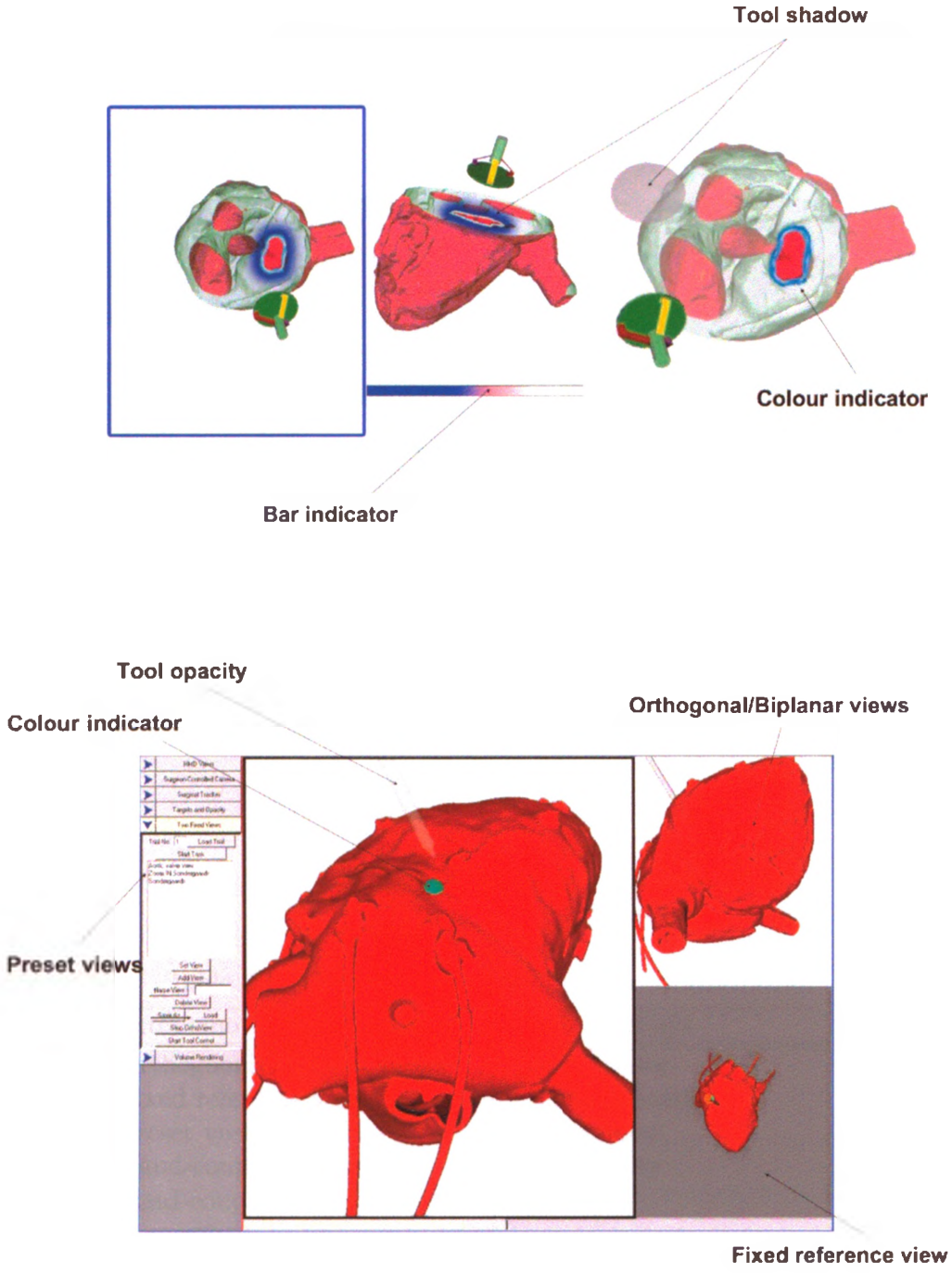

Figure 2.4: Mockups of conceptual interface designs.

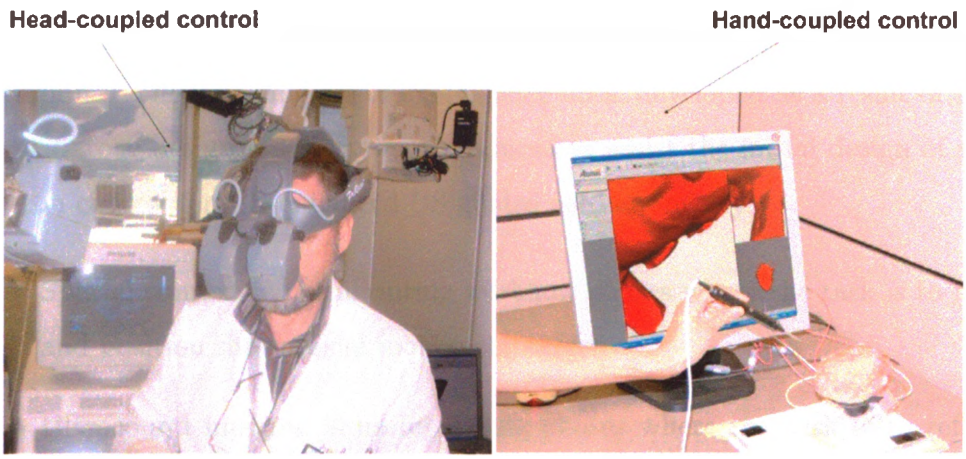

Figure 2.5: Photographs of conceptual interface designs.

Table 2.1: The two categories of design enhancements.

Autonomous View Control	Visual Overlays
-Orthogonal/Biplanar views -Fixed reference view -Preset views -Hand-coupled control -Head-coupled control (immersive)	-Bar indicator -Colour indicator -Tool shadow -Tool opacity

2.3.2 Key Findings

Based on our domain-centered discussion, it is clear that delegating view control requires both effort and good communication between surgeon and technician. If misunderstandings arise, this form of view control can be disorienting and time consuming. The following set of expert guidelines, developed as a result of this first participatory design meeting, were produced to assist the initial design of effective surgeon control for VR-enhanced US guided cardiac surgery.

- Spatial context is useful during workflow phases where the task is to navigate and position the surgical tool on the target;
- Ultrasound imaging is useful during phases where the task is to refine tool position;
- The optimal view of the target is task-specific and dependant on the surgical approach. A head-on view of the target is typically sought during open surgery;
- Surgeons are accustomed to referencing the anatomy using standard "textbook" views;
- A secondary view is required in cases where the target becomes obstructed by the approaching virtual tools;
- A reference view should be provided to help the user recover from getting "lost" in the virtual world

The ability to perceive tool-to-target depth, adjust the view of the 3D virtual environment and establish correct tool orientation throughout navigation and positioning tasks are essential for implementing autonomous view control.

2.4 Summary

Undertaking the challenge of tailoring the infinite combinations and variations of visualization parameters is large and overwhelming. This chapter introduces the strategic approach adopted to decompose the overall problem into smaller and more manageable goals. As a result of steps (1) user task analysis and (2) expert evaluation, suitable guidelines were established to aid the development and evaluation of a UI specific to the VR-guided valve/patch implantation.

Design recommendations have been seperated into two categories based on the type of human factors they emphasize. The first category, autonomous view control has motivated the development of three surgeon-controlled interface prototypes described in Chapter 4 of this thesis. A similar follow up shall examine the second category, visual overlays. However, this second category of design improvements extends beyond the direction of this thesis, and thus will be planned for future work.

Chapter 3, discusses the next step of the overall design strategy: the investigation of VE parameters on user performance through a controlled user study.

Chapter 3

Empirical Design Strategy: A User-Based Study

3.1 Purpose

The goal of the work presented in this chapter was to evaluate targeting performance under the influence of various VE design components using a controlled user-based study. As opposed to full-fledged interface designs, individual design parameters, such as the presence of spatial context, the use of fixed orthogonal view angles, and the use of stereovision, were investigated.

3.2 Task and Experimental Setup

The task was an abstracted version of a mitral valve implantation, involving the docking of a prosthetic mitral valve onto three target 'valve annuli' located within a poly(vinyl) alcohol cryogel (PVA-C) heart box phantom (See Figure 3.1). Readily available in the laboratory, PVA-C was chosen for this experiment due to its deforming property and texture that can provide realistic tactility and mechanical properties similar to that of human tissue [72].

The heart box phantom was essentially comprised of three target 'valve annuli' constructed using sheets of PVA-C and immersed in water. The prosthetic valve/patch insertion tool was tracked using the Aurora® magnetic tracking system. The AtamaiViewer® platform [40], running on a 3.2 GHz dual processor workstation, was used to integrate all hardware and software components.

To mimic the conditions for a VR-assisted IGI, the phantom was covered with surgical drapes, forcing subjects to rely on virtual information, rather than direct vision, for guidance. The entry point of the heart box phantom was 7.5 cm long and less than 5 cm wide to simulate a minimal incision.

Because this experiment was designed specifically as a low-level docking task, I was not concerned with the subject's interpretation of the target 'valve annuli' or the valve's true anatomical surroundings. Rather, the focus was on whether subjects could effectively perceive depth and converge visual and motor information using only VE design components.

The influence of each VE component on visual depth perception and visuomotor coordination during navigation and positioning was compared by measuring user targeting accuracy, and task completion time. Subjects completed a post-study questionnaire upon finishing the experiment.

Although important for an actual procedure, for the purpose of this study, I was not concerned with the use of ultrasound, since this visualization modality would confound the comparision of strictly VE components. Also, because the physical targets were fixed and defined virtually by using a shape based registration technique (described in [40]), dynamic real-time target information was not critical for accuracy. The use of a fixed reference body could ensure accurate and consistent tracking in a stable, unmoving test environment; thereby, justifying the VR-only approach.

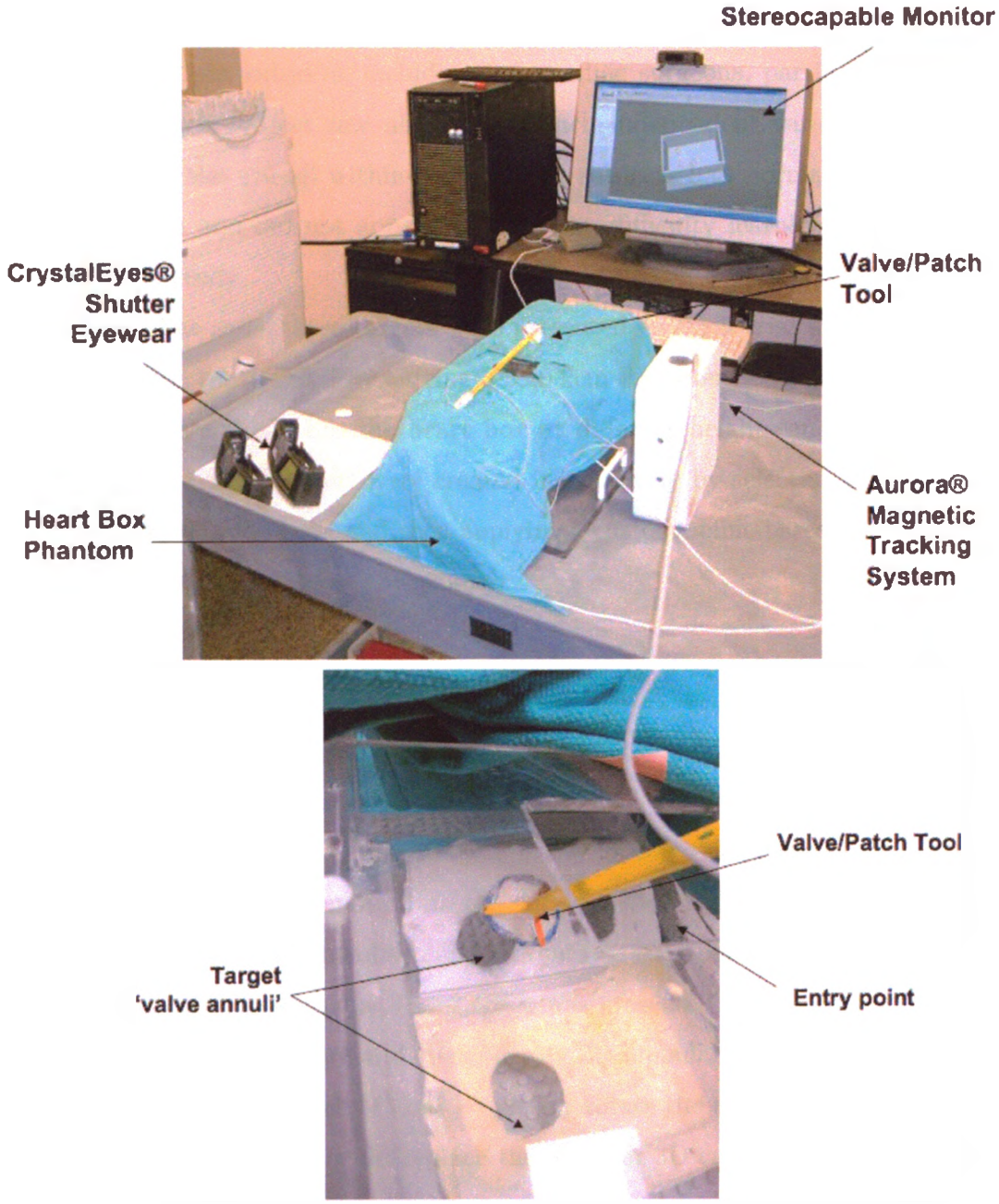

Figure 3.1: Experimental set up with valve/patch insertion tool and PVA-C heart box phantom. *Top*: Heart box phantom covered. *Bottom*: Heart box phantom uncovered.

3.3 Study Design

Subjects. Eight subjects, including two cardiac surgeons, participated in the experiment. Each subject saw all levels of each independent variable, thus the dependent variables are all within-subject. All subjects had normal or corrected-to-normal vision and did not appear to have any difficulty learning the task and completing the study.

Targets. The same three targets were presented for each trial to each subject. The targets were constructed by cutting 3.3±0.2cm diameter holes into the PVA-C sheets that were affixed onto the heart box at different heights and angles. The virtual models of these targets were created interactively by selecting points along the circumference of the hole, and then applying a custom spline-based segmentation function, used in [36], to define the geometry.

3.3.1 Visualization Factors

A summary of the three visualization factors is presented in Table 3.1. A description for each factor is detailed.

Control Type. In the existing VE setup, the surgeon's viewpoint of the VE was delegated to a technician remote from the surgical site. To explore the concept of eliminating intermediary-controlled camera rotation, a virtual 'mirroring' technique was employed. In this technique, the surgeon was presented with two simultaneous orthogonal views of the target and tool, offering 2D biplanar depth perception that could compensate for the lack of ability to freely rotate the model. I predicted that two fixed orthogonal views could replace the need for 3D camera manipulation by the technician while maintaining accuracy and improving targeting efficiency. For the purpose of this experiment, I have referred to the existing viewing condition as the 'Delegated Control' condition as the action of panning and rotating the camera view is delegated to a technician. The orthogonal viewing condition is known as the

'Orthogonal Views' condition.

Disparity. I compared the use of stereo versus mono visualization when viewing the VE. The VE was displayed on a stereoscopic monitor designed for viewing with active eyewear. To enable stereo viewing, the display mode was set to 'Stereo' and subjects wore a set of liquid crystal shutter eyewear, CrystalEyes3® (StereoGraphics®, San Rafael, CA). Because depth perception would be critical for the task, stereovision was expected to improve the accuracy of targeting compared to the 'Delegated Control' condition under monovision. It was suspected that the orthogonal view angles provided by the 'Orthogonal Views' condition would provide sufficient depth perception and thus, would not be affected by either stereo or monovision.

Spatial Context. To investigate the effect of spatial context on subject performance, two display conditions were set up: (1) A VE where the target and its surroundings were presented; and (2) an environment where only the target was shown with no surrounding context displayed. It was predicted that the display of context would improve depth perception for initial tool-to-target navigation but would not make a significant contribution to final tool positioning.

Table 3.1: The visualization factors.

Variable	No. Levels	Levels
Control Type	2	Delegated (1 Perspective View), Orthogonal Views (2 Fixed Views)
Disparity	2	Mono, Stereo
Context	2	Yes, No

These three variables formed the following visualization conditions displayed in both mono and stereovision, shown (mono only) in Figures 3.2 and 3.3. A total of 8 visualization conditions were created, 2 (control type) x 2 (disparity) x 2 (context).

The trials were grouped by these 8 distinct visualization conditions, resulting in 24 measurements per group. Using a factorial nesting of independent variables, a total of 192 accuracy and time measurements were collected (8 (subject) x 3 (target) x 2

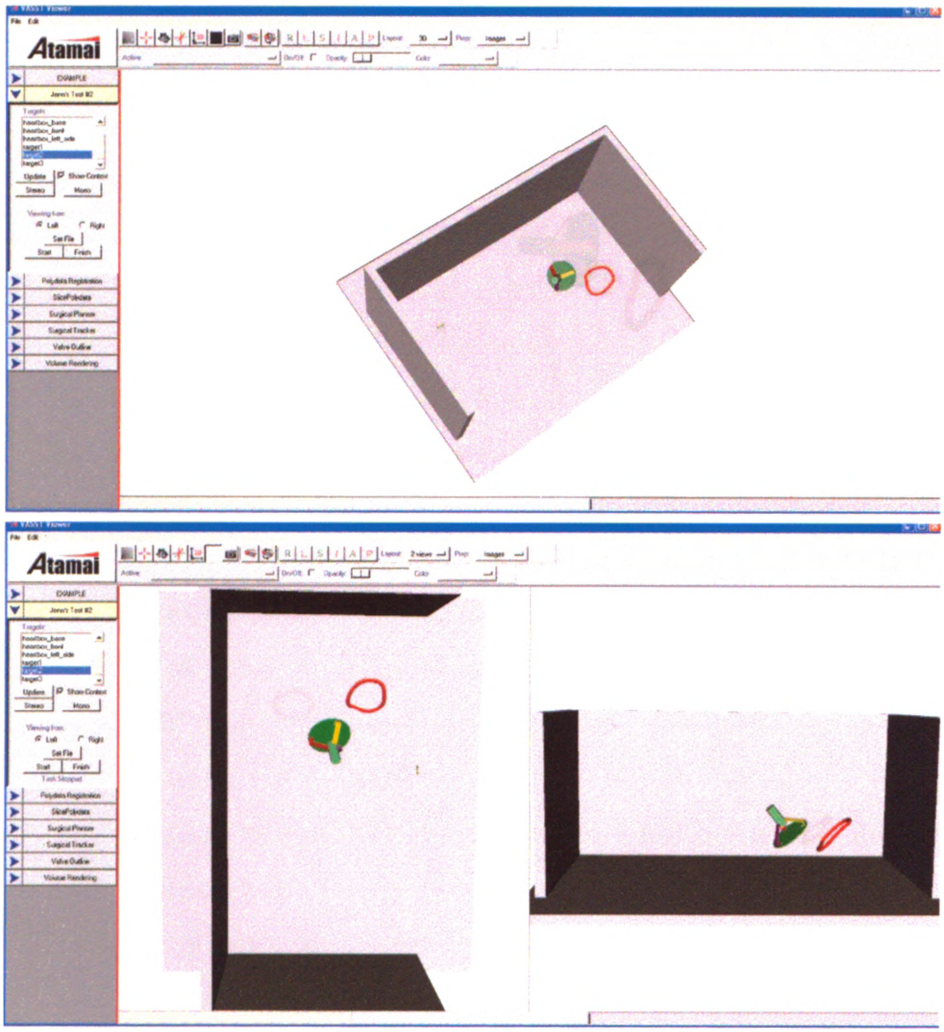

Figure 3.2: The visualization conditions, each displayed using both mono and stereo disparity, are (from *top* to *bottom*): (1) Context with 'Delegated Control' and (2) Context with 'Orthogonal Views'.

