



University of Kentucky
UKnowledge

University of Kentucky Master's Theses

Graduate School

2006

INFORMATION TECHNOLOGY FOR NEXT-GENERATION OF SURGICAL ENVIRONMENTS

JESUS CABAN

University of Kentucky, jesuscaban@gmail.com

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

CABAN, JESUS, "INFORMATION TECHNOLOGY FOR NEXT-GENERATION OF SURGICAL ENVIRONMENTS" (2006). *University of Kentucky Master's Theses*. 229.
https://uknowledge.uky.edu/gradschool_theses/229

This Thesis is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Master's Theses by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

ABSTRACT OF THESIS

INFORMATION TECHNOLOGY FOR NEXT-GENERATION OF SURGICAL ENVIRONMENTS

Minimally invasive surgeries (MIS) are fundamentally constrained by image quality, access to the operative field, and the visualization environment on which the surgeon relies for real-time information. Although invasive access benefits the patient, it also leads to more challenging procedures, which require better skills and training. Endoscopic surgeries rely heavily on 2D interfaces, introducing additional challenges due to the loss of depth perception, the lack of 3-Dimensional imaging, and the reduction of degrees of freedom.

By using state-of-the-art technology within a distributed computational architecture, it is possible to incorporate multiple sensors, hybrid display devices, and 3D visualization algorithms within a flexible surgical environment. Such environments can assist the surgeon with valuable information that goes far beyond what is currently available. In this thesis, we will discuss how 3D visualization and reconstruction, stereo displays, high-resolution display devices, and tracking techniques are key elements in the next-generation of surgical environments.

KEYWORDS: Computer Vision, Visualization, High-Resolution Display Systems,
Minimally Invasive Surgery, Information Technology

INFORMATION TECHNOLOGY FOR NEXT-GENERATION
OF SURGICAL ENVIRONMENTS

By

JESUS CABAN

Director of Thesis

Director of Graduate Studies

Date

INFORMATION TECHNOLOGY FOR NEXT-GENERATION
OF SURGICAL ENVIRONMENTS

By

JESUS CABAN

Co-Director of Thesis

Co-Director of Thesis

Director of Graduate Studies

Date

RULES FOR THE USE OF THESIS

Unpublished theses submitted for the Masters degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this thesis for use by its patrons is expected to secure the signature of each user.

Name

Date

[illegible]

THESIS

JESUS CABAN

THE GRADUATE SCHOOL
UNIVERSITY OF KENTUCKY

2005

INFORMATION TECHNOLOGY FOR NEXT-GENERATION
OF SURGICAL ENVIRONMENTS

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
College of Engineering at the University of Kentucky

By

Jesus Caban

Lexington, Kentucky

Director: Dr. Brent Seales, Professor of Computer Science

Lexington, Kentucky

2005

INFORMATION TECHNOLOGY FOR NEXT-GENERATION
OF SURGICAL ENVIRONMENTS

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
College of Engineering at the University of Kentucky

By

Jesus Caban

Lexington, Kentucky

Dr. James Griffioen, Professor of Computer Science
and Dr. Jerzy Jaromczyk, Professor of Computer Science

Lexington, Kentucky

2005

MASTER'S THESIS RELEASE

I authorize the University of Kentucky
Libraries to reproduce this thesis in
whole or in part for purposes of research.

Signed: _____

Date: _____

ACKNOWLEDGMENTS

First, I would like to thank my family: William, Annie, William Jr. and Shandri Cabán for their support and encouragement to stay focused and concentrated in my work during my difficult and arduous times.

I would like to thank my advisor Brent Seales for his guidance, motivations and financial support. Also, I would like to thank the other members of my committee, James Griffioen and Jerzy Jaromczyk.

Many thanks to my friends and colleagues: Brennan Flora, Daniel Olson, Charles Carpenter, Jess Muettert, Enerlyn Lozada, George Landon, Jose Ortiz and Yun Lin. Without their help, support, and comfort I would have never made it.