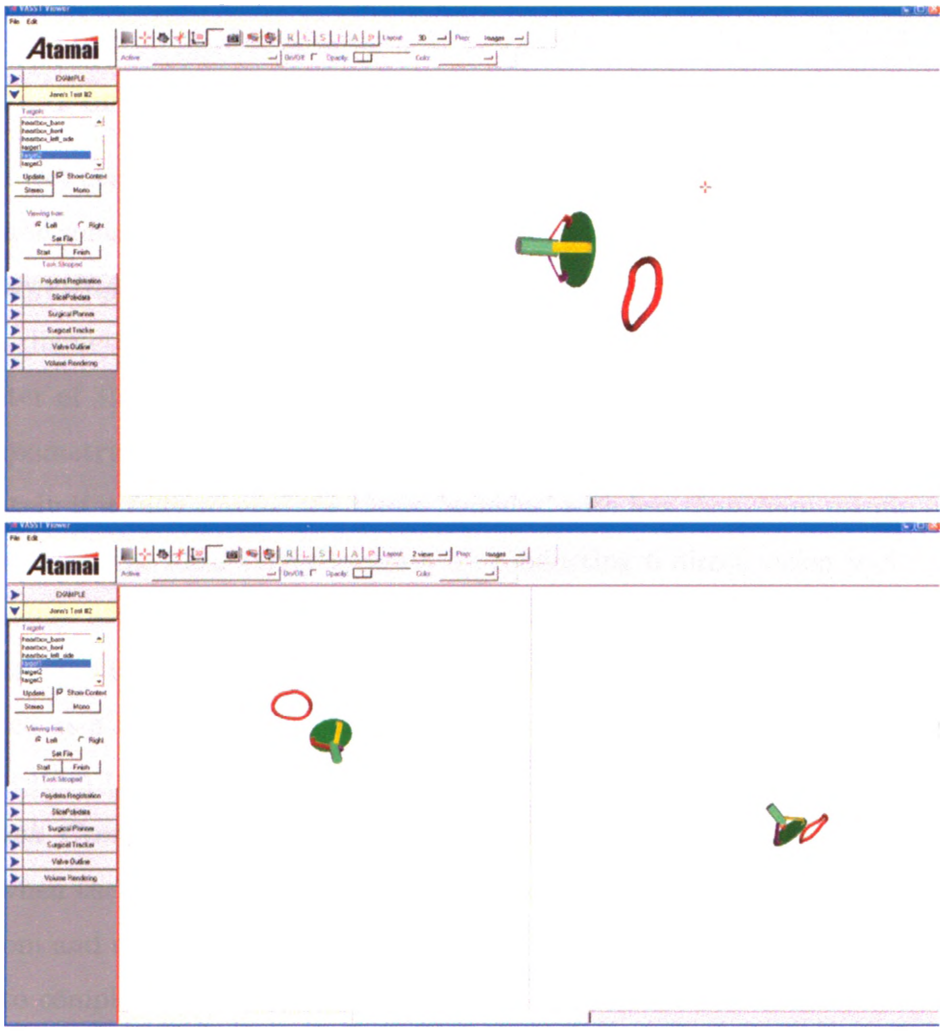

Figure 3.3: The visualization conditions, each displayed using both mono and stereo disparity, are (from *top* to *bottom*): (1) No Context with 'Delegated Control' and (2) No Context with 'Orthogonal Views'.

(control type) x 2 (disparity) x 2 (context)). Potential order effects were controlled by alternating the presentation of control type among subjects and generating random permutations for the order of targets, spatial context and disparity.

3.3.2 Performance Measures

Subject performance was measured in terms of targeting accuracy, task completion time, and quantitative and qualitative feedback from a post-study questionnaire.

Targeting Accuracy. Because the patch tool tip was approximately circular with a diameter of 3.8cm and each of the targets were similarly shaped with mean diameter of 3.3±0.2cm, the 'true' target location was difficult to identify based on pure geometry. From a clinical perspective, the patch was considered to be placed accurately if it fully covered the target 'annulus' with less than 2mm margin of error. I determined the ideal target location by conducting 6 direct vision tool-to-target placements after each subject trial. The mean (x,y,z) coordinate of the end-tool location over the 48 direct vision trials was set as the gold standard. Accuracy was reported as a RMS positional error between this gold standard and the actual end tool location tracked by the Aurora®.

Task Completion Time. Completion time in seconds was measured starting from when the subjects held the tool above the opening of the insertion point to the phantom and ending at the point when the subject felt he or she had positioned the valve to completely cover the target 'annulus'.

Post-Study Questionnaire. Participants were asked to rate the amount of confidence they felt when performing the task in the presence of each of the eight VE components using a 5-point Likert scale (1 = low confidence, 5 = high confidence), to provide a quantitative comparison of each subject's confidence level. Subjects were also asked to comment on the advantages and disadvantages of each VE component to provide qualitative assessment. For further detail, the questionnaire is provided in the Appendix B.

3.3.3 Hypotheses

Prior to conducting the study, the following hypotheses were made:

1. Stereo will result in improved accuracy as it provides depth information that is unattainable with colour and shading cues offered by monovision

2. Stereo is not necessary if two orthogonal views are shown. There will be no difference between 'Orthogonal Views' with Stereo and 'Orthogonal Views' with Mono.

3. Monoscopic vision will provide sufficient depth information in the 'Orthogonal Views' condition where two orthogonal views are displayed. There will be no difference between 'Orthogonal Views' with Mono and 'Delegated Control' with Stereo.

4. 'Orthogonal Views' will result in more efficient targeting as the need to communicate view adjustments to a technician is eliminated.

5. Context is required for tool navigation (initial targeting accuracy).

6. Context may not be needed to facilitate tool positioning (final targeting accuracy).

3.4 Results

Statistical and graphical analysis of the data obtained for all performance measures were conducted using MATLAB® Statistics Toolbox and Microsoft® Excel.

The collected RMS errors and task completion times were tested for normality before proceeding. The histograms in Figure 3.4 indicate that the collected data, from all subjects, did not follow a normal distribution.

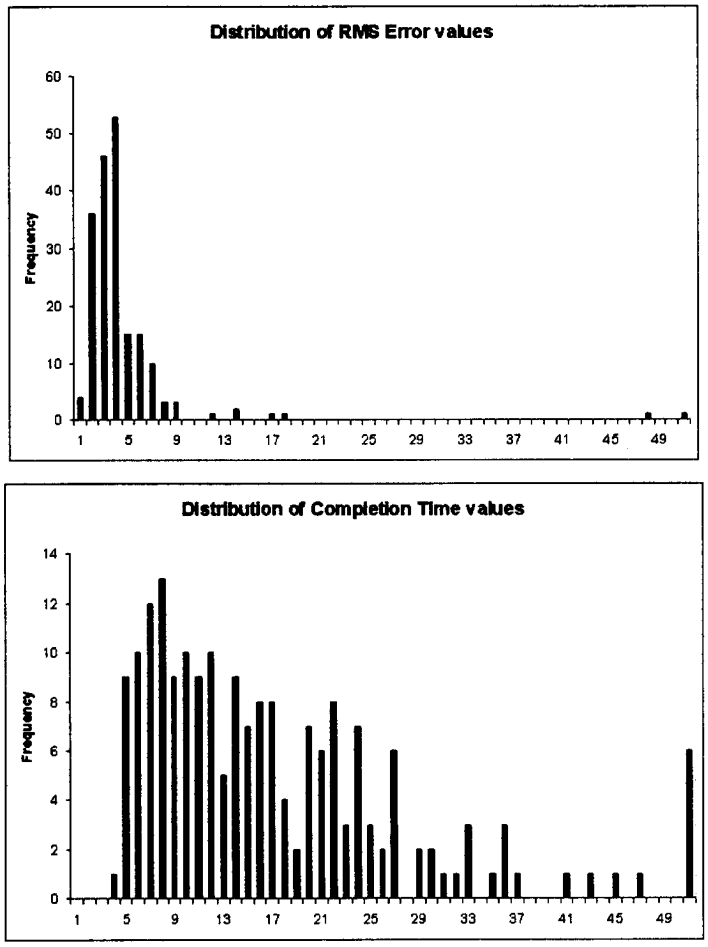

Figure 3.4: The distribution of RMS error and task completion time data values from all subjects (shown on *top* and *bottom* respectively).

As a result, the Friedman's non-parametric repeated measures test was applied as an alternative to the one-way Analysis of Variance (ANOVA) [73]. No significant effects were found for target 1, 2 or 3 in terms of targeting accuracy ($\chi^2(2)$=1.23, p=0.540) and task completion time ($\chi^2(2)$=0.800, p=0.830). Thus, all targets were considered equal difficulty.

Main effects for each factor, Control Type, Disparity and Context, were tested using a χ^2 test statistic with $\alpha = 0.05$. The results of the statistical tests are summarized in the Table 3.2.

Table 3.2: Main effects - Results of Friedman's analysis on task performance measures. The * indicates a significant effect with $\alpha = 0.05$.

Factor	Dependent Variable	df	χ^2	p
Control Type	RMS_{error}	1	1.40	0.236
	Time	1	7.15	< 0.01*
Disparity	RMS_{error}	1	0.480	0.488
	Time	1	0.500	0.481
Context	RMS_{error}	1	0.440	0.507
	Time	1	9.00×10^{-2}	0.767

Friedman's analysis does not offer a method of testing for interactions between independent variables; however, several non-parametric tests to detect interaction effects have been devised. The Adjusted Rank Transform was selected for its ease of implementation and performance in detecting interactions in a set of simulated data [74]. The Adjusted Rank Transform method operates on the ranks of the data, by first eliminating any main effects of the independent variables from the data, ranking the resulting data and then analyzing the ranks with an ANOVA. Main effects are removed by subtracting an unbiased estimator (i.e. the mean) from the original data. The results of the Adjusted Rank Transform test on interactions between variables is shown in Table 3.3.

Table 3.3: Interaction effects - Results of the Adjusted Rank Transform test. The * indicates a significant effect with $\alpha = 0.05$.

Factor	Dependent Variable	*df*	*F*	*p*
Control Type by Context	RMS_{error}	1	7.00×10^{-2}	0.789
	Time	1	0.210	0.646
Control Type by Disparity	RMS_{error}	1	4.04	< 0.05*
	Time	1	1.00×10^{-2}	0.913
Disparity by Context	RMS_{error}	1	1.11	0.293
	Time	1	6.00×10^{-2}	0.803

3.4.1 Effects on Accuracy

No main effect on accuracy was observed for any of the visualization factors, such as Control Type ($\chi^2(1)$=1.40, p=0.236), Disparity ($\chi^2(1)$=0.480, p=0.488), or Context ($\chi^2(1)$=0.440, p=0.507). However, an interaction effect between Control Type by Disparity on accuracy was observed (F(1,188)=4.04, p<0.05).

The combination of each visualization factor on accuracy is shown in Figure 3.5. The interaction effect of Disparity by Control Type on accuracy is shown in Figure 3.6.

A post-hoc pairwise comparison was applied to determine the significant effect between the levels of interacting variables, Control Type and Disparity. First, the data were grouped according to all level combinations of the Control Type by Disparity factors as shown in Figure 3.7. Then, the Wilcoxon Signed-Rank test[75] was used to compare each combination. Subject accuracy was significantly worse when using the 'Delegated Control & Mono' condition (12.1±6.8mm) compared to all other conditions (p<0.05). Accuracy improved by 67% when using 'Orthogonal Views & Mono' (4.02±1.5mm), 66% when using 'Orthogonal Views & Stereo' (4.05±1.4mm), and 70% when using 'Delegated Control & Stereo' condition (3.63±1.0mm).

[1]RMS error is given with a 95% confidence interval for $\alpha = 0.05$ using t-statistic for n=8. The t-distribution is used because of the small sample size. The resulting measure ($RMS_{error} \pm t$*Standard Error) represents the range of values that is likely to cover the true mean [76].

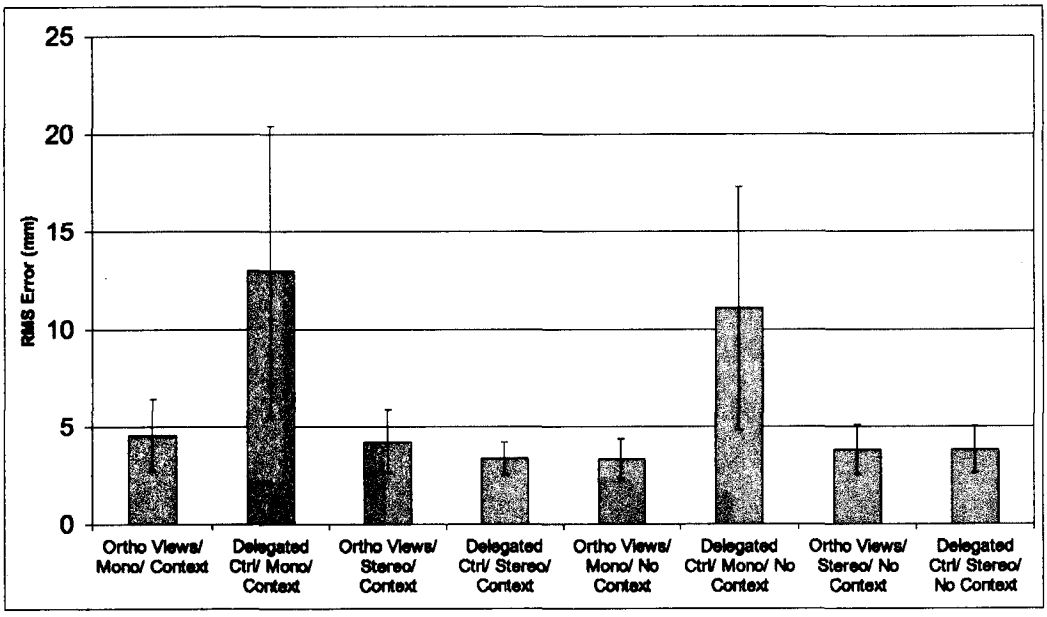

Figure 3.5: Effect of each visualization condition on accuracy. Y-error bars represent the 95% confidence interval of the RMS error [1].

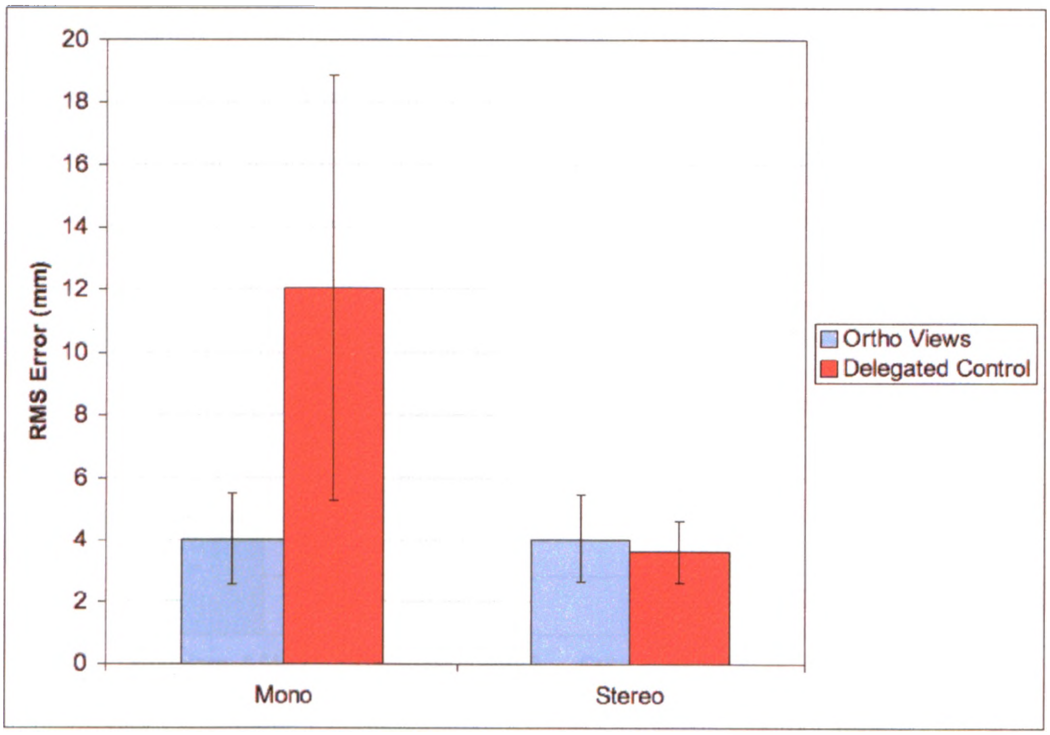

Figure 3.6: Effect of Disparity by Control Type on accuracy displayed as an interaction plot. Y-error bars represent the 95% confidence interval of the RMS error.

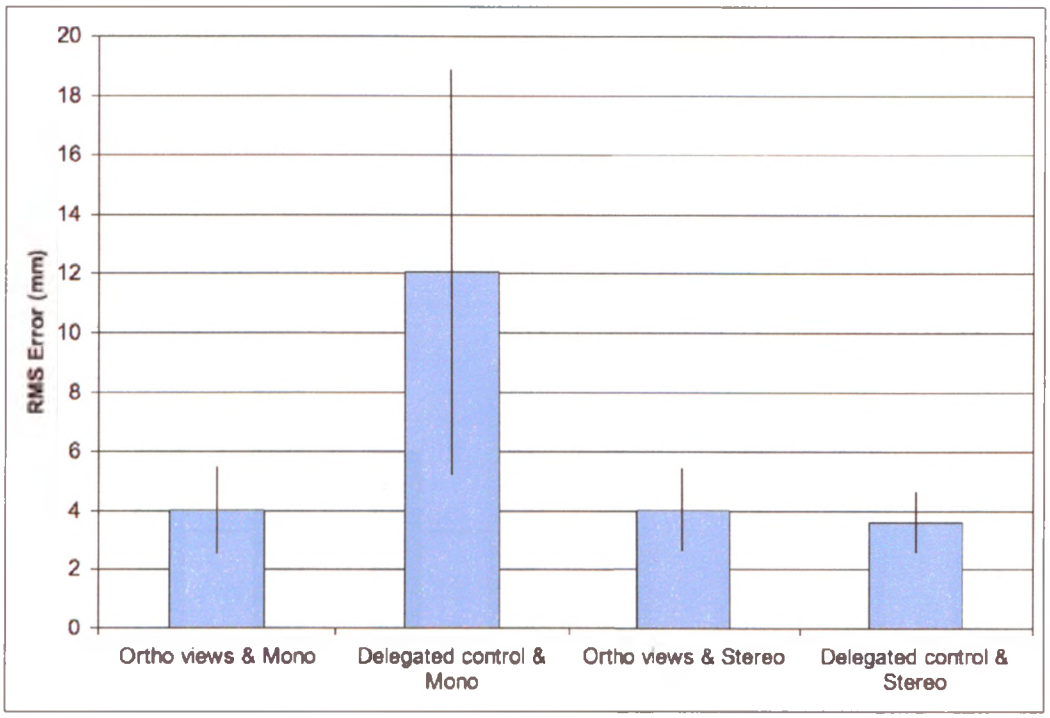

Figure 3.7: Mean accuracy for conditions involving factors Disparity and Control Type. Y-error bars represent the 95% confidence interval of the RMS error.