Also, I would like to thank Adrian Park, Ivan George, and Rosemary Klein from the University of Maryland Medical Center for their support and for inspiring me to continue ahead with what I like and with my future.

Most of all, I would like to thank God for giving me the health, dedication, courage, and many wonderful blessings that made possible the completion of this long journey.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	iii
LIST OF FIGURES	v
CHAPTER 1 INTRODUCTION	1
1.1 Overview	1
1.2 Limitations	2
1.3 Proposed Architecture	5
1.3.1 Processing	6
1.3.2 Display	6
1.4 Thesis	7
1.5 Thesis Content	8
CHAPTER 2 PROCESSING	9
2.1 Introduction	9
2.2 Instrument Tracking	12
2.2.1 Previous Work	13
2.2.2 Endoscope Calibration	14
2.2.3 Reconstruction Method	19
2.2.4 Tracking Experiments	22
2.3 Image Enhancement	26
2.3.1 Distortion	27
2.3.2 Edges	28
CHAPTER 3 DISPLAY	30
3.1 Introduction	30
3.1.1 Flexible Display	33
3.2 Related Work	36
3.3 Implementation	39
3.3.1 Chromium	40
3.3.2 Geometric Calibration	41
3.3.3 Photometric Calibration	50
3.3.4 Improving the Calibration Accuracy	52
3.3.5 Color Correction	53
3.3.6 Implementation of Calibration in Chromium	53
3.4 Applications	54
3.4.1 Scalable, Adaptive Resolution	55

3.4.2	Flexibility	55
3.4.3	Medical Visualization	56
3.4.4	Layering	56
3.5	Software Distribution	57
CHAPTER 4 STUDIES AND RESULTS		59
4.1	Performance Analysis and Evaluation	59
4.2	Heterogeneous Display Systems	61
4.2.1	Multi-context Display System	61
4.2.2	Hybrid Display Systems	63
4.3	Distributed Video	64
4.3.1	Performance Analysis	65
4.4	Immersive Medical Visualization	66
4.5	Maryland Deployment	66
4.5.1	Maryland SimCenter	67
CHAPTER 5 CONCLUSION		68
REFERENCES		69
VITA		74

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Proposed architecture for the next-generation of surgical environments. Given medical imagery, scope video, CT images and other preoperative medical results, we can distribute such data and compute additional, valuable information to assist the surgeon.	5
2-1 (left) Current minimally invasive architecture where cameras are directly mapped to the display device. (right) A diagram of the processing part of the new-generation of surgical environments which tries to decouple the direct mapping camera-display by connecting the video devices and other medical images to a distributed processing environment which enable new ways to enhance images and compute valuable information.	10
2-2 Images used for the endoscope's calibration process	16
2-3 Visualization of the extrinsic results after the camera calibration step	17
2-4 (left) Scope's image before removing distortion. (right) Scope's image after removing distortion	18
2-5 Surgical stapler with identifiable marks and known size distance between them.	19
2-6 Diagram showing how we tracking surgical tools	20
2-7 A computer simulation of a surgical tool used to test and prove our algorithms for tracking and computing the 3D position and orientation of laparoscopic instruments	23
2-8 The setup to test our method for estimating the 3D orientation of surgical instruments	24
2-9 Tracking of a surgical stapler	26
2-10 (left) Distortion model of a 30 degrees Stryker endoscope (right) Distortion model of a Vista stereo endoscope	28
2-11 Computed edges of the surgical video	29