3.4.2 Effects on Completion Time

The effect of each visualization condition on task completion time is displayed in Figure 3.8. A main effect was observed for Control Type on completion time ($\chi^2(1)$=7.15, p<0.01) and is shown in Figure 3.9. Timing improved by 25% under 'Orthogonal Views' (14.7±6s) compared to 'Delegated Control' (19.5±11s). Main effects on completion time were neither observed for Disparity ($\chi^2(1)$=0.500, p=0.481) nor Context ($\chi^2(1)$= 9.00×10^{-2}, p=0.767). There were no notable interaction effects between any of the independent variables.

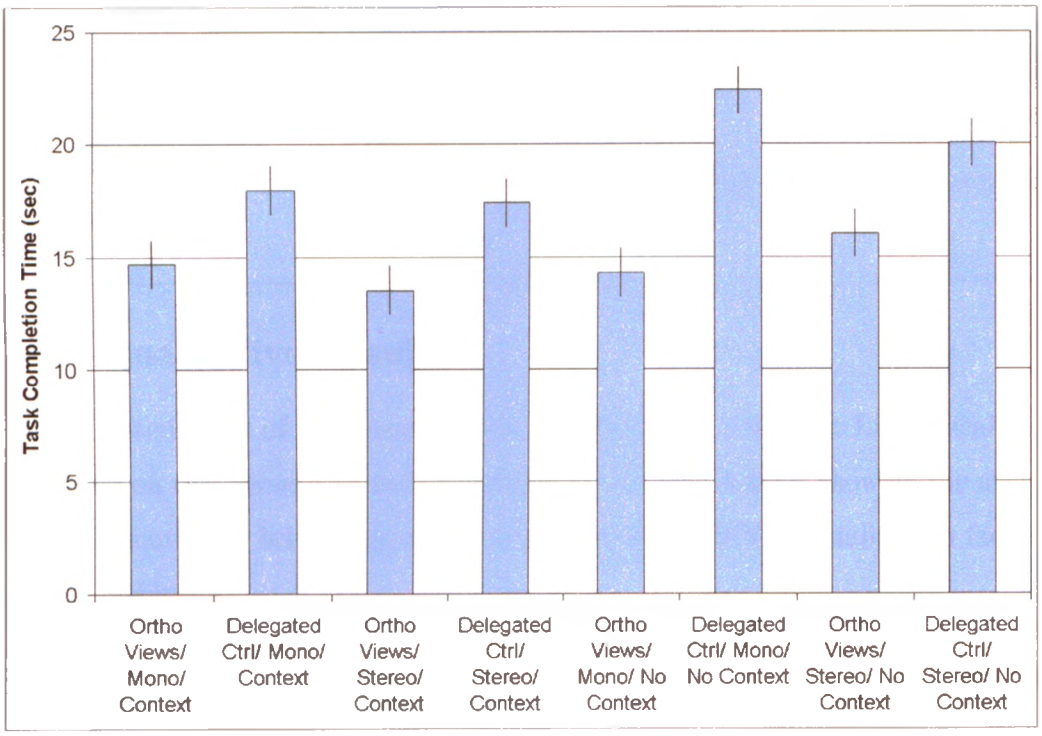

Figure 3.8: Effect of each visualization condition on task completion time. Y-error bars represent the standard error.

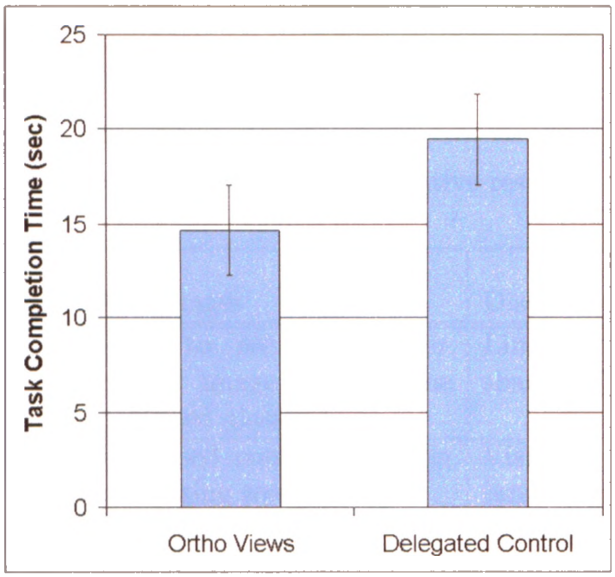

Figure 3.9: Effect of Control Type on task completion time. Y-error bars represent the standard error.

3.4.3 Qualitative Results

Users rated each of the visualization factors using a 5-point Likert scale and commented on the advantages and disadvantages of each. A table showing the average confidence score and brief summary of comments for each VE visualization factor is shown in Table 3.4.

3.5 Discussion

Based on the results, the first four hypotheses were supported. The use of 'Stereo' under the 'Delegated Control' condition resulted in a marked improvement in accuracy compared to the case where monovision was used. Thus, it can be concluded that using monovision with the existing setup is suboptimal for depth perception tasks such as positioning and navigation. On the other hand, no notable improvement was

Table 3.4: Summary of qualitative results.

Factor	Average Rating	Advantages	Disadvantages
Monovision	3.5	Able to see a brighter, crisper image without the polarized glasses	Limited to only colour and shading depth cues
Stereovision	4.0	Increased confidence when navigating toward target	Visibility hindered by polarized glasses; Stereo did not make a huge difference during the positioning task given the small field of view
Orthogonal Views	4.5	Able to learn task quickly; Images were placed close enough together to facilitate simultaneous interpretation	Unable to verify the final tool location at different angles by free rotation
Delegated Control	4.0	Ability to see all angles whenever necessary	Inefficient for viewing target from all angles; Distance to target from perpendicular angle was usually incorrect and thus required rotation of the camera to the orthogonal angle
Context	4.5	Easy to establish tool orientation, especially during initial motion towards target; More confident when approaching target	*None*
No Context	3.5	Easy to position target once tool is close enough	Initial tool orientation is often misinterpreted; Much time spent on orienting tool at the beginning of the task

found between the use of 'Stereo' and 'Mono' under the 'Orthogonal Views' condition, suggesting that using either 'Mono' and 'Stereo' can be sufficient for depth information given the display of two orthogonal views. In addition, because there was no significant difference between 'Orthogonal Views with Mono/Stereo' and 'Delegated Control with Stereo', it can be suggested that the depth information supplied by orthogonal viewing is comparable to stereo viewing. Further, because the task completion time was significantly improved under the 'Orthogonal Views' compared to 'Delegated Control' for all visualization conditions, it is evident that targeting can be performed more efficiently without compromising accuracy. Thus, by providing a method of perceiving depth without having to communicate view adjustments to the technician, the surgical navigation and positioning tasks can be performed more effectively.

Displaying the virtual representation of the target surroundings (i.e. Context) had significant impact on neither completion time nor targeting accuracy. However, subjective assessment from a post-study questionnaire revealed that displaying peripheral context was necessary for initial trajectory planning, but this information was ignored once the subject had approached the target within a close proximity. All subjects reported that by seeing the target's surroundings, they were able to quickly establish tool orientation and felt more confident in approaching the target. Although there was no quantitative evidence to support the idea that context could improve navigation, the qualitative feedback suggested otherwise. This supported the hypothesis that context is not essential for tool positioning but could be helpful in navigation.

Major outliers were observed in the accuracy measurements of the following two conditions: (1) 'Delegated Control' with monovision and context; and (2) 'Delegated Control' with monovision and no context. Under these two conditions, since the common denominator was 'Delegated Control' and monovision, it was possible for the subject to believe they were accurately positioned over the target when in fact they did not have adequate depth perception. A comparision of the X,Y,Z location of these two outliers illustrate that the largest error is in the X-direction (See Figure

3.10). Since free rotation is an aspect of the 'Delegated Control' viewing condition, the distance corresponding to depth could have been along any axis depending on the camera's viewpoint. For the particular outliers, the misinterpretation of target depth appears to be coming from the X-direction.

The large RMS error values observed for these two conditions (RMS error=13.0±7.4mm with Context; RMS error=11.1±6.7mm with No Context) indicate greater chance of misinterpreting depth compared to conditions where 'Orthogonal Views' was the common factor (RMS error=4.01±2.2mm). This may suggest that fixed orthogonal views reduces the presence of major errors.

3.6 Key Findings

Through user-based analysis, it is determined that certain VE components can influence or diminish user performance in a simple tool-to-target docking task guided solely by VR. The following conclusions can be generalized about VE components: Control Type, Disparity and Context, and their interactions with each other.

- **For a single view perspective, stereovison is critical for targeting accuracy.**

Stereovision could significantly improve targeting accuracy when the existing 'Delegated Control' mode was used. Colour and shading, the main depth cues offered by monovision, was found insufficient for depth information. However, by introducing simultaneous orthogonal views, monovision was as effective as stereovision in providing 3D information with one view.

- **Fixed orthogonal view angles can improve accuracy, efficiency and reduce the probability of major accuracy outliers.**

Not only was the 'Orthogonal Views' approach an effective method of displaying 3D information on a 2D computer monitor, the user was able to dock the target without

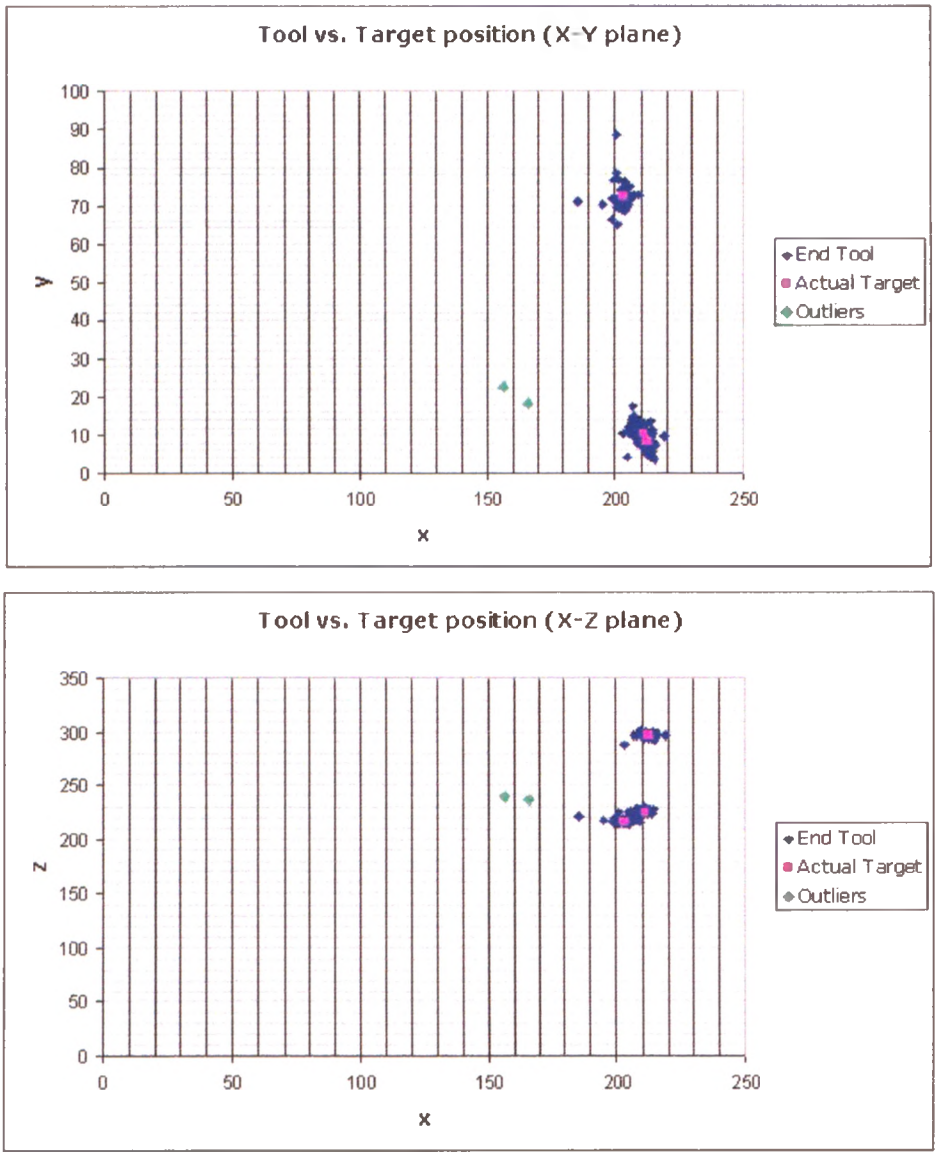

Figure 3.10: *Top*: Position error on X-Y plane. *Bottom*: Position error X-Z plane.

manual view rotation. By removing the need to communicate view adjustments to a camera technician, one could perform the docking task more efficiently. Evidenced by the absence of extreme outliers under 'Orthogonal Views', it is suggested that the chance of large errors are reduced with fixed viewing poses.

- **Displaying background context improves user confidence in navigation and tool orientation.**

It was demonstrated qualitatively that context is useful for the users' initial orientation but is less relied upon once trajectory toward target had begun. Future work should be directed to determine at what point during a surgical procedure is VR context information necessary and when it is frivolous.

Chapter 4

The 'Surgeon-Controlled' Interface

4.1 The 'Delegated Control' Problem

The VR-enhanced US-guidance system developed at the VASST lab, uses 3D cardiac surface models generated from MRI or CT data to provide anatomical context to 2D intra-operative US images. With this surgical platform, surgeons are able to see the position of their tools relative to the surgical targets in real-time. In addition, the virtual camera can be manipulated with 6 DOF, allowing surgeons to view the target and tools from any angle, at any distance.

As it currently stands, the typical OR offers no accessible computer input modality for the surgeon, whose hands are usually occupied. The full functionality and flexibility of the VR-enhanced system can only be exploited with the assistance of an additional technician who has the job of rotating, panning, and tilting the VR-US hybrid scene based on the surgeon's directions. Surgeon-to-computer interaction with the existing setup, shown in Figure 4.1, was observed in a first pass *in vivo* acute porcine valve surgery conducted at the CSTAR testing centre.

While the surgeon could successfully perform the surgical tasks, guidance using a VR-enhanced US-based system was cumbersome and suboptimal. Like most computer-assisted surgery systems, at least one technician must be present to serve

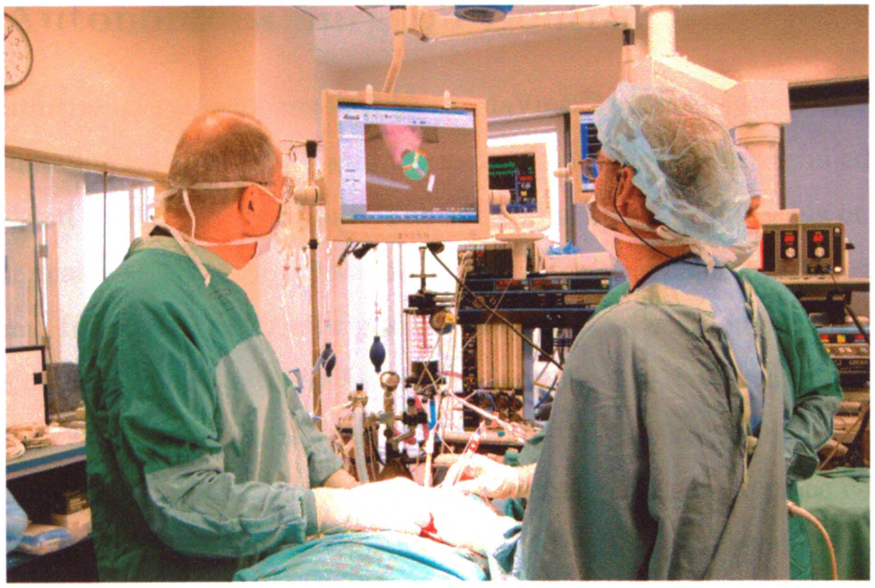

Figure 4.1: Current implementation of VR-enhanced cardiac surgery platform as seen in an acute porcine study at CSTAR in April 2008.

---

as an intermediary between the surgeon and machine [77] to whom surgeons must communicate camera view adjustment commands. Delegated control errors caused by a miscommunication between surgeon and technician could lengthen procedure time and cause frustration [78]. Inadequate navigation of the camera by the technician can impair the surgeon's visualization, cause frustration and compromise patient safety [43]. Moreover, the surgeon would likely experience greater cognitive load having to instruct the technician on which task to perform or button to click, while focusing on the surgery itself. Evidently, there needs to be a way for surgeons to control these technological tools as VR becomes more pervasive in modern ORs.

4.2 Autonomous Surgeon Control

The intuitive user control provided by the daVinci® surgical system (IntuitiveSurgical, Mountain View, CA) was one of the main reasons for the successful adoption of robot-assisted cardiac interventions. The surgeon is able to control his or her view of the surgical area by using a foot pedal to 'clutch' the camera, leaving surgical instruments frozen at a particular position, and then re-positioning the viewpoint. This level of manipulation is useful for viewing large medical data sets through a small field of view. In another example, the AESOP2000® (Computer Motion Inc., Goleta), a video-directed endoscope that could be positioned with simple voice commands, eliminates the need for a human assistant [79]. Falk *et al.* [79] demonstrated that a voice-controlled robotic arm would perform comparably to pedal controlled devices because verbal control is part of the normal concentration pattern of the operating surgeon.

Surgeons may also require automonous control of system processes for safety reasons. For instance, the 'start/stop' voice command in [78], and the semi-automatic system shutdown with verbal cue 'microphone' in [57], put users in charge of the functionality of their own surgical equipment. By offering a simple method to safely operate surgical machinery, the usability and acceptance of these technologies can be increased.

In an attempt to make computer equipment more surgeon-accessible, M/ORIS, a Medical/Operating Room Interaction System was designed with the intention of giving surgeons direct control over their computer systems [78]. The proposed multimodal framework integrated implicit user control via gestures, explicit control via voice recognition, and task-specific control via devices such as foot pedals, joysticks and sterilized keyboards [78].