3-1	(left) Multiple views of medical visualization in a limited display space with limited resolution. (right) Multiple views of medical visualization in a multi-wall high-resolution display	34
3-2	(left)An SGI Onyx 2 supercomputer used for visualization and for active stereo visualization (right) Cluster of computers connected by a high-speed network creating an scalable distributed computing system by using commodity hardware	38
3-3	A set of casually aligned projectors used to create a seamless high-resolution display system	39
3-4	Diagram of our camera-based multi-projector calibration process . . .	42
3-5	Chessboard calibration patterns	44
3-6	Blobs calibration patterns	46
3-7	(left) By overlapping an image of the calibration grid, of a six-projectors display system, with a medical visualization in the same display we can see how the surface is approximated by the tessellation process (right) A zoom to a portion of the display where multiple projectors overlap	47
3-8	Seamless display system in a curved screen	48
3-9	(left)A high-resolution display system created by casually aligning 8 projectors in a curved screen. (center) A grid displayed through the display system to show the calibration results. (right) Medical visualization is possible after the calibration results	49
3-10	Calibration step of 8 casually-aligned projectors. By projecting equally spaced fiducials to each of the projectors that are part of the display we can generate the warping required to create a seamless images .	49
3-11	Alpha images used for the photometric calibration step	51
3-12	Photometric results. A set of casually aligned projectors can be dynamically calibrated, but without the correct photometric correction, the user will see brighter areas which can be distracting	51
3-13	Calibration results and errors caused by the camera lens distortion . .	52
3-14	SPU diagram	53
3-15	By using multiple projectors and our calibration technique, it is possible to create stereo display system that, by wearing the required glasses, the user can perceive depth	57
4-1	Plot of the performance analysis experiment	60

4-2	OpenGL software that creates a unified display that can receive multiple real-time video, tracking information, and other pre-operative imagery	62
4-3	A multi-context high-resolution display system used to display multiple images and videos in a single display space.	62
4-4	The hybrid display system	64
4-5	Distributed Video	65
4-6	High-resolution medical visualization	67
4-7	Maryland SimCenter	67

CHAPTER 1

INTRODUCTION

1.1 Overview

Minimally Invasive Surgery (MIS) is intentionally constrained in order to benefit the patient by minimizing the incision and the number of surgical cuts. Through small incisions or natural body openings, the surgeon performs the surgery or makes a diagnosis without the necessity of more dramatic and more traumatizing incision usually required during open surgeries.

A minimally invasive surgery is usually performed in the following way: a small incision is made near the umbilicus. Through that opening, a sheath, a scope containing a light and lens system are inserted. The camera (endoscope) sends continuous, real-time images to one or more video monitors. An insufflator pumps carbon dioxide into the abdomen cavity under automated pressure control to provide the space necessary to operate and to examine the abdominal contents. Secondary sheaths are inserted through incisions in the sides of the abdomen to allow the introduction of different surgical instruments used during the procedure. By looking at the video monitors and through the small incisions in the sides of the abdomen, the surgeon and his assistant can efficiently accomplish complex surgical tasks.

Advancements in video imaging, endoscope technology, and instrumentation have made it possible to convert a number of procedures, in many surgical specialties, from open surgeries to endoscopic ones. Rapid progress in imaging techniques and

advances in computer technology throughout the past decade have had a major effect on surgery and radiology, as well as a strong influence on related clinical fields.

The number of minimally invasive procedures has increased in the last several years, primarily because of the benefits that MIS provides. MIS has several advantages when compared to the usual open surgery. Some of the advantages are:

- less loss of blood
- lowered risk of postoperative complications and infections
- less pain, less strain on organs and less tissue trauma
- smaller surgical scars
- shorter hospital stay
- faster recovery

Although image-based information has always influenced treatments and therapies, the process has been revitalized by the improved quality and content of digital data and digital images. The use of computers, high-quality sensors, high-resolution display devices, computer vision, and 3D visualization techniques promises to facilitate complex endoscopic procedures by assisting the surgeons in the operating room. Also it promises to facilitate endoscopic procedures by enhancing digital images and digital data, and by improving the ability to learn new complex operations through virtual simulators and trainers.

1.2 Limitations

The limited space of an invasive surgery and the goal of reducing patients' trauma and injuries creates special challenges for endoscopic approaches. These and a number of other limitations make minimally invasive surgeries more challenging than the same procedure in open surgery. Here are some of the limitations and drawbacks of MIS.