For more complex 3D multi-modal data visualization, Hinckley *et al.* [80] and Cooperstock *et al.* [81], described two methods for effective user interaction: two-

handed virtual manipulation and hand gesture control, respectively. Under conditions of low latency and accurate input, these modalities were more efficient than keyboard and mouse control for the task of manipulating neurosurgical planning information [80].

4.3 Surgeon Controlled Interfaces

I explored the potential for autonomous view control by developing the following three interface prototypes:

1. Preset view angle control with monoscopic biplanar vision
2. Tracked head-mount display (HMD) control with stereoscopic vision
3. Tangible 'camera' control with monoscopic biplanar vision

4.3.1 Preset View Angles

This interface was based on the contention that the surgeon's desired viewing perspective is specific to the surgical approach. For mitral and aortic valve replacement procedures, surgeons generally opted to see a direct, head-on view of the target valve annulus. Other commonly used view angles specified by collaborating surgeons were: (1) the standard anterior-posterior anatomy view; (2) the orthogonal-to-target view; and (3) a close-up view of the surgical probe tip.

This interface was able to support as many pre-determined view angles as necessary for the operation. By associating a keyword to each view, the surgeon was able to control the views by a simple verbal command.

The chosen view was presented in the largest left-most display pane while its corresponding orthogonal view was displayed in a smaller pane (4:1 area ratio) located at the top-right corner (see Figure 4.2). Since human gaze naturally moves from left

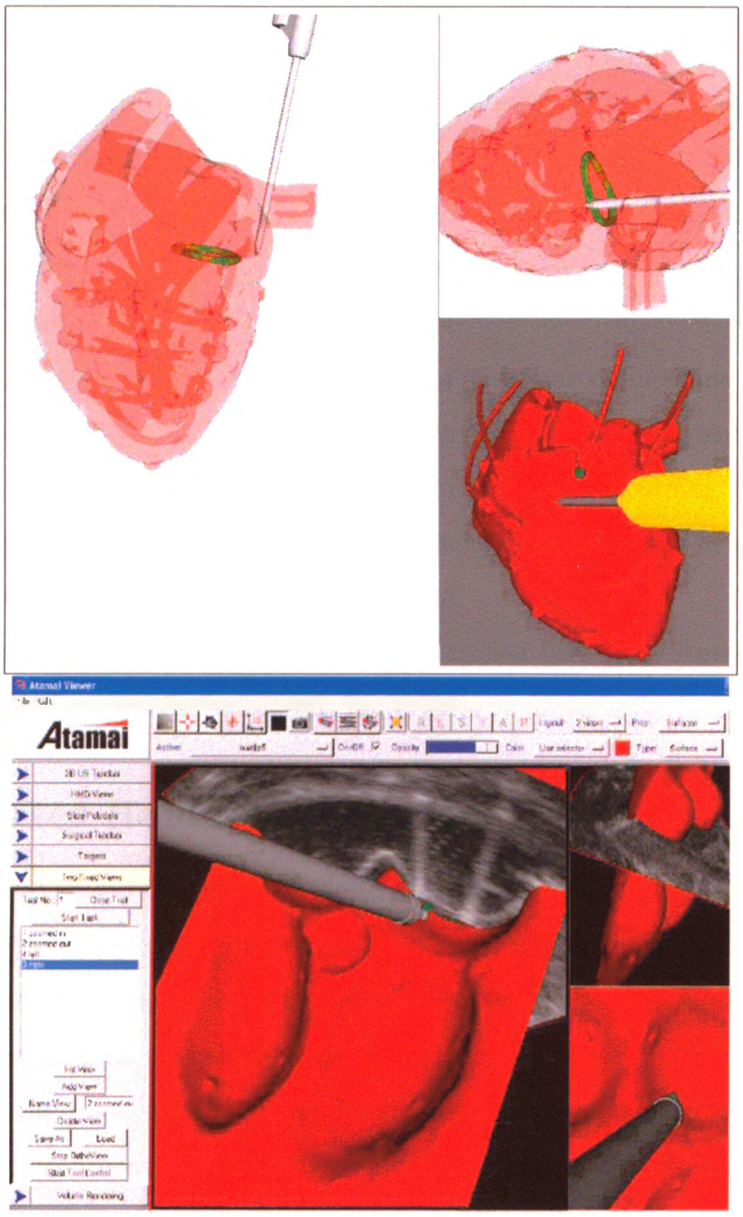

Figure 4.2: *Top*: Conceptual design of the 'Preset View Angles' interface displaying a 3D surface model of the Chamberlain heart phantom; *Bottom*: Implementation of the 'Preset Views' in the experimental setup displaying a 3D surface model of the PVA-C phantom specific to this study.

to right in North-American culture [82], this layout should enable depth perception in a comfortable manner.

The bottom-right area displayed a close-up probe-to-target view (4:1 area ratio), as requested by surgeon feedback. The screen resolution was 1280 x 1024 pixels. It was anticipated that the reduction in camera motion during surgical navigation could improve the surgeon's orientation in the virtual world, resulting in better accuracy and faster completion time.

For prototyping purposes, four 'Preset View' angles were defined and selected using keywords such as 'Zoom in', 'Zoom out', 'Standard AP', 'Left View', and 'Right View'. In reality, any number of views could be defined and named according to the needs of the surgical team. User control could be implemented using speech or gestural input, or using a hands-free device such as a foot pedal.

Surgical tools and phantom location were tracked using the Aurora® magnetic tracking system from Northern Digital, Inc (NDI).

4.3.2 Tracked Head-Mount Display

Head-mount displays (HMDs) provide an immersive VR experience. Studies showed that when the immersive environment was an exact virtual replica of the observer's physical environment, depth could be perceived without the typical distortion found when VR represented an unrealistic setting [83]. The ProView^{TM} XL35 HMD (Kaiser Electro-Optics Inc.) was used to display a stereoscopic computer-generated replica of the physical environment enhanced by real-time US video. Screen resolution of the HMD was 1024 x 768 pixels, the field of view was 30° diagonal and 21° vertical x 28° horizontal, and the weight was 35 ounces [84].

The HMD was designed to accommodate the subject at his or her most comfortable viewing position. I individually calibrated the vantage point of the HMD to each subjects' desired stance, which was typically with the subject's head tilted at a 10-20° angle below the horizontal line of sight and in direct alignment with their hands. The

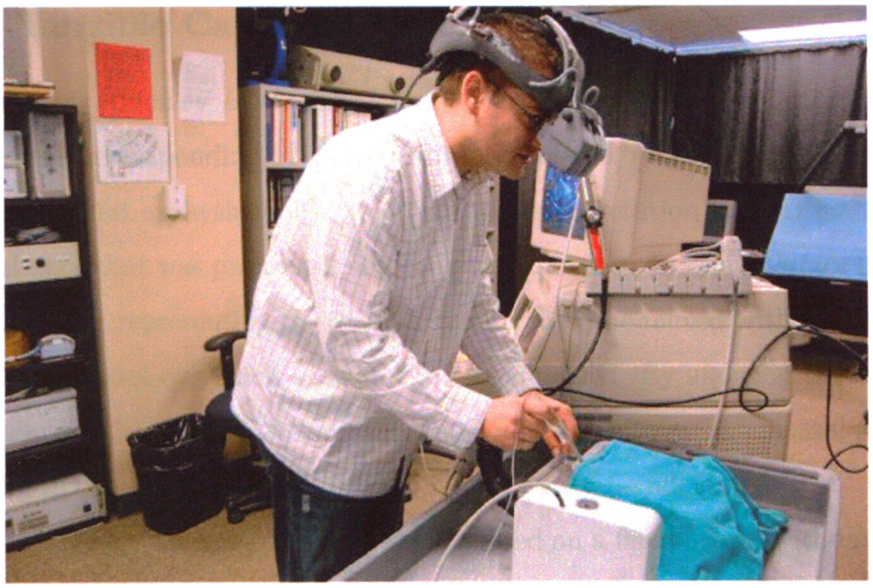

Figure 4.3: Implementation of the 'Tracked Head-mount Display' in the experimental setup.

---

implementation of the setup is shown in Figure 4.3.

Six passive markers were placed on the head-mount display creating a passive tool that was tracked optically with the Polaris® (NDI). An optical tracking system was used to track the HMD as it could provide a larger field of view. All other surgical tools and phantom location were magnetically tracked using the Aurora® system. By tracking the HMD in real-time, the orientation and position of the subject's head were mapped to the camera's direction of projection. Spatial orientations of the physical and virtual worlds were represented in the same coordinate space, thereby reducing the number of reference frames that must be mentally merged and interpreted by the surgeon. It was predicted that the immersive and interactive view control provided by the HMD interface would speed task performance. However, it was uncertain whether accuracy would improve if the user was relying solely on stereovision for depth perception.

4.3.3 Tangible Camera

Inspired by Hinckley *et al.* [80], I implemented a tangible user interface with a base reference corresponding to a physical replica of the 'surgical' site. Hinckey *et al.* demonstrated that a physical object can aid the interpretation and 3D manipulation of medical data that was presented on a computer monitor. In this interface prototype, a physical object representing the virtual camera could be moved with 6 DOF, giving subjects full control over their view of the virtual world. By allowing the subject to physically grasp the 'camera' to pan, tilt, and rotate the display, the mental processing involved in 3D spatial reasoning on a 2D monitor may be alleviated.

For prototyping purposes, the tool was mounted on a flexible arm with interlocking junctions that could be moved in 6 DOF. Placed at the base of the arm was a miniature representation of the actual phantom environment, which served as a frame of reference for the user. The concept and its resulting implementation are shown in Figure 4.4. Surgical tools, phantom and tangible camera were tracked using the Aurora® tracking system.

For future implementations, the base frame of reference could be replaced with a 3D physical replica of the target organ or be completely independent of a confined object. For instance, the surgeon's dominant hand, likely occupied with a surgical instrument, could serve as a base frame of reference for the other hand, which could be positioning the 'camera' to adjust perspective. Motivated by Guiard's Kinematic Chain model which states that the dominant hand usually finds its spatial reference based on the motion of the non-dominant hand [85], the envisioned interaction would be similar to the movement of a flashlight (with one hand) over the area of focus.

Similar to the 'Preset View' condition, the display was shown with a 1280 x 1024 pixel screen resolution and 32-bit color quality. Monocular biplanar views provided depth information. The larger left panel of the display showed the main view controlled by the 'Tangible Camera' while a smaller right panel (3:1 area ratio) showed a close-up of the target area at a direct tool-to-target angle. The right panel view contained

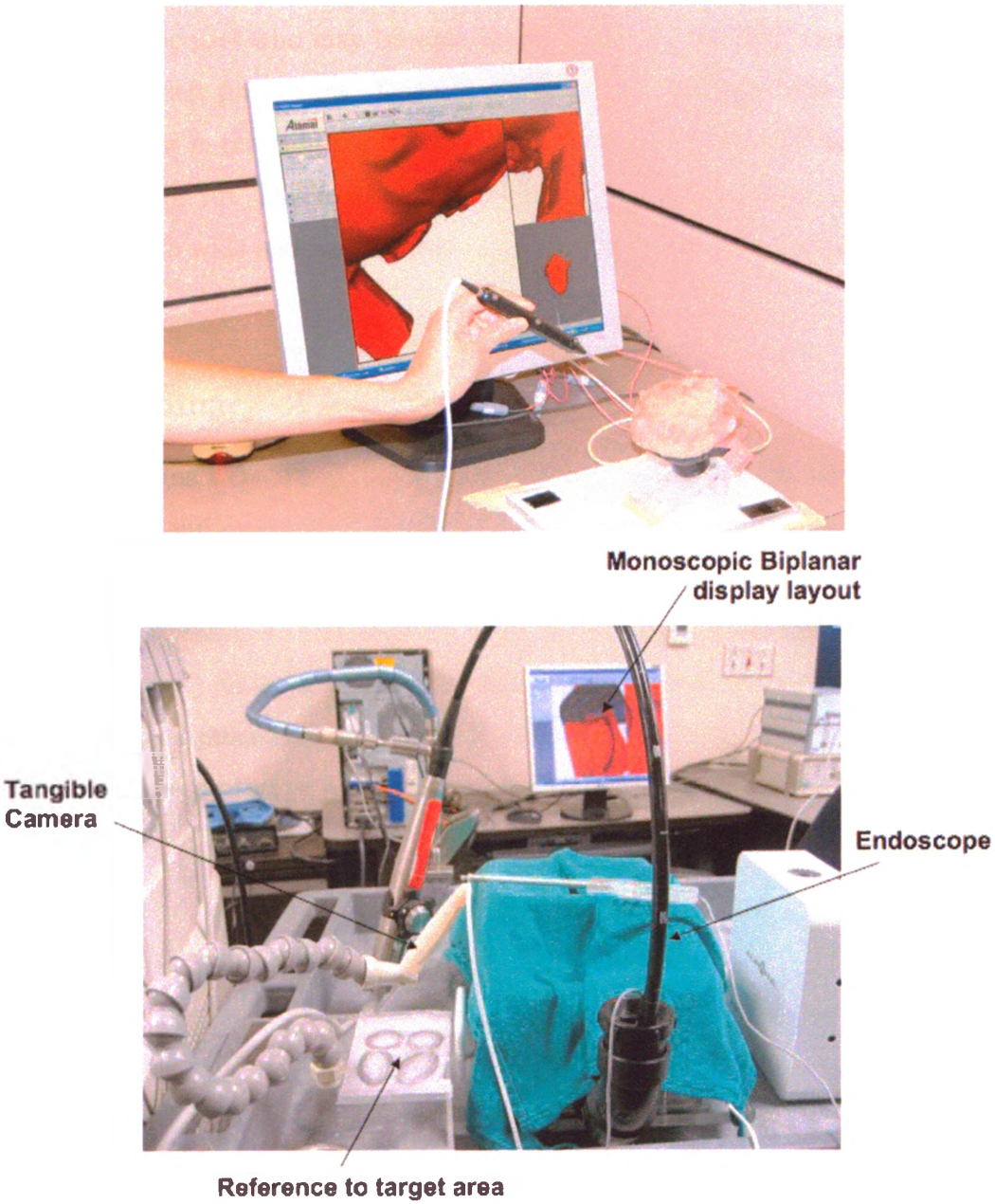

Figure 4.4: *Top*: Conceptual design of the 'Tangible Camera' interface; *Bottom*: Implementation of the 'Tangible Camera' in the experimental setup.

a fixed perspective which served as a stable reference for the user, and worked to complement the user who may become disoriented with the 360° viewing flexibility offered by the left panel view. During 'positioning', users were able to rely on the stable fixed view in the right panel while maneuvering the camera to display an angled view of the probe tip. This fixed close-up view may provide the same benefit observed when using both wide and narrow field-of-view (FOV) together to assist a positioning task [86].

4.3.4 Summary of Interface Prototypes

These designs were based on the requirements gathered from preliminary task-based analysis and expert advice described in Chapter 2 as well as the design factor considerations explored in Chapter 3.

All three prototypes eliminated the need for the intervention of a display control technician by: (1) allowing surgeons to directly manipulate the virtual environment during a surgical procedure; and (2) providing a means for depth perception. A brief summary of the prototype components and intended method of use is described in the Table 4.1.

Table 4.1: Summary of interface prototype specifications and descriptions.

Interface Prototype	Surgeon Control	Depth Perception	Description
Preset View Angles	Keyword-coupled (indirect)	Monoscopic biplanar display	Surgeon sets up virtual space prior to operation and can toggle between the preset views during the procedure using simple verbal commands
Tracked Head-mount Display	Head-coupled immersive control (direct)	Stereoscopic display	Using the movement of their head, the surgeon can control the viewpoint in an intuitive manner
Tangible Camera Control	Hand-coupled non-immersive control (direct)	Monoscopic biplanar display	Using their hand, the surgeon can 'clutch' and 're-adjust' the viewpoint by directly moving a physical representation of the 'virtual camera' in 6 DOF

Chapter 5

Empirical Evaluation of Surgeon-Controlled Interfaces

5.1 Objectives

As a continuation of the UI design methodology described in Figure 2.2 of Chapter 2, this chapter describes the user-centered evaluation (step 5) and the comparative evaluation (step 6) of the three user interface prototypes detailed in the previous chapter.

5.1.1 User-Centered Evaluation

The goal of a user-centered evaluation was to quantifiably assess and improve user interaction [87]. Empirical and observational results were used to determine the performance of each interface compared to a control. The outcomes were design validation and requirement recommendations.

Each prototype was individually assessed by comparing user performance results, such as accuracy, timing, and motion trajectory during navigation and positioning, to the existing setup. For the purpose of this experiment, the existing setup was referred

to as the 'Delegated' condition, since the user must delegate the control of their view to a human assistant. It was predicted that ***surgeon-control could improve task performance*** while not imposing any additional workload on the user compared to the 'Delegated' condition.

5.1.2 Comparative Evaluation

In contrast, comparative evaluation was the quantitative assessment of one interface prototype compared to other evolving prototypes for performing the same user tasks [87]. The outcome was an increased understanding of which prototype, or components of a prototype, support or hinder surgeon control.

By comparing each of the three interface prototypes with respect to each other, conclusions were drawn about the difference between (1) stereoscopic versus monoscopic biplanar display of depth information; and (2) the use of keyword-coupled indirect control ('Preset Views') versus head-coupled direct control ('Tracked HMD') versus hand-coupled direct control ('Tangible Camera').

As reported by current literature, there exists contradictory evidence regarding the value of using stereovision. The advantage of using either mono and stereovision is speculated to be dependent on the users' task as well as the users' experience. For instance, neurosurgical path planning [81] and target localization within a skull phantom [45], are more efficient and more accurate when conducted under stereovision. On the other hand, laparoscopic tasks involving dissection, fastening, suturing and knot-tying are not significantly improved by 3D endoscopic video displayed in stereo [88] [89]. I am interested in the prospect of using stereovision in a valve/patch implantation task.

Based on the user study discussed in Chapter 3, it was suggested that for a probe-to-target placement task, ***two monoscopic planar views positioned at orthogonal angles to each other, would be better for probe-to-target depth perception*** than stereoscopic vision with one rotatable view. Orthogonal viewing planes was

deemed sufficent for overcoming the problem of 3D object occlusion characteristic of VR environments [90]. I sought to build on my previous understanding of depth perception using biplanar views gathered from the user-based study results in Chapter 3.

It was uncertain as to whether keyword, head or hand-coupled control would be more advantageous for intra-operative surgeon control. Ideally, minimal hand control would be preferred for the OR due to the strict maintenance of sterility and the abundance of handheld instruments. It was predicted that for a clinical OR environment, the ***keyword-coupled indirect control offered by the 'Preset Views' interface would be optimal*** for enforcing surgeon autonomy as it did not require any new apparatus. However, further understanding of head and hand-coupled modalities could be useful for guiding the direction of future interface designs.

5.2 Task and Experimental Setup

5.2.1 Imaging System and Phantom

Contrary to my previous experiment which tested design parameters, full-fledged interface designs were evaluated here. Thus, US was necessary for a complete interface evaluation. The imaging workstation comprised a Philips Sonos 7500 US machine, a Philips adult TEE transducer probe (M/N:T6210) with frequency of 5MHz, and a 3.2GHz dual-CPU machine with 2GB RAM. US images were captured at 30Hz at a resolution of 640x480 pixels. The image-guidance software platform, AtamaiViewer®, integrated all components such as the VTK surface models, the real-time B-mode 2D US feed and tracking information from the Aurora® and Polaris® tracking systems.

A custom phantom was designed for this experiment using PVA-C and 2.0mm glass bead targets. A tissue-mimicking material, PVA-C was chosen as it is suitable for US, MRI and CT imaging [72]. The bead targets were embedded within the phantom

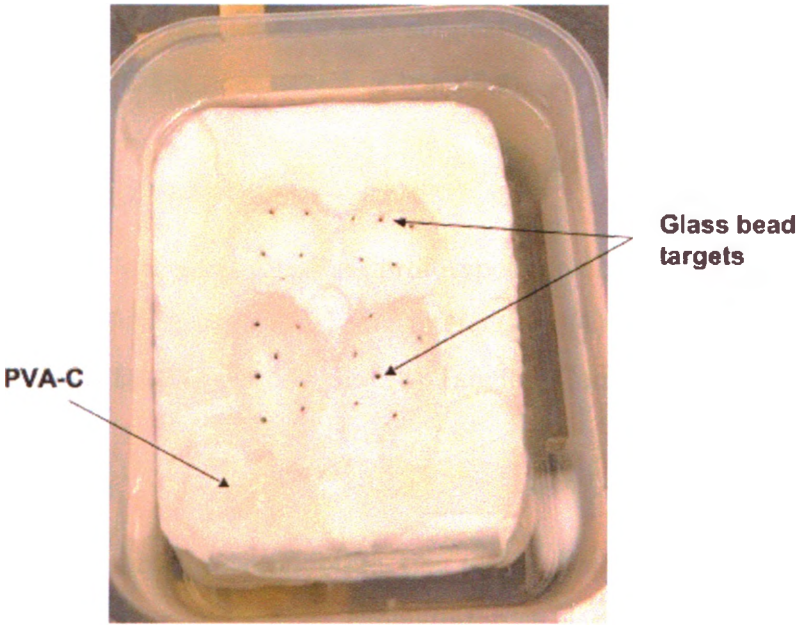

Figure 5.1: PVA-C phantom with embedded glass bead targets.

---

whose shape was designed as an abstract representation of a cross-section of the four chambers of the heart (see Figure 5.1). The location of the beads were generalized as intricate target points that could exist within the heart.

This non-realistic, abstract cardiac model was ideal for specifying clinically relevant views of the heart, while at the same time being sufficiently general for primitive targeting tasks. Using this model, expert users could apply medical terminology when describing views in the 'Preset Views' interface yet the model could be just as easily interpreted by novices.

A 3D surface reconstruction of the phantom was generated using x-ray CT images with a slice thickness of 0.625mm, and exposure factors of 300mA and 70keV. The phantom with dimensions 16.5 x 15 x 9cm was placed in a plastic tank (23 x 19 x 14.5cm), immersed in water at room temperature.

To mimic a minimally-invasive closed-chest procedure, direct vision to the phantom

was removed by covering the phantom with surgical cloaking. Subjects were guided solely by the hybrid VR-US environment (Figure 5.2).

5.2.2 Interface Prototypes

Each of the three proposed interface prototypes, described in Chapter 4, was used as the method of data display and user interaction: (1) Preset View Angles, (2) Tracked Head-Mount Display (HMD), and (3) Tangible Camera.

5.2.3 The Task

The utility of the three proposed interaction techniques were assessed through a VR-enhanced US-guided targeting task. Probe-to-target placement, common to many image-guided procedures, required the user to reach and touch a specified location using a refined probe tip to the nearest 0.01mm. This simulated task was executed using a surgical probe within a customized phantom that represented the cross-coronal section of the four heart chambers. Such a surgical simulation was appropriate for cardiac surgery applications, yet sufficiently abstract for general MI applications using VR-enhanced US-guided techniques.

5.3 Study Design

5.3.1 Subjects

Eight subjects (4 male and 4 female graduate students) performed a probe-to-target positioning task to locate six targets embedded within the PVA-C phantom. With no direct view of the phantom, subjects were asked to rely solely on the hybrid VR-US environment for guidance in each of the interface conditions. All subjects had normal or corrected-to-normal vision.

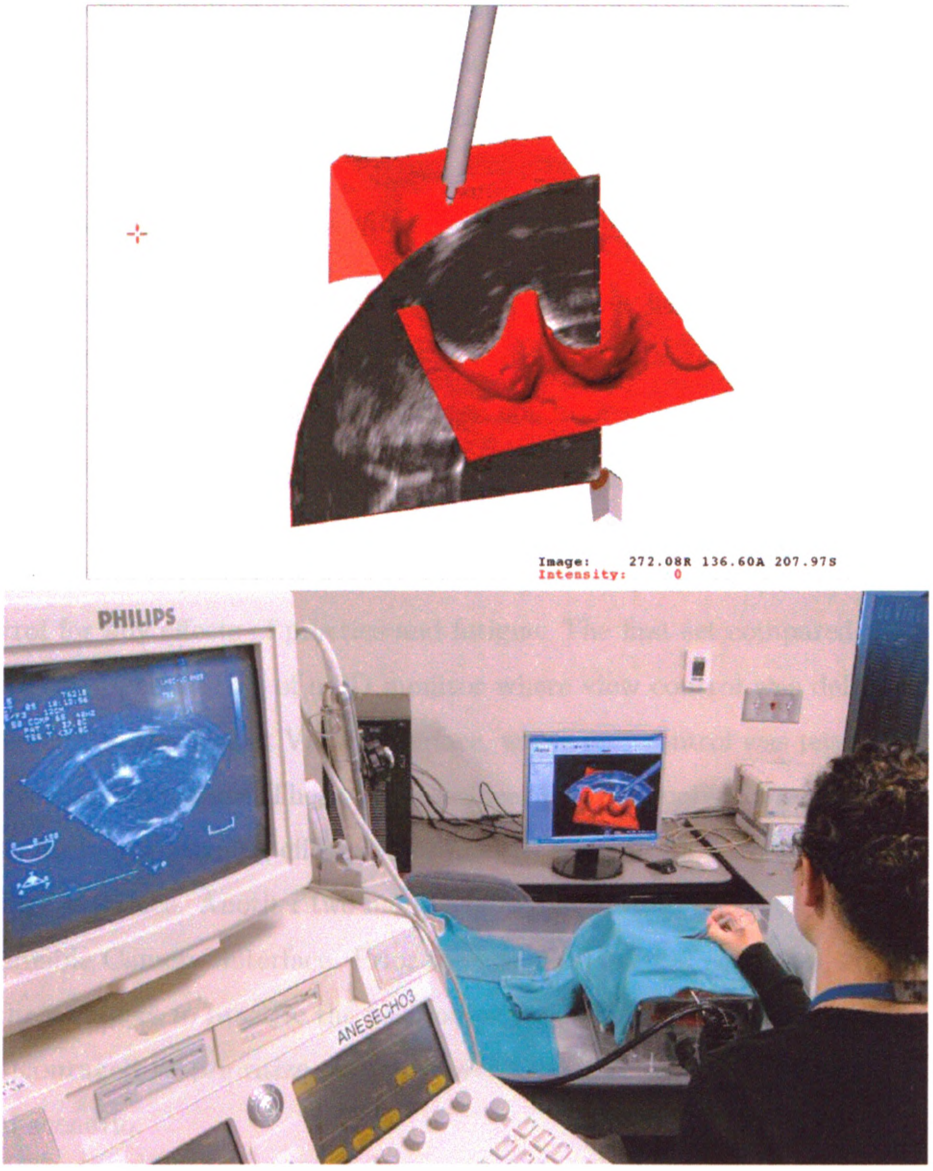

Figure 5.2: *Top*: Surface model of the PVA-C phantom in the VR-US hybrid environment. *Bottom*: Experimental set-up showing the user who is reaching for a target under VR-US guidance.

Using a power of 0.8, probability of Type I error (alpha) = 0.05, and probability of Type II error (beta) = 0.2, the use of eight subjects was deemed acceptable for a *within-subjects* experiment to determine any significant effects. For detail on the sample size calculation, please see Appendix C.

5.3.2 Experimental Design

A within-subjects study was conducted to reduce the effect of any predisposed subject preferences, and to account for variance in spatial reasoning abilities among subjects. To moderate interference between conditions, multiple sets of experiments were conducted using control techniques such as counterbalancing and temporal separation.

Three sets of experiments were conducted, each separated by 2-month intervals to control for any effects of practice and fatigue. The first set compared the existing interface setup, consisting of a 2D monitor where view control was delegated to an intermediary, to the 'Preset Views' interface, where user control was regained through simple verbal commands. These two interfaces were presented in a counterbalanced order to overcome learning effects. Two months later, in the second set, users tested the 'HMD' interface. Another two months later, in the third and final set, users tested the 'Tangible Camera' interface. Prior to each experimental condition, subjects were given three targets to practice the positioning task. During the testing trials, subjects were encouraged to prioritize accuracy over timing as would be preferred in a real surgical scenario.

5.3.3 Performance Measures

Subject performance was evaluated by the following variables: (1) Targeting accuracy; (2) total task completion time; (3) navigation and (4) positioning time; and (5) amount of workload experienced.

Targeting Accuracy. Actual target location ($\overline{p}_{bead}$) was determined by taking the mean end-tool coordinates gathered from 32 direct vision probe-to-target placements per target (one after each subject trial for each condition). The resulting mean provides a "gold standard" target location that addresses the accuracy offset inherent of the glass bead radius.

Subjects were asked to navigate and position the probe as close to the target using only the information displayed by the VR-enhanced US environment. To evaluate the outcome of the probe-positioning task, the Euclidean distance error, $\left|\overline{D}_{err}\right|$, between the tracked end-probe tip position, $\overline{p}_{tip} = (x_{tip}, y_{tip}, z_{tip})$, and the actual bead location, $\overline{p}_{bead} = (x_{bead}, y_{bead}, z_{bead})$, was measured to the nearest 0.01mm (Equation 5.1).

$$\left|\overline{D}_{err}\right| = \sqrt{(x_{tip} - x_{bead})^2 + (y_{tip} - y_{bead})^2 + (z_{tip} - z_{bead})^2} \tag{5.1}$$

Accuracy, a measure of trueness and precision, was calculated using the RMS error (RMS_{err}) as shown in 5.2.

$$RMS_{err} = \sqrt{\sum_{i=1}^{n} \frac{\left|\overline{D}_{err}\right|^2}{n}} \tag{5.2}$$

According to collaborating surgeons, a 2mm margin of error was clinically acceptable. However, for the purpose of testing tools in development, all error measurements were used to enable comparison between interface prototypes.

Total Task Completion Time. Total time to completion was the univariate global measure of motor behaviour [91] encompassing both the time and motion needed to 'navigate' and 'position' the probe to the target. Total task completion time, measured to the nearest 0.01 second, was recorded from when the subject held the probe above the tool entry point, to the point when the subject felt he or she had positioned the probe over the target. Although, this metric could assess the overall motor behaviour, further analysis was used to describe the behaviour observed in the 'navigation' and 'positioning' phases.

Navigation and Positioning Time. The total completion time was separated into 'Navigation Time' and 'Positioning Time'. By manually inspecting each user's velocity-distance-time profiles, the pattern of each user's trajectory formulated the criteria for separating the total task into 'navigation' and 'positioning'. Areas of low, steady velocities where distance-from-target exhibited minimal change was associated with fine motor movement in tool positioning. On the contrary, areas of high, fluctuating velocities and where large and rapid distance changes from the target existed were characteristic of the ballistic motion in navigation. The time instance at which extreme values of instantaneous velocity and distance drop to a relatively low and steady pattern was considered to be the 'Switch Over' point. This is the point where I anticipate the user has moved from 'navigation' to 'positioning'. Figures 5.3 and 5.4 illustrate examples of my methodology.

The time elapsed between the tool's starting position to the 'Switch Over' point was deemed to be the user's navigation time. The remaining time, between the 'Switch Over' point to the end-tool location, was the positioning time.

Task Workload. Commonly used to measure workload in numerous domains such as civil and military aviation, driving, air traffic control and nuclear power plant control systems [92], the NASA task load index (TLX) test enables quantitative analysis of subjective workload. Recently, this test has also been applied to healthcare research and has been adopted by Crossan *et al.* [93] in VR medical training studies.

The test is conducted through a computer program. The user determines how much difficulty they experienced by dragging a graphical slider to the left (least difficult) or right (most difficult) for the following factors: mental, physical and temporal demands, as well as the performance, effort and frustration experienced. Users also assign a weight to each of these factors to indicate the extent they felt each had contributed to the difficulty of the task. The result is the TLX score, a weighted rating of six workload factors experienced by the subject [94].

I used the TLX technique to quantify the amount of workload experienced by each

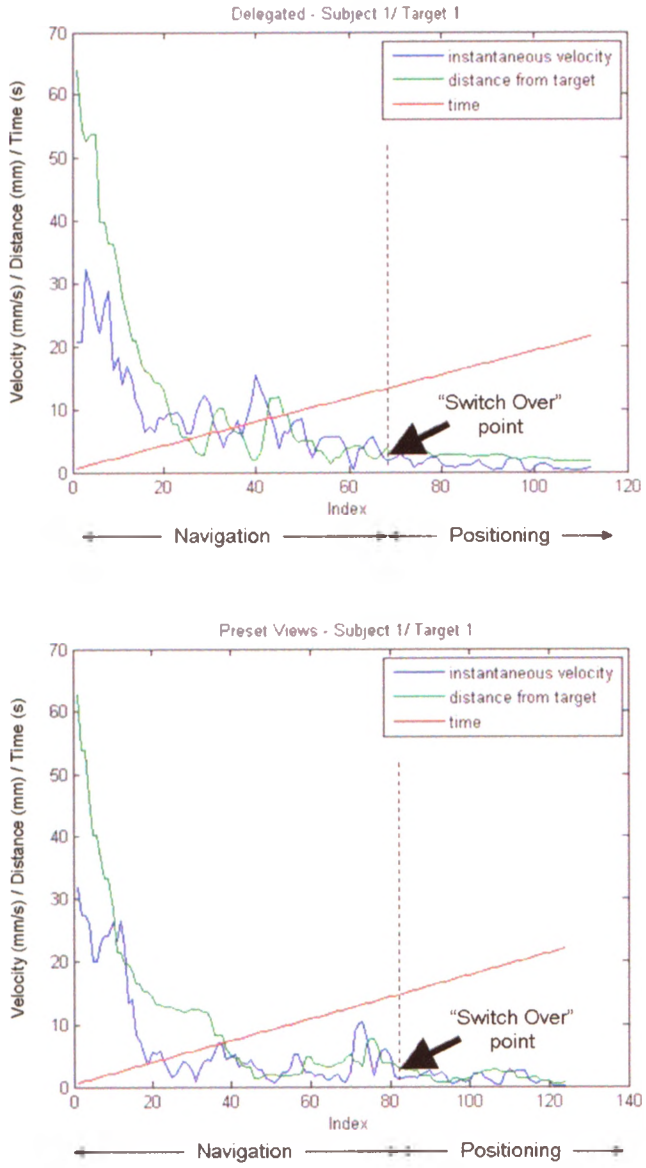

Figure 5.3: The velocity-distance-time profiles for Subject 1 and Target 1 for (from *top* to *bottom*): (1) 'Delegated' and (2) 'Preset Views'. Labeled on the profile are the 'Switch Over' points, 'Navigation' and 'Positioning' phases determined by a consistent separation method.

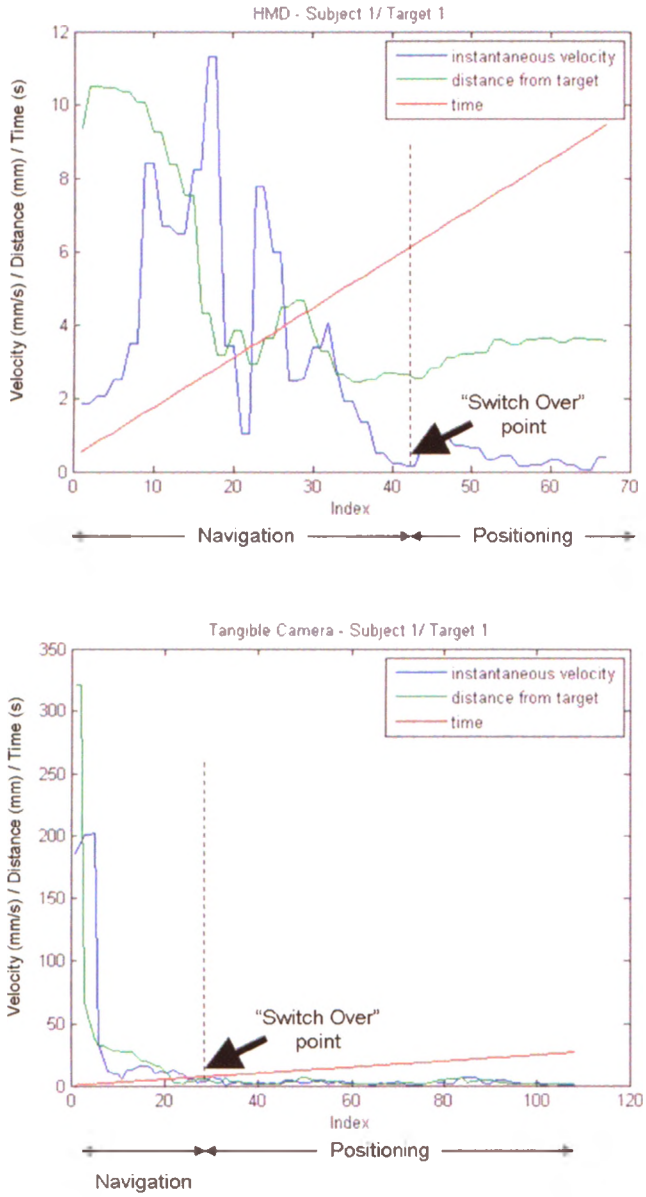

Figure 5.4: The velocity-distance-time profiles for Subject 1 and Target 1 for (from *top* to *bottom*): (1) 'HMD' and (2) 'Tangible Camera'. Labeled on the profile are the 'Switch Over' points, 'Navigation' and 'Positioning' phases determined by a consistent separation method.

subject for each interface display. The assessment was conducted after each set of experiments. Users were asked to rate their experience with each interface independent of one another.

5.3.4 Data Analysis

The same data set was used for the user-centered and comparative evaluations. Because each evaluation differed in its objectives, they were distinguished by separate hypotheses and results analysis. The two-tailed Wilcoxon Signed-Rank test was used to compare each interface with the 'Delegated' condition (user-centered evaluation), and also to compare the new interface conditions (comparative evaluation). Using the Wilcoxon Signed-Rank statistic in [95], p-values were determined for n=8. Results, shown in Table 5.1, were considered statistically significant at the $p<0.05$ level.