- **Degree of Freedom**

The degrees of freedom available in open surgeries are lost in minimally invasive

procedures because of the limitation of space and the restricted range of motion of the instruments.

- **2D vs. 3D imagery**

3-dimensional imaging is lost on a 2-dimensional LCD screen. The single screen video monitor, which the surgeon uses as the image source, has no stereo information, no spatial information, and no concept of the depth of field. The surgeon can only estimate the depth of structures by moving the camera around or physically probing the structures.

- **Field of view**

The laparoscope has a small field of view. The surgeon must frequently adjust the camera position and orientation to regions of the operative field not visible through the scope. Such camera movements require skilled coordination between the surgeon and the surgeon's assistant.

- **Coordination**

The procedure requires significant hand-eye coordination. The laparoscopic camera does not generally face the direction in which the surgeon is facing. As a result, the instrument's on-screen movements will not match the surgeon's hand movements. For example, with a 90-degree rotation of the camera, an intended movement to the upper side of the image will result in a display movement to the left or right.

- **Multiple devices**

Image-guided surgeries still use a single display device to see the video and multiple independent display devices to see other important information. We believe that multiple independent and self-contained display devices can distract the surgeon and make it harder to have a complete, coherent, and cohesive idea about the status of the surgery at any time during the procedure. The

surgeon is responsible for joining preoperative images, different display devices and medical images with the current scope images and anatomy.

- **No tactile perception**

The loss of tactile sensation and perception of tissue qualities during minimally invasive surgery is a real drawback. Tactile perception of the anatomy is a key characteristic available in open surgeries that is not present during minimally invasive surgeries.

These limitations put more pressure on the surgeon and make any simple surgical intervention a challenging task. By using computer technology, we can find potential solutions to the current surgical limitations of minimally invasive surgeries.

Enhancement, undistortion, and smoothing of 2-dimensional images are some of the benefits that we can get by applying image processing to the laparoscopic imaging system. By applying computer vision and shadowing techniques, it is possible to create the illusion of 3-dimensions as shown by Tan et al [45]. Merging of non-visual imaging techniques, including 3-dimensional modeling and reconstruction of imaging data from computerized tomography, magnetic resonance imaging, and ultrasound, provide a real potential to merge different imaging modalities and different types of preoperative data, in the operating room, to better assist the surgeon. Similar to the above mentioned benefits that information technology can bring to the operating room, it is possible to increase the accuracy of the procedures and increase the surgeon's overall performance.

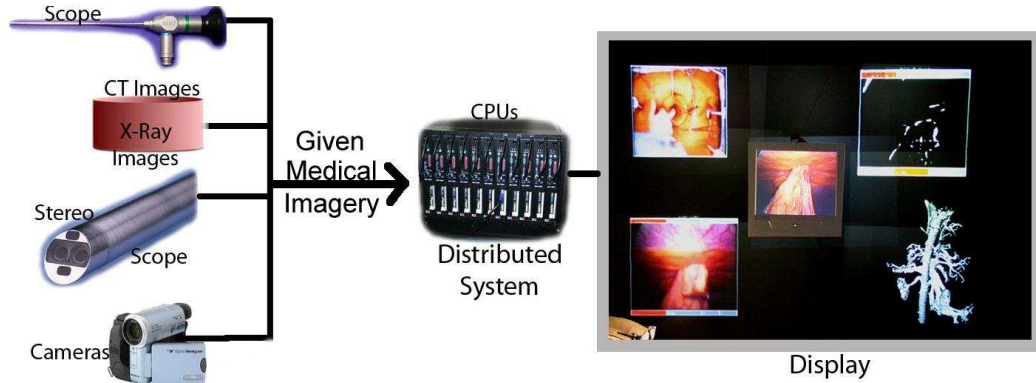


Figure 1–1: Proposed architecture for the next-generation of surgical environments. Given medical imagery, scope video, CT images and other preoperative medical results, we can distribute such data and compute additional, valuable information to assist the surgeon.