5.4 Results - User-Centered Evaluation

Targeting Accuracy. Subjects demonstrated a 39% improvement in accuracy using the 'Preset Views' interface (1.78±0.73mm) compared to the 'Delegated' condition (2.90±1.1mm) ($p<0.05$). While the use of the 'Tangible Camera' resulted in a lower mean RMS error (2.04±0.61mm), the difference was not significant ($p=0.2$). No improvement in accuracy was seen when using the 'HMD'. Figure 5.5 shows the RMS error for all interface conditions.

Total Task Completion Time. Subjects performed 26% faster using the 'HMD' (25.1±6.3s) compared to the 'Delegated' setup (33.8±7.4s) ($p<0.01$). Although average completion time was 13% faster using the 'Preset' condition (29.1±9.2s) compared to the 'Delegated' condition, this difference was not significant ($p=0.2$). The comparision of completion times is shown in Figure 5.6. No significant difference was found between 'Tangible Camera' and the 'Delegated' condition ($p>0.2$).

Navigation and Positioning Time. For the navigation phase, subjects

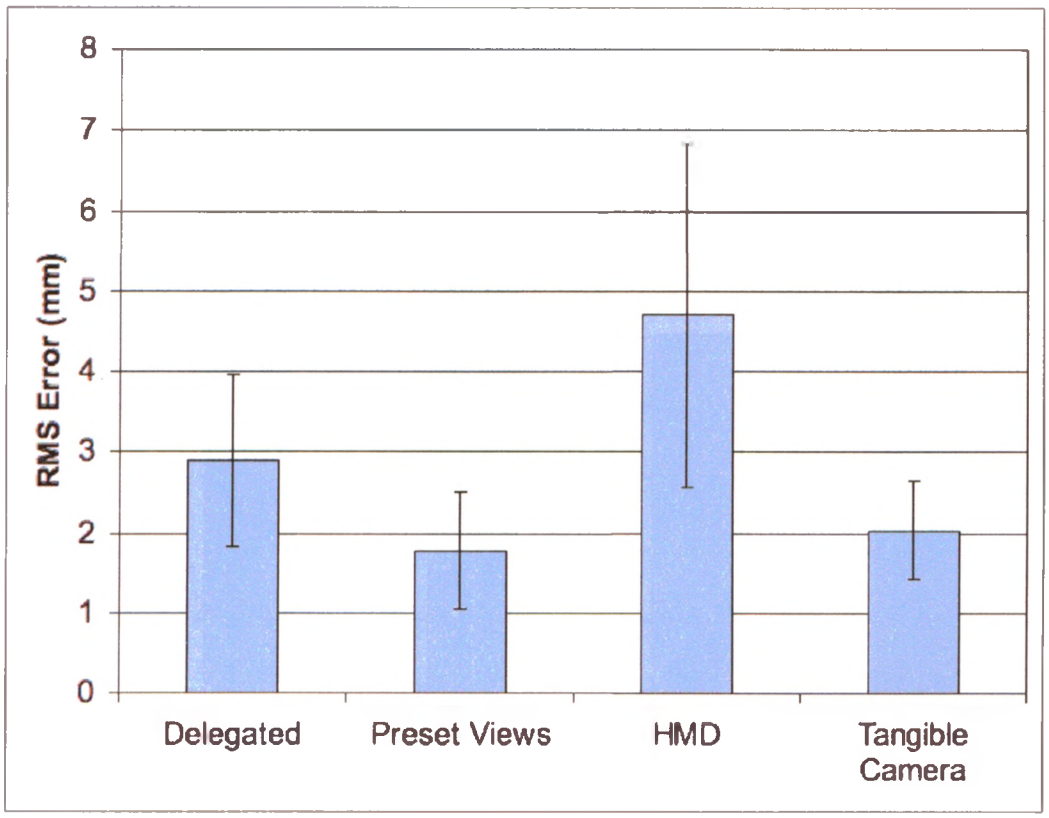

Figure 5.5: Mean accuracy for all four interface conditions. Y-Error bars represent the 95% confidence interval of the RMS error.

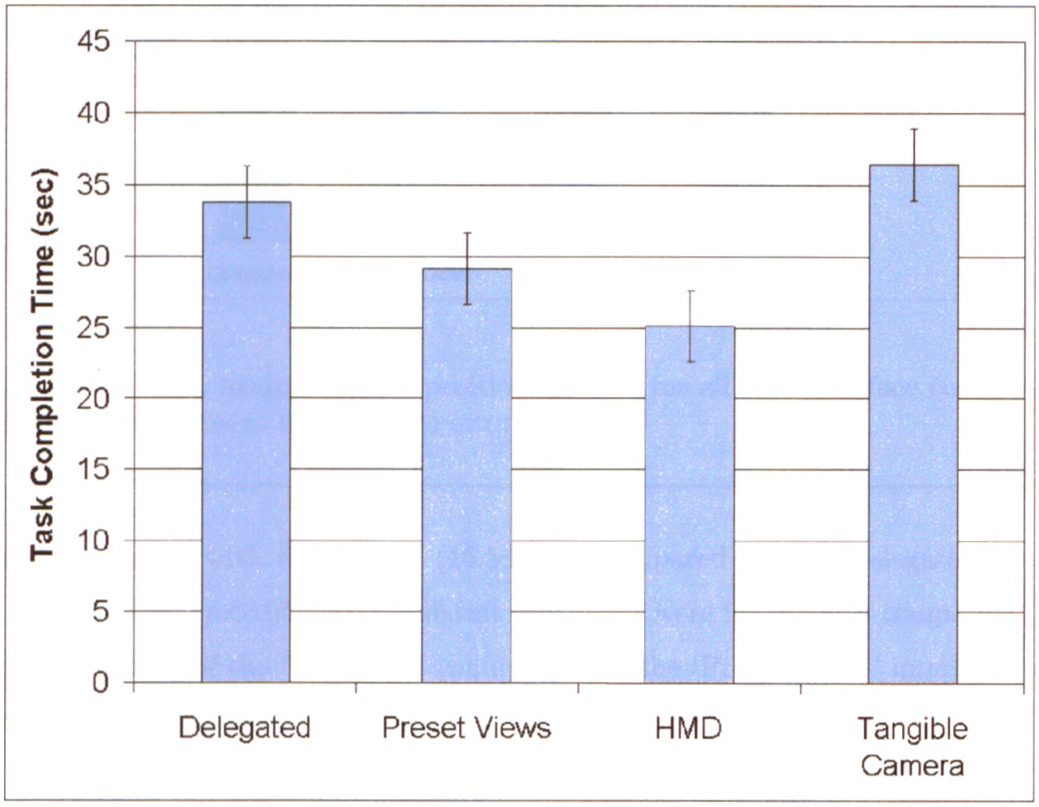

Figure 5.6: Mean task completion time for all four interface conditions. Y-Error bars represent the standard error.

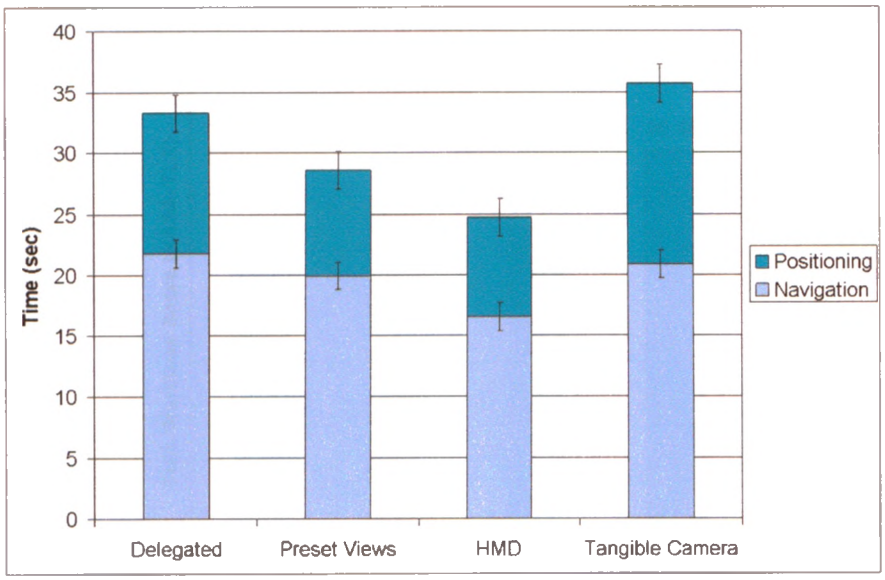

Figure 5.7: Mean navigation and positioning timg for all four interface conditions. Y-Error bars represent the standard error.

---

were 24% faster with the 'HMD' (16.5±3.3s) compared to the 'Delegated' interface (21.8±6.5s) ($p<0.05$). No significant differences were found when comparing the navigation time of the 'Delegated' condition with the 'Preset Views' interface, nor with the 'Tangible Camera' interface.

In terms of positioning, subjects were 29% faster using the 'HMD' (8.21±4.2s) compared to the 'Delegated' setup (11.5±4.2s) ($p<0.05$). Similarly, significant differences in positioning were not observed between the 'Delegated' and 'Preset Views' interfaces, nor the 'Delegated' and 'Tangible Camera' interfaces. The graph in Figure 5.7 shows the average navigation and positioning times recorded for each condition.

Task Workload. No significant workload differences were observed between any of the new interfaces and the 'Delegated' condition ($p>0.2$ for all comparisons, except 'Tangible Camera' where $p=0.15$). Workload rating scores are displayed in Figure 5.8.

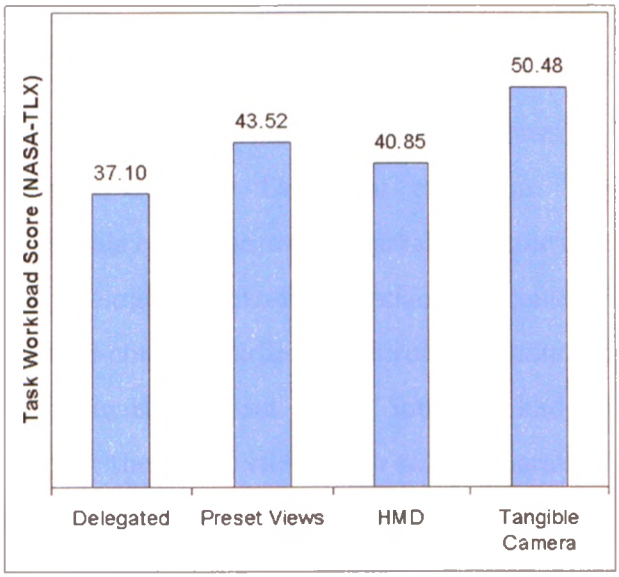

Figure 5.8: Mean task workload index score calculated based on the NASA-TLX test, an assessment which combines mental, physical and temporal load with frustration, effort and performance.

Table 5.1: Performance results for each of the four interface conditions.

	Delegated	Preset Views	HMD	Tangible Camera
Accuracy (mm ± 95% Confidence Interval)	2.90 ± 1.1	1.78 ± 0.73	4.70 ± 2.1	2.04 ± 0.61
Total Task Completion Time (s)	33.8 ± 7.4	29.1 ± 9.2	25.1 ± 6.3	36.4 ± 11
Navigation Time (s)	21.8 ± 6.5	19.9 ± 7.1	16.5 ± 3.3	20.9 ± 9.6
Positioning Time (s)	11.5 ± 4.2	8.69 ± 3.0	8.21 ± 4.2	14.9 ± 7.8
Workload Rating (TLX score)	37.1	43.5	40.9	50.5

5.5 Discussion

5.5.1 Preset Views vs. Delegated

In terms of efficiency and ease of use, the 'Preset Views' did not manifest any exceptional improvements over the 'Delegated' condition. However, because the accuracy attained using this condition was remarkably better than the existing setup, the design of this type of surgeon-controlled interface was considered empirically valid.

Through the keyword-coupled indirect control and monoscopic biplanar views, subjects performing under the 'Preset Views' interface had the highest targeting accuracy. By customizing their own views prior to the operation, subjects appeared to have a better understanding of the 3D perspective of the target and were able to optimize the view of the US fan prior to the task.

It was noted that with the 'Preset Views', mean task completion time was faster than the 'Delegated' condition. Although not significant, this observation suggested that less camera motion and limited vantage points could actually facilitate, if not improve, targeting compared to an interface that supported 6 degrees of camera motion and unlimited perspective.

Because the 'Preset Views' did not introduce any new equipment or require any behavioural adjustments to the typical workflow, no difference was found in workload rating score between this interface and the 'Delegated' setup. The keyword-coupled indirect control and monoscopic biplanar views would be easily integrated into any OR with a standard computer monitor.

5.5.2 HMD vs. Delegated

Subjects performed significantly faster using the direct head-coupled view control of the 'HMD' compared to the existing setup. This demonstrated that the entire process of targeting, from navigation to positioning, could be hastened without having

to depend on an intermediary to change the view angle.

No significant difference was observed between targeting accuracy for the 'HMD' and the 'Delegated' condition. This was consistent with results from Maithel *et al.* [56], who also found no difference in task performance when using a head-mount display versus a traditional video monitor for simulated laparoscopic procedures. However, the efficacy of the 'HMD' was unclear as the average RMS error achieved with the 'HMD' was worse than the 'Delegated' condition and in fact, nearly double the tolerable 2mm margin of error. This outcome exemplified the dangerous consequences of enabling direct view control without adequate depth display. Although subjects felt confident that they had achieved the target quickly using the 'HMD', they were consistently incorrect in their estimation of depth. This may have been attributed to the poor visibility of the US fan image due to the inferior color contrast and screen resolution of the HMD.

Despite subjects' comments on the "usability" and "intuitiveness" of the 'HMD', its poor accuracy results have disqualified it from being an optimal surgeon-controlled interface for VR-enhanced US-guided cardiac surgery. However, stereoscopic HMD technology may still be clinically useful in surgical environments already accustomed to head-mounted instruments, such as frameless stereostatic neurosurgery [45], or for complex 3D tasks such as navigation through complex vasculature [81]. The design of suitable HMDs should consider the design improvements described at the end of this section.

5.5.3 Tangible Camera vs. Delegated

On average, the targeting accuracy with the 'Tangible Camera' interface was better than the existing setup, although the difference was not significant. The ability to support direct view adjustments and present depth information using monocular biplanar views proved to offer adequate depth information.

The workload rating for this method was the highest among all other interfaces

including the 'Delegated' condition. Subjects reported that the arm mounting the physical camera tool was too rigid and thus, difficult to maneuver smoothly. Aside from the hardware implementation of the prototype, the concept of the 'Tangible Camera' was well received by all subjects.

Subjects enjoyed having *ad hoc* flexibility to change the viewing angle without having to wear head-mounted equipment. This eagerness could have likely contributed to the long navigation time observed for the 'Tangible Camera'. Subjects manipulated the view at the beginning of the task even though camera movement was unnecessary.

Some subjects developed a successful strategy that involved positioning the camera to the ideal view of the US fan and target prior to starting the task; then slowly repositioning the camera with minor adjustments to view the distance gap between the probe and target. Users slowly moved the camera towards an orthogonal view of their probe tip while relying on the fixed stable reference on the smaller right panel for guidance. Despite this enhanced form of user control during positioning, little difference was observed between the positioning time for this interface and the existing. As a matter of fact, positioning time was the highest for this interface. It is possible that the physical effort required to manipulate the tangible camera may have been at least as time consuming and as difficult as mentally verbalizing view changes.

The 'Tangible Camera' is comparable to the existing image-guided configuration. However, with further design improvements, it is possible that this surgeon-controlled interface might outperform the current 'Delegated' control interface.

5.5.4 Design Considerations

Only under the 'Preset Views' interface was subject performance within the acceptable 2mm margin of error. For all other conditions, the accuracy was not clinically successful. This may have been due to the subjects' lack of experience with performing VR-enhanced US-guided tasks. Based on results seen with each of the three surgeon-controlled techniques compared with the standard delegated control

setup, it would appear that if constructed appropriately, surgeon-control would be feasible and have the potential to improve user performance.

A few considerations for the next iteration of designs were:

- ***Auto-sweeping and/or rotating US fan.*** The act of having to precisely adjust the fan angle with each change in view of the VE was time consuming. An improvement could be made by incorporating a 3D TEE with an automatically sweeping or rotating US fan to reconstruct a 3D image [96]. For 'Preset Views' and 'Tangible Camera', multiplanar reformatting of 3D US images can faciliate the viewing of the US image at the chosen biplanar angles [97].

- ***Flexible arm for Tangible Camera.*** To improve the fluidity of the camera arm movement in the 'Tangible Camera' interface, the construction of the camera arm should be implemented with more flexible materials.

- ***ON/OFF for HMD tracking.*** Because the motion tracking of the 'HMD' may not be beneficial for the navigation phase, allowing the user to turn the tracking on and off could make its use more efficient.

- ***Smooth camera motion.*** A predictive tracking algorithm can be implemented in the next design iteration for both direct control interfaces (i.e 'HMD' and 'Tangible Camera'). Established methods, such as Kalman and extended Kalman filter-based predictors [98], can be used to smooth the perceived latency between user motion and display.

5.6 Results - Comparative Evaluation

Using the same data in Table 5.1, accuracy results among the new surgeon-controlled techniques were compared.

Targeting Accuracy. Subject accuracy was highest for both the 'Tangible Camera' and 'Preset Views' condition. Compared to the 'HMD' interface (4.70±2.1mm),

accuracy improved by 62% using the 'Preset Views' interface (1.78±0.73mm)(p<0.05). When using the 'Tangible Camera' interface, accuracy improved by 57% (2.04±0.61mm)(p<0.01). No notable differences were observed between 'Preset Views' and 'Tangible Camera'.

Total Task Completion Time. Total completion time was 45% longer when using the 'Tangible Camera' (36.4±11s) compared to the 'HMD' (25.1±6.3s)(p<0.05). A closer inspection shows that subjects required more time for tool 'positioning' when using the 'Tangible Camera'. Details are explained in the following sub-section.

Navigation and Positioning Time. There were no significant differences found in navigation time for any of the comparisions. When comparing positioning time, the 'HMD' (8.21±4.2s) was found to facilitate quicker positioning (45% faster) compared to the 'Tangible Camera' (14.9±7.8s) (p<0.05). 'Preset Views' (8.69±3.0s) also demonstrated a significantly faster (41%) positioning time compared to the 'Tangible Camera' (p<0.05). No difference was found between the positioning time for the 'HMD' and 'Preset Views'.

Task Workload. No differences were observed in task workload scores between the stereoscopic display of the 'HMD' and the monoscopic biplanar display of the 'Preset Views' and 'Tangible Camera'. Although 'Preset Views' and 'Tangible Camera' received the highest TLX score relative to the other interface conditions, significant differences were not found between these conditions.

5.7 Discussion

5.7.1 Stereoscopic vs. Monoscopic Biplanar

The remarkable improvement in targeting accuracy when using the monoscopic biplanar displays of both the 'Preset Views' and 'Tangible Camera' revealed the benefit of two different viewing angles, whether orthogonal or not, and challenged the need for stereo disparity.

With or without stereo disparity, graphical presentation of perspective and object occlusion offer depth cues in a VR environment. Yet during targeting, as the VR tool approaches target, the tool occludes the view of the target, hampering the user's accuracy. Displaying an angled view of the probe at all times, the depth from VR probe to VR target can be readily perceived. In fact, most real-world tasks are not fully 3D, and it has been demonstrated in VR studies that 2D tasks are cognitively simpler than 3D tasks [90]. Thus, by providing 2D alternatives, one could increase the usability of 3D image interpretation.

Moreover, the success of the 'Preset Views' interface could be attributed to the layout of the biplanar display itself. The subject was able to focus on the chosen view, located on the left, referring to the orthogonal angle as needed, on the right.

Cao *et al.* [86] found that simultaneous use of wide and narrow FOV is associated with better performance in MIS. Having a supplemental close-up probe-to-target view in the lower right-corner offered the same advantage. This was consistent with the higher accuracy results found for the 'Preset Views' and 'Tangible Camera' interfaces.

Despite improvements in task performance, the 'Preset Views' and 'Tangible Camera' received higher workload rating (TLX) scores compared to the standard configuration. In both cases, the monocular biplanar views could lead to higher mental workload associated with this presentation of depth. Greater mental activity is typically correlated with the process of merging two or more frames of reference simultaneously. In a psychophysical study, Klatzy *et al.* [43] reports that having to align disparate frames of reference increases cognitive load. Thus, mental stress may have been experienced when using two monocular biplanar views to display 3D information as the subject must mentally merge the two panes to identify spatial relationships.

On the other hand, the increased cognitive load imposed by the monocular biplanar views, might counter-intuitively serve as a safety feature. The visuomotor task of navigating and positioning the probe to the target location using two adjacent but

dissimilar views might afford a heightened sense of concentration and caution on behalf of the surgeon. The opposite to this concept has been demonstrated with the 'HMD'. An interface that was easy to use or appeared intuitive, might not always be the most optimal design, especially if the information was not presented in an appropriate manner.

5.7.2 Keyword, Head and Hand-coupled Control

The subjects' trajectories provided insight into the performance benefits and usability of keyword-coupled, head-coupled, and hand-coupled control modalities.

As differences were not found for navigation time, significant differences in total task completion time were attributed to positioning only. Positioning time was fastest for the 'Preset Views' and 'HMD' with no notatble difference between them. However, the positioning times for the 'HMD' and the 'Preset Views' interfaces were both remarkably faster than the 'Tangible Camera'. Cao *et al.* [86] demonstrated that with the ability to see both narrow and wide FOV easily, performance in tool positioning can improve. Keyword-coupled and head-coupled control might offer an easier and more intuitive method for zooming in and out of wide and narrow FOVs compared to hand-coupled control. Although both the layouts of the 'Preset Views' and 'Tangible Camera' displayed a constant narrow FOV in one of the panes, the hand-coupled control provided by the 'Tangible Camera' required more physical effort for view changing. This physical effort, likely due to hardware limitations, could explain the longer amount of time required to position the probe onto the target.

Identifying which control method could provide optimal 'positioning' ability would be particularly useful for designing interfaces to support complex surgical tasks such as suturing and knot-tying. Under these tasks, the user would already operating within the area of the 'positioning' phase - the target vicinity.

5.8 Conclusion

The user-centered evaluation and comparative evaluation revealed:

- The potential for surgeon-controlled interfaces for VR-guided cardiac surgery;
- The necessity of design improvements for each prescribed user interface;
- The advantage of monoscopic biplanar vision over stereovision for the task of probe-positioning within a cardiac phantom; and
- The advantage of using keyword-coupled indirect control over head-coupled and hand-coupled control.

Overall, the current design of the 'Preset Views' interface was considered most suitable for clinical integration. Based on navigation time, 'Preset Views' performed comparably with the existing setup. Although it was not proven that less camera rotation and fewer view manipulations could actually hasten navigational ability, there was no evidence against this proposition. Due to it's simple implementation, this interface required no additional equipment nor adjustments to workflow. Although the monocular biplanar (orthogonal) view angles might be difficult to merge initially, this multiple-view layout may afford user caution as demonstrated by the improvement in accuracy. Compared to the 'HMD' and 'Tangible Camera', positioning ability using the 'Preset Views' was respectively more accurate and efficient.

The final chapter of this thesis discusses the contributions and limitations of the presented work and suggests future directions for study.

Chapter 6

Contributions, Limitations and Future Work

6.1 Contributions

The work described in this thesis contributes to the advancement of interface design for virtual reality-enhanced image guided surgical platforms. Contributions can be grouped into three areas: understanding of design parameters, display development, and a methodology that can be used for the future interface development in the VASST lab.

The evolving VR-enhanced surgical platform was used for creating, validating and refining user interfaces and interaction techniques. Interface requirements were extracted from task analyses and participatory design meetings with domain experts. User-based studies uncovered which of three design parameters (i.e. control type, disparity and spatial context) or combination of parameters were responsible for accurate depth perception, faster task completion and user confidence in task performance. The findings from these task, domain and user-based studies have all contributed to the understanding of what is required to build an optimal user interface geared for valve and patch implantation in a cardiac intervention. As a result, three user

interface techniques were developed as alternatives to the current UI. Based on the current platform's ability to display tracked surgical instruments relative to a fused VR-US environment, the new interfaces offer enhancements such as:

- The ability to set and toggle between medically relevant preset views
- The ability to control viewpoint using a tracked HMD; and
- The ability to control viewpoint using an external hand-held tool (i.e. a tangible camera)

User-centered and comparative evaluation demonstrate the efficacy of these interface techniques in aiding tool navigation and positioning. Design enhancements were conceived by usage observation and gathering qualitative feedback from test participants.

Empirical findings demonstrate the monoscopic biplanar views could outperform the use of a stereoscopic head-mount display. Indirect keyword-coupled control provided by 'Preset Views' was found to be the optimal method of autonomous control based on accuracy and task completion time outcomes. After this first design iteration, the benefits of surgeon-control techniques are foreseeable so long as the proposed enhancements are made.

Methodological innovations include an automated random target display module and the capability to capture and process path trajectory information in terms 'navigation' and 'positioning' phases. These additions have created a testing environment that puts emphasis on human factors.

6.2 Limitations

Consideration of the limitations in the hardware, experimental design and statistical analysis, is critical for the overall understanding of the results and conclusions drawn.

6.2.1 Hardware

It is conceivable that the materials used to construct the 'Tangible Camera' prototype had an impact on user response and thus performance. The rigidity of the camera arm may have increased physical workload and caused mild frustration for the user as evidenced by the high TLX score. Improvements to the hardware may lead to faster performance and better impression of the technology.

6.2.2 Experimental design

The choice of subjects may have influenced the workload TLX scores for each of the interface conditions. For example, although the 'Preset View' condition had the best accuracy, it received the highest TLX score. This may be due to the fact that all subjects were graduate students in engineering and medical biophysics with little to no medical training. As a result, they may have been more accustomed to the free-form manipulation of 3D computer graphics than the stationary views, characteristic of medical diagrams. Results may have differed if medical students, residents and clinicians were recruited as study participants.

Furthermore, the findings of this study are limited to the probe-positioning task used for assessing performance. For the purpose of a mitral valve replacement or ASD patch implantation, this particular task is appropriate. Yet, most cardiac interventions involve more than simple tool-to-target navigation and positioning. According to a collaborating surgeon, "once [the surgeon] has reached the target, this is the point where the surgery begins". This suggests that different tasks may reflect different levels of performance on certain interfaces. For instance, the use of autonomous surgeon control may have been too complicated for the targeting task, a straightforward probe-positioning task that could easily be performed under 'Delegated' control. A more challenging and realistic scenario consisting of either complex anatomy or a physically restrictive entry point, may have exposed greater performance advantages

for all three of the surgeon-control conditions.

For this study, the emphasis was placed on accuracy over speed. Consequently, the analysis of each condition on navigation did not show any significant effects compared to positioning. The data used to represent tool navigation is not as reliable for analysis compared to the data used to represent tool positioning. A navigation-specific task should be introduced in order to draw solid conclusions about navigation. Participants should be instructed to begin at a standard starting point and move toward the target as quickly as possible, valuing speed over accuracy.

6.2.3 Statistical Analysis

Interpretation of empircal results requires acknowledgement of the low statistical power of some of the tests. In both the user-based study and interface analysis, more participants may have provided greater statistical power for the observed effects. However, the study was conducted within-subjects. This allowed for stronger comparisons to compensate for a small sample size and to account for any subjective bias between subjects. Thus, the comparisons made should demonstrate reliable effects, though these effects are smaller in size than expected.

6.3 Future Work

This thesis elaborates findings from the first iteration of my design strategy. The presented work sets a foundation for solving the larger problem - the development of an optimal surgeon-computer interface for VR-enhanced cardiac surgery. Subsequent design iterations following this work should include:

- Re-validation of the presented conclusions with collaborating clinicians. Currently the 'Preset Views' interface is the recommended method to display and deliver VR-enhanced information. This resolution should be discussed with the

clinicians in a second participatory design meeting similar to the one described in Chapter 2.

- Refinement of the surgeon-controlled user interfaces using the proposed enhancements elicited from surgeon feedback and empirical evaluation results.
- Design and evaluation of the second category of interface enhancements, 'The Visual Overlays', as described in Chapter 2.
- An experiment that involves a more challenging experimental task. The task may involve a more complex phantom that better resembles a realistic anatomical case or involve complex motor control such as longer navigation through a 3D structure.
- An experiment with a navigation-specific component. Instructions given will impose stricter regulations on a set starting point and emphasize speed over accuracy. For this task, the use of VR context can be revisited in a user-based study, similar to Chapter 3, as well as an evaluation of interface prototypes.
- An experiment that involves a reciprocal tapping task can account for the speed and accuracy tradeoff under different visual display conditions. By applying 3D Fitts' Law [99], one can determine the index of difficulty for various displays.
- Recruitment of medically trained individuals such as surgeons, domain specialists, residents, and medical students to participate in the evaluation of the next iteration of surgical interface prototypes. This may reduce the average RMS error to at least 2mm, the clinically acceptable margin of error.

By conducting another iteration of the design strategy, the VASST laboratory will be one step closer to attaining a more integrated and usable interface for valve replacement and patch implatation using VR-enhanced US-guided techniques. In addition, the structure of my design methodology encourages a deeper understanding of

image perception and interaction generalized to any VR-enhanced surgical application. There is no doubt that conducting human factor studies, along side systems engineering, is essential to the development of usable and thus, effective surgical technology.

Appendix A - MTS Registration Transformation

The Aurora magnetic tracking system (MTS) was used to track tools and the US. All surgical tools used had 6 DOF sensors located at the tool tip. In addition, 6 DOF sensors were used to track the US transducer and to provide a fixed reference for the phantom. The US transducer was calibrated using a Z-bar phantom using 16 screen captures with an RMS error of 1.55mm.

By default, the MTS has an inherent coordinate frame based on the magnetic field generator. As a result, all tool coordinate vectors are set to be relative to the field generator.

To maintain consistency between the tracked tool coordinates and the phantom coordinate frame, regardless of field generator motion, we must transfer the base reference to a fixed rigid body attached to the phantom tank. This is known as a reference tool.

The following equations define the transformations required to measure accuracy.

$$\overline{p}_{mts} = T_{ref}^{-1}\left(\overline{p}_{fg}\right), \tag{6.1}$$

where $(\overline{p}_{mts})$ represents the default MTS tracked tool coordinates, and T_{ref} represents the transform of the reference tool. The tracking coordinate frame was registered to the VR model using a tracked sphere-tipped tool that mated with the divots on the outside of the phantom box, generating a very accurate point-based registration.

Using this point based registration transform ($T_{phantom}$), we translated the measured needle coordinates ($\overline{p}_{mts}$) into the VR coordinate frame:

$$\overline{p}_{meas} = T_{phantom}\left(\overline{p}_{mts}\right), \tag{6.2}$$

where $\overline{p}_{meas}$ is the measured needle tool tip location VR space. Trueness is then the distance defined by the magnitude of the error vector,

$$d_{err} = \left|\overline{p}_{error}\right| = \left|\overline{p}_{meas} - \overline{p}_{target}\right|, \tag{6.3}$$

where $\overline{p}_{target}$ is the closest point on the VR phantom model as defined in the VR space.

Appendix B - Post-Study Questionnaire

1. Do you have any vision conditions (i.e. colour blindness, no depth perception, glasses)? If so, name:

2. If you need glasses, are you wearing them today? Yes or No

3. Do you have any experience with Virtual Reality surgery procedures? Yes or No

4. Using a scale of 1 to 5 (*1=least confident* and *5=most confident*), rate the effect of the following on your confidence level? Provide comments if possible:

Table 6.1: Visualization Parameters, Rating and Comments.

Effect	**Rating**	**Comments**
Monoscopic View		
Stereoscopic View		
Two orthogonal planes with no assisted view positioning		
One view with assisted view positioning		
Seeing "the box" AND the targets		
Seeing the target ONLYahahahahah and "no box"		

Appendix C - Sample Size Calculation

A pilot study was conducted using one subject performing six trials under four conditions (1 Control, 3 Experimental). This calculation was conducted to determine the sample size needed to observe significant effects according to two measured variables: targeting accuracy (distance error) and task completion time.

Based on similar human factor studies conducted by Yee [91], Du [100] and Morton [101], the following formula was used to estimate sample size:

$$n = \frac{(z_\alpha - z_\beta)^2}{(\mu - \mu_o)^2/\sigma^2} \tag{6.4}$$

$\alpha = 0.05$, probability of having a Type I error (false positive)

$z_\alpha = 1.645$, standard score corresponding to α

$\beta = 0.2$, power of $1 - \beta = 0.8$, probability of avoiding a Type II error (false negative)

$z_\beta = -0.845$, standard score corresponding to $1 - \beta$

The μ - μ_o and σ were estimated from the pilot study for each of the display modalities tested.

According to Cohen's d, the average effect size for all conditions was medium to large (ES=0.7). Due to the effect size and the within-subjects design of my user study,

Table 6.2: Sample size calculation - Results from a single user under four conditions (1 Control, 3 Experimental)

Condition	Mean (μ)	Mean$_o$ (μ_o)	Std. Dev. (σ)	Sample Size (n)
Accuracy (mm)				
Preset Views	1.54	2.32	0.82	6.69
HMD	2.03	2.32	0.92	61.59
Tangible Camera	1.91	2.32	0.86	26.56
Task Completion Time (sec)				
Preset Views	19.80	21.85	2.29	7.74
HMD	18.64	21.85	5.98	21.47
Tangible Camera	27.66	21.85	14.42	38.10

sample sizes of over 20 were considered to be prohibitively large.

Human factors studies by Du [100] and Morton [101] were conducted using sample sizes of n=6 and n=8, respectively. The similarity of my within-subjects design to that of Du and Morton has justified my sample size of n=8.

Bibliography

[1] R. Galloway and T. Peters. Overview and history of image-guided interventions. In T.M. Peters and Cleary K., editors, *Image-Guided Interventions: Technology and Applications*, chapter 1, pages 1–21. Springer, 2008.

[2] W.R. Chitwood Jr. and E. Rodriguez. Minimally invasive and robotic mitral valve surgery. In L.H. Cohn, editor, *Cardiac Surgery in the Adult, Third Ed.*, chapter 45, pages 1079–1100. McGraw-Hill, 2008.

[3] A. Mehta, C. Sapir, S. Shirazian, and D. Wobbekind. Minimally Invasive Cardiac Surgery. http://biomed.brown.edu/Courses/BI108/BI108_2000_Groups/Heart_Surgery/, 2000.

[4] H. Niinami, H. Ogasawara, Y. Suda, and Y. Takeuchi. Single-vessel revascularization with minimally invasive direct coronary artery bypass*. *Chest*, 127(1):47–52, 2005.

[5] E. Buffolo, J.N.R. Branco, L.R. Gerola, L.F. Aguiar, C.A. Teles, J.H. Palma, and R. Catani. Off-pump myocardial revascularization: Critical analysis of 23 years' experience in 3,866 patients. *Annals of Thoracic Surgery*, 81(1):85–89, 2006.

[6] H. Vanermen, F. Farhat, F. Wellens, R. Geest, I. Degrieck, F. Praet, and Y. Vermeulon. Minimally invasive video-assisted mitral valve surgery: From port-access towards a totally endoscopic procedure. *Journal of Cardiac Surgery*, 15(1):51–60, 2000.

[7] T. Murakami, M. Kuinose, M. Takagaki, and E. Inagaki. Mitral valve replacement through right thoracotomy after previous coronary artery bypass grafting. *Japanese Journal of Thoracic and Cardiovascular Surgery*, 52(1):26–29, 2004.

[8] D.P. Bichell, T. Geva, E.A. Bacha, J.E. Mayer, R.A. Jonas, and P.J. del Nido. Minimal access approach for the repair of atrial septal defect: The initial 135 patients. *Annals of Thoracic Surgery*, 70(1):115–118, 2000.

[9] D.M. Cosgrove III, J.F. Sabik, and J.L. Navia. Minimally Invasive Valve Operations. *Annals of Thoracic Surgery*, 65(6):1535–1538, 1998.

[10] J.I. Fann, N.B. Ingels Jr., and D.C. Miller. Pathophysiology of mitral valve disease. In L.H. Cohn, editor, *Cardiac Surgery in the Adult, Third Ed.*, chapter 41, pages 973–1012. McGraw-Hill, 2008.

[11] Lucile Packard Children's Hospital at Stanford. Health Library/Cardiovascular Disorders. http://www.lpch.org/DiseaseHealthInfo/HealthLibrary/cardiac/0435-pop.html, 2009.

[12] R.O. Bonow, B. Carabello, A.C. de Leon Jr., L.H. Edmunds Jr., B.J. Fedderly, M.D. Freed, W.H. Gaasch, C.R. McKay, R.A. Nishimura, P.T. O'Gara, R.A. O'Rourke, S.H. Rahimtoola, J.L. Ritchie, M.D. Cheitlin, K.A. Eagle, T.J. Gardner, A. Garson Jr., R.J. Gibbons, R.A. O'Rourke, R.O. Russell, T.J. Ryan, and S.C. Smith Jr. ACC/AHA guidelines for the management of patients with valvular heart disease: A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines (Committee on management of patients with valvular heart disease). *Journal of American College of Cardiology*, 32(5):1486–1582, 1998.

[13] D.J. McCormick. Atrial septal defect: Pathophysiology, diagnosis, and treatment. *Medscape Cardiology*, 2006.

[14] R.J. Craig and A. Selzer. Natural history and prognosis of atrial septal defect. *Circulation*, 37(5):805–815, 1968.

[15] H. Laks, D. Marelli, M. Plunkett, and J. Myers. Adult congenital heart disease. In L.H. Cohn, editor, *Cardiac Surgery in the Adult, Third Ed.*, chapter 61, pages 1431–1464. McGraw-Hill, 2008.

[16] H.J. Geiler, C. Schlensak, M. Südkamp, and F Beyersdorf. Heart valve surgery today-indications, operative technique, and selected aspects of postoperative care in acquired valvular heart disease. *Deutsches Äerzteblatt International*, 106(13):224–233, 2009.