1.3 Proposed Architecture

With the integration of computers, information technology, and algorithms, some of the limitations and drawbacks of minimally invasive surgeries can be solved, and more information can be presented to the surgeon to assist him during surgical procedures.

In this thesis we propose a new, state-of-the-art architecture for the next-generation of surgical environments, which mainly consists of two modules and sections. With improvements, studies, and research in each of the pipeline modules, the integration and fusion of technology and computers within regular operating rooms will be easier and will be greatly beneficial for surgeons.

The two main areas of focus that will make the integration of state-of-the-art technology in operating rooms possible are: *Processing of Medical Data* and *Display*.

Figure 1–1 shows a diagram of the pipeline that will support and create the next-generation of surgical environments. Understanding the importance of each of the modules and putting together ideas about how computers, digital data, real-time data manipulation and processing, and display devices can assist surgeons in the operating

room is how we believe that the next-generation of surgical environments will become a reality.

Medical acquisition systems, sensors, and medical imaging devices have dramatically improved in speed, resolution, and accuracy in the last several years. At the same time, the raw data generated from each image-acquisition device has increased in size and is being generated in quantities that we have not seen before.

1.3.1 Processing

A parallel and distributed processing environment is proposed to enable a number of new ways to enhance, manipulate, and work with the medical data before sending it to the display system. Due to the increase in resolution, accuracy, and size of current medical data, a distributed system would prevent the processing module of the architecture from becoming a bottleneck in the next-generation of surgical environments' pipeline.

Currently, in minimally invasive surgeries, the direct mapping from the endoscope's video to the LCD display does not scale well and does not allow the integration of new algorithms to process the real-time video on-the-fly. The proposed distributed module of the next-generation of surgical environments is created with inexpensive and commodity hardware connected through the network. Such a computer cluster can be easily upgraded, is scalable, and has the capability to join or remove computers from the distributed environment as is needed. This way, the system is robust, affordable, and scalable.

1.3.2 Display

The continued popularity of 3D datasets and increases in resolution, scale, and complexity is now making the back-end display technology become an informational bottleneck of critical concern. As the acquisition systems improve and provide better resolution and accuracy, surgeons are still using the same LCD displays to visualize

the data and surgical video. In most cases current display devices down-sample the data to be able to show it through regular CRT or LCD monitors. Clearly, down-sampling the data is not the preferred option because of loss of detail. Surgeons and radiologists need to see and analyze details of the data being shown.

We propose a high-resolution display system within a distributed environment to show high-resolution 3D data sets and high-definition video. The display system is capable of creating a hybrid and heterogeneous presentation by showing multiple data sets, multi-context images, and multiple video images simultaneously to better assist the surgeon.

1.4 Thesis

The thesis of this work is that minimally invasive surgeries can be supported and improved by new architectures and algorithms for managing medical data and can be improved through advanced visualization and display techniques. To support this thesis, we have designed and tested an architecture that has demonstrated benefits.

In addition, we have designed, deployed and tested a display system that moves beyond what is normally and currently accepted for minimally invasive procedures. The results we present support the thesis that minimally invasive surgeries can be improved in a number of ways through the advances we present here.

In particular, we focus our research, studies and results in an architecture that can support the processing and enhancement of medical data, and a design that can display complex visualization models in a distributed and multi-modal way.

Specifically, some of the research and experiments presented in this thesis address the following problems, which have demonstrated benefits to the surgeon during minimally invasive procedures:

- Methods to track the surgical tools in 3D space in order to know the orientation, position, and location of the tip of the instrument.

- Procedures to compute 3D points of parts of the surgical field. Such points can be used to better assist the surgeon and give them information about the distance between instrument and anatomy.
- Fusion and integration of multiple, independent, and self-contained display devices.
- High-resolution display systems to enhance medical visualization and surgical videos.
- Stereo endoscopes and ways to capture, process and display stereo video.
- Computer-based performance analysis for surgical trainings.