[17] T.K. Rosengart, T. Feldman, M.A. Borger, T.A. Vassiliades Jr., A.M. Gillinov, K.J. Hoercher, A. Vahanian, R.O. Bonow, and W. O'Neill. Percutaneous and minimally invasive valve procedures: A scientific statement from the American Heart Association Council on Cardiovascular Surgery and Anesthesia, Council on Clinical Cardiology, Functional Genomics and Translational Biology Interdisciplinary Working Group, and Quality of Care and Outcomes Research Interdisciplinary Working Group. *Circulation*, 117(13):1750–1767, 2008.

[18] P.C. Saunders, E.A. Grossi, R. Sharony, C.F. Schwartz, G.H. Ribakove, A.T. Culliford, J. Delianides, F.G. Baumann, A.C. Galloway, and S.B. Colvin. Minimally invasive technology for mitral valve surgery via left thoracotomy: Experience with forty cases. *Journal of Thoracic and Cardiovascular Surgery*, 127(4):1026–1032, 2004.

[19] D.D. Glower, K.P. Landolfo, F. Clements, N.P. Debruijn, M. Stafford-Smith, P.K. Smith, and F. Duhaylongsod. Mitral valve operation via port access versus median sternotomy. *European Journal of Cardio-Thoracic Surgery*, 14(1):143–147, 1998.

[20] L.H. Cohn, D.H. Adams, G.S. Couper, D.P. Bichell, D.M. Rosborough, S.P. Sears, and S.F. Aranki. Minimally invasive cardiac valve surgery improves

patient satisfaction while reducing costs of cardiac valve replacement and repair. *Annals of Surgery*, 226(4):421–428, 1997.

[21] P. Modi, A. Hassan, and W.R. Chitwood Jr. Minimally invasive mitral valve surgery: A systematic review and meta-analysis. *European Journal of Cardio-Thoracic Surgery*, 34(5):943–952, 2008.

[22] W.H. Ryan, J. Cheirif, T.M. Dewey, S.L. Prince, and M.J. Mack. Safety and efficacy of minimally invasive atrial septal defect closure. *Annals of Thoracic Surgery*, 75(5):1532–1534, 2003.

[23] R.M. Mentzer Jr., M.S. Jahania, and R.D. Lasley. Myocardial protection. In L.H. Cohn, editor, *Cardiac Surgery in the Adult, Third Ed.*, chapter 15, pages 443–464. McGraw-Hill, 2008.

[24] Hammon J.W. Extracorporeal circulation: Perfusion system. In Cohn L.H., editor, *Cardiac Surgery in the Adult*, chapter 12A, pages 350–370. McGraw-Hill, 2008.

[25] T.M. Dewey and M.J. Mack. Myocardial revascularization without cardiopulmonary bypass. In Cohn L.H., editor, *Cardiac Surgery in the Adult*, chapter 23, pages 633–654. McGraw-Hill, 2008.

[26] P.J. Lin, C. Chang, J. Chu, H. Liu, F. Tsai, P. Chu, C. Chiang, M. Yang, M. Shyr, and P. Tan. Video-assisted mitral valve operations. *Annals of Thoracic Surgery*, 61(6):1781–1786, 1996.

[27] W.R. Chitwood Jr., J.R. Elbeery, W.H.H. Chapman, J.M. Moran, R.L. Lust, W.A. Wooden, and D.H. Deaton. Video-assisted minimally invasive mitral valve surgery: The "micro-mitral" operation. *Journal of Thoracic and Cardiovascular Surgery*, 113(2):413–414, 1997.

[28] T.M. Peters, C.A. Linte, A.D. Wiles, N. Hill, J. Moore, C. Wedlake, D. Jones, D. Bainbridge, and G. Guiraudon. Development of an augmented reality approach for closed intracardiac interventions. In *2007 IEEE InternationalSymposium on Biomedical Imaging: From Nano to Macro*, pages 1004–1007. IEEE, 2007.

[29] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Proceedings of the SPIE Conference on Telemanipulator and Telepresence Technologies*, volume 2351, pages 282–292. Boston, Massachusetts, USA, 1995.

[30] Q. Zhang, R. Eagleson, and Peters T. Dynamic real-time 4D cardiac display using GPU-accelerated volume rendering. *The Journal of Computerized Medical Imaging and Graphics*, 5:555–556, 2009.

[31] D.F. Pace, A.D. Wiles, J. Moore, C. Wedlake, D.G. Gobbi, and T.M. Peters. Validation of four-dimensional ultrasound for targeting in minimally-invasive beating-heart surgery. In *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, volume 7261. SPIE, 2009.

[32] C.A. Linte, M. Wierzbicki, J.T. Moore, G.M. Guiraudon, S.H. Little, and T.M. Peters. Towards subject-specific models for the dynamic heart for image-guided mitral valve surgery. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Medical Image Computing and Computer Assisted Interventions*, volume 4792-2, pages 94–101. LNCS, 2007.

[33] C.L. Cheung, C. Wedlake, J. Moore, S.E. Pautler, A. Ahmad, and T.M. Peters. Fusion of stereoscopic video and laparoscopic ultrasound for minimally invasive partial nephrectomy. In *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, volume 7261. SPIE, 2009.

[34] F.J. Rybicki, T. Sheth, and F. Chen. Cardiac surgical imaging. In L.H. Cohn, editor, *Cardiac Surgery in the Adult, Third Ed.*, chapter 6, pages 179–198. McGraw-Hill, 2008.

[35] J.A. Fox, V. Formanek, A. Friedrich, and S.K. Shernan. Intraoperative echocardiography. In L.H. Cohn and L.H. Edmunds Jr., editors, *Cardiac Surgery in the Adult*, chapter 10, pages 283–314. McGraw-Hill, 2003.

[36] C.A. Linte, M. Wierzbicki, J. Moore, G. Guiraudon, D.L. Jones, and T.M. Peters. On enhancing planning and navigation of beating-heart mitral valve surgery using pre-operative cardiac models. In *IEEE Engineering and Medicine and Biology Society*, pages 475–478. IEEE, 2007.

[37] J.B. Hummel, M.R. Bax, M.L. Figl, Y. Kang, C. Maurer Jr., W.W. Birkfellner, H. Bergmann, and R. Shahidi. Design and application of an assessment protocol for electromagnetic tracking systems. *Journal of Medical Physics*, 32(7):2371–9, 2005.

[38] Northern Digital Inc. NDI Aurora Manual. Report, NDI International Headquarters, Waterloo, Canada, 2009.

[39] A. Goshtasby. *2-D and 3-D Image Registration*, pages 156–166. John Wiley & Sons, 2005.

[40] C.A. Linte, J. Moore, A.D. Wiles, Wedlake C., and T.M. Peters. Virtual reality-enhanced ultrasound guidance: A novel technique for intracardiac interventions. *Computer Aided Surgery*, 13(2):82–94, 2008.

[41] A. Wang, S.M. Mirsattari, A.G. Parrent, and T.M. Peters. Fusion of intraoperative cortical images with preoperative models for neurosurgical planning and guidance. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 7261. SPIE, 2009.

[42] A. Wang, S.M. Mirsattari, D. Gobbi, P. Das, Q. Zhang, and T.M. Peters. Interactive multimodality display environment with photographic overlay enhancement for epilepsy surgical planning. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 6918. SPIE, 2008.

[43] B. Wu, R.L. Klatzky, D. Shelton, and G.D. Stetten. Psychophysical evaluation of in-situ ultrasound visualization. *IEEE Transactions on Visualization and Computer Graphics*, 11(6):684–693, 2005.

[44] P.J. Edwards, L.G. Johnson, D.J. Hawkes, M.R. Fenlon, A.J. Strong, and M.J. Gleeson. Clinical experience and perception in stereo augmented reality surgical navigation. In *Medical Imaging and Augmented Reality*, pages 369–376. Springer Berlin/Heidelberg, 2004.

[45] W. Birkfellner, M. Figl, C. Matula, J. Hummel, R. Hanel, H. Imhof, F. Wanschitz, A. Wagner, F. Watzinger, and H. Bergmann. Computer-enhanced stereoscopic vision in a head-mounted operating binocular. *Physics in Medicine and Biology*, 48:N49–N57, 2003.

[46] F. Banovac, J. Bruno, J. Wright, and K. Cleary. Thoracoabdominal interventions. In T.M. Peters and K. Cleary, editors, *Image-Guided Interventions: Technology and Applications*, chapter 13, pages 387–407. Springer, 2008.

[47] A.D. Wiles, D.G. Thompson, and D.D. Frantz. Accuracy assessment and interpretation for optical tracking systems. In R.L. Galloway Jr., editor, *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, volume 5367, pages 421–432. SPIE, 2007.

[48] G.S. Guthart and J.K. Salisbury Jr. The Intuitive^{TM} telesurgery system: Overview and application. In *Proceesings of the 2000 IEEE Conference on Robotics and Automation*, volume 1, pages 618–621, 2000.

[49] V. Falk and F.W. Mohr. Minimally invasive myocardial revascularization. In Cohn L.H., editor, *Cardiac Surgery in the Adult*, chapter 26, pages 697–710. McGraw-Hill, 2008.

[50] R.E. Link, S.B. Bhayani, and L.R. Kavoussi. A prospective comparison of robotic and laparoscopic pyeloplasty. *Annals of Surgery*, 243(4):486–491, 2006.

[51] Z. Obrenovic, D. Starcevic, and E. Jovanov. Multimodal presentation of biomedical data. In M. Akay, editor, *Wiley Encyclopedia of Biomedical Engineering*. John Wiley & Sons, 2006.

[52] O. Bimber and R. Raskar. Modern approaches to augmented reality. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses*, pages 1–86. ACM, 2006.

[53] L.G. Johnson, Edwards P., and Hawkes D. Surface transparency makes stereo overlays unpredictable: The implications for augmented reality. In *Studies in Health Technology and Informatics*, volume 94, pages 131–6, 2003.

[54] M. Menozzi. Visual ergonomics of head-mounted displays. *Japanese Psychological Research*, 42(4):213–221, 2000.

[55] M.A. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *Computer Graphics*, 26:203–210, 1997.

[56] S.K. Maithel, L. Villegas, N. Stylopoulos, S. Dawson, and D.B. Jones. Simulated laparoscopy using a head-mounted display vs traditional video monitor: An assessment of performance and muscle fatigue. *Surgical Endoscopy*, 19:406–411, 2005.

[57] H. Visarius, J. Gong, C. Scheer, S. Haralamb, and L.P. Nolte. Man-machine interfaces in computer assisted surgery. *Computer Aided Surgery*, 2:102–7, 1997.

[58] L. Rossi, D. Sacerdoti, B. Billi, G. Lesnoni, M. Orciuolo, T. Rossi, D. Sacerdoti, and L. Bertollini. Automatic speech recognition in vitreo-retinal surgery: A project for a prototypal computer-based voice-controlled vitrectomy machine. *European Journal of Ophthamology*, 6(4):454–9, 1996.

[59] M.E. Allaf, S.V. Jackman, P.G. Schulam, J.A. Cadeddu, B.R. Lee, R.G. Moore, and L.R. Kavoussi. Laparoscopic visual field. *Surgical Endoscopy*, 12:1415–8, 1998.

[60] J. Wachs, H. Stern, Y. Edan, M. Gillam, C. Feied, M. Smith, and J. Handler. A real-time hand gesture interface for medical visualization applications. In A. Tiwari, J. Knowles, and E. Avineri, editors, *Applications of Soft Computing: Recent Trends*, chapter 4, pages 153–162. Springer, 2006.

[61] C. Grätzel, T. Fong, S. Grange, and C. Baur. A non-contact mouse for surgeon-computer interaction. *Technology and Health Care*, 12(3):245–257, 2004.

[62] K. Hinckley, R. Pausch, J.C. Goble, and N.F. Kassell. Passive real-world interface props for neurosurgical visualization. In *CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 452–458, 1994.

[63] A.G. Hauptmann. Speech and gestures for graphic image manipulation. *SIGCHI Bulletin*, 20(SI):241–245, 1989.

[64] D. Holmes III, M. Rettmann, and R. Robb. Visualization in image-guided interventions. In T.M. Peters and Cleary K., editors, *Image-Guided Interventions: Technology and Applications*, chapter 3, pages 45–80. Springer, 2008.

[65] W. Royce. Managing the development of large software systems. In *Proceedings of IEEE WESCON*, pages 1–9, 1970.

[66] B. Boehm. A spiral model of software development and enhancement. *IEEE Computer*, 21(5):61–72, 1988.

[67] D. Hix and H.R. Hartson. *Developing user interfaces: Ensuring usability through product and process*, page 416. New York, NY, John Wiley & Sons, 1993.

[68] J.L. Gabbard and J. E. Swan II. Usability engineering for augmented reality: Employing user-based studies to inform design. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):513–525, 2008.

[69] C.L. Mackenzie, J.A. Ibbotson, C.G.L. Cao, and A.J. Lomax. Hierarchical decomposition of laparoscopic surgery: A human factors approach to investigating the operating room environment. *Minimally Invasive Therapy and Allied Technologies*, 10(3):121–127, 2001.

[70] M. Good. Participatory design of a portable torque-feedback device. In *CHI '92: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 439–446, 1992.

[71] J.L. Blomberg and A. Henderson. Reflections on participatory design: Lessons from the trillium experience. In *CHI '90: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 353–360. ACM, 1990.

[72] K.J.M. Surry, H.J.B. Austin, A. Fenster, and T.M. Peters. Poly(vinyl alcohol) cryogel phantoms for use in ultrasound and MR imaging. *Physics in Medicine and Biology*, 49(24):5529–5546, 2004.

[73] D. Clark-Carter. *Doing Quantitative Psychological Research*, pages 257–258. Taylor & Francis - Psychology Press, 1997.

[74] S. Sawilowsky. Nonparametric tests of interaction in experimental design. *Review of Educational Research*, 60(1):91–126, 1990.

[75] P.H. Kvam and B. Vidakovic. *Nonparametric Statistics with Applications to Science and Engineering*, pages 126–129. John Wiley & Sons, 2007.

[76] R.G. Congalton and K. Green. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices (2nd Ed.)*, pages 41–55. CRC Press - Taylor & Francis Group, 2009.

[77] M.B. Bloom, A.D. Salzberg, and T.M. Krummel. Advanced technology in surgery. *Current Problems in Surgery*, 39(8):733–830, 2002.

[78] S. Grange, T.W. Fong, and C. Baur. M/ORIS: A medical/operating room interaction system. In *International Conference on Multimodal Interfaces (ICMI)*. ACM, October 2004.

[79] V. Falk, T. Walther, R. Autschbach, A. Diegeler, R. Battellini, and F.W. Mohr. Robot-assisted minimally invasive solo mitral valve operation. *Journal of Thoracic and Cardiovascular Surgery*, 115(2):470–471, 1998.

[80] K. Hinckley, R. Pausch, D. Proffitt, and N.F. Kassell. Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction*, 5(3):260–302, 1998.

[81] J.R. Cooperstock and G. Wang. Stereoscopic display technologies, interaction paradigms, and rendering approaches for neurosurgical visualization. In A.J. Woods, N.S. Holliman, and J.O. Merritt, editors, *Stereoscopic Displays and Applications XX*, volume 7237-1, pages 723703–723714. SPIE, 2009.

[82] M.R. Lehto and J.R. Buck. *Introduction to Human Factors and Ergonomics for Engineers*, pages 69–70. CRC Press - Taylor & Francis Group, 2007.

[83] V. Interrante, B. Ries, and L. Anderson. Distance perception in immersive virtual environments, revisited. In *VR '06: Proceedings of the IEEE Conference on Virtual Reality*, pages 3–10, 2006.

[84] Kaiser Electro-Optics Inc. Head Mounted Displays: Proview XL35 & XL50. Report, Kaiser Electro-Optics Inc. A Rockwell Collins Company, 2003.

[85] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behaviour*, 19(4):486–517, 1987.

[86] A. Cao, R.D. Ellis, E.D. Klein, G.W. Auner, M.D. Klein, and A.K. Pandya. Comparison of supplemental wide field of view versus a single field of view with zoom or performance in minimally invasive surgery. *Surgical Endoscopy*, 22:1445–1451, 2007.

[87] J.L. Gabbard, D. Hix, and J. E. Swan II. User-centered design and evaluation of virtual environments. *IEEE Computer Graphics and Applications*, 19(6):51–59, 1999.

[88] E.M. McDougall, J.J Soble, J.S.J. Wolf, S.Y. Nakada, O.M. Elashry, and R.V. Clayman. Comparison of three-dimensional and two-dimensional laparoscopic video systems. *Journal of Endourology*, 10:371–374, 1996.

[89] G.B. Hanna and A. Cuschieri. Influence of two-dimensional and three-dimensional imaging on endoscopic bowel suturing. *World Journal of Surgery*, 24(4):444–448, 2000.

[90] D.A. Bowman, S. Coquillart, B. Froehlich, M. Hirose, Y. Kitamura, K. Kiyokawa, and W. Stuerzlinger. 3D user interfaces: New directions and perspectives. *IEEE Computer Graphics and Applications*, 28(6):20–55, 2008.

[91] A.S. Yee. An investigation of the effect of display augmentation with instantaneous movement information on telemanipulation performance under visual-motor mismatch. MASc Thesis, University of Toronto, Toronto, 2008.

[92] N. Stanton. *Human Factors Methods: A Practical Guide for Engineering and Design*, pages 318–319. Ashgate Publishing, 2005.

[93] A. Crossan, S.A. Brewster, S. Reid, and D. Mellor. Multi-session VR medical

training - The HOPS simulator. In *In Proceedings of the 16th British HCI Group Annual Conference*, pages 213–226. BCS HCI, Springer, 2002.

[94] Human Performance Research Group. NASA task load index (TLX) v. 1.0 Computerized Version. Report, NASA Ames Research Center California, 2003.

[95] R.B Darlington. The Wilcoxon signed-ranks test: A table of Wilcoxon p's for N thru 50. http://comp9.psych.cornell.edu/Darlington/wilcoxon/wilcox51.htm, 1996.

[96] R.C. Houck, J.E. Cooke, and E.A. Gill. Live 3D echocardiography: A replacement for traditional 2D echocardiography? *American Journal of Roentgenology.*, 187(4):1092–1106, 2006.

[97] A. Fenster, D.B. Downey, and H.N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 26(2001):R67–R99, 2000.

[98] A. Kiruluta, E. Moshe, and P. Subbarayan. Predictive head movement tracking using a Kalman filter. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, 27(2):326–331, 1997.

[99] T. Grossman and R. Balakrishnan. Pointing at trivariate targets in 3D environments. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 447–454. ACM, 2004.

[100] W. Du. An investigation of the potential benefits of history trail displays for human spatial performance under different visual-motor mappings. MASc Thesis, University of Toronto, Toronto, 2002.

[101] A. Morton. An investigation of an augmented reality display of predictive and historical trajectory information for manual control under misaligned visual-motor mappings. MASc Thesis, University of Toronto, Toronto, 2004.