1.5 Thesis Content

The content of this thesis is distributed the following way:

- **Chapter 2 (Processing):** We show different algorithms and techniques that can be applied to the next-generation of surgical environments pipeline to better assist the surgeon. By using image processing algorithms, advanced computer vision and realistic computer graphics, it is possible to help the surgeon during the understating and interpretation step of medical data.
- **Chapter 3 (Display):** We present our research in a casually aligned, auto-calibrated, multi-projector, high-resolution display system and its uses for medical visualization. The display system is used to better show, integrate, fuse and display medical imaging for surgeons.
- **Chapter 4 (Results):** We present our results in each of the parts of the pipeline. The results include tracking outcome, medical visualization improvement with high-resolution display system, distributed scope video analysis, and the deployment of the different components in an operating room.
- **Chapter 5 (Conclusion):** We summarize our work and conclude how information technology can be used in surgical environments and the benefits that the proposed architecture bring to minimally invasive surgeries.

CHAPTER 2

PROCESSING

2.1 Introduction

The processing environment is the joint and the glue that connects, makes compatible, and reconciles data from different sensors and different image modalities as well as prepares that data to be rendered and displayed. That fusion involves incorporating sensor data, 2D and 3D data collected prior to the procedures, and on-the-fly computed 2D and 3D data with the real-time video from the operation. The modeling and processing modules must transform, change, and process the raw data obtained from the devices or databases into a digital form that can be presented via the display interface and display system to the viewer.

The processing portion of the proposed architecture is where we apply computational algorithms, image processing techniques, distributed processing, and advanced computer vision algorithms to enhance the images and video. In addition, it is where we compute extra information from the real-time data and the pre-operative data, and where we compute and estimate the 3D position of the surgical tools. All of this extra valuable information can assist the surgeon during minimally invasive procedures.

There is little need for a modeling and processing layer if we are only interested in displaying real-time video and if the resolution of the acquired data from sensors exactly matches the display's resolution. The current laparoscopic environment has

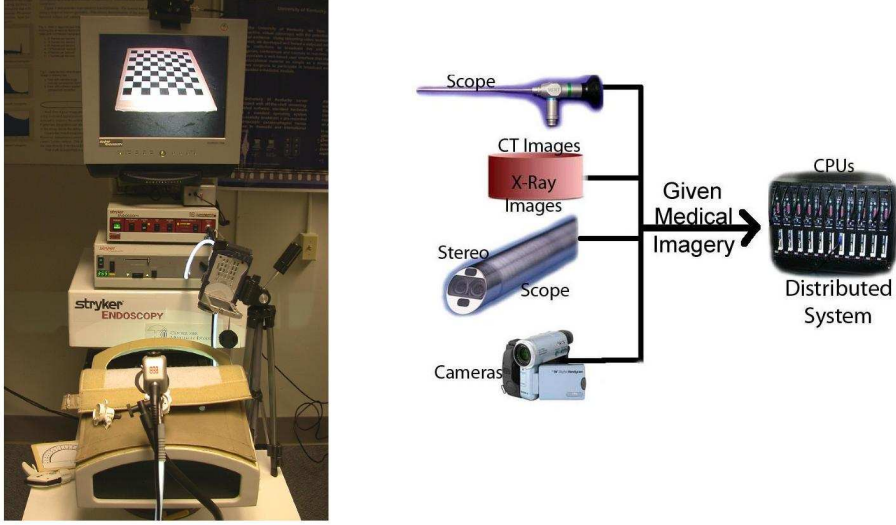


Figure 2-1: (left) Current minimally invasive architecture where cameras are directly mapped to the display device. (right) A diagram of the processing part of the new-generation of surgical environments which tries to decouple the direct mapping camera-display by connecting the video devices and other medical images to a distributed processing environment which enable new ways to enhance images and compute valuable information.

been engineered so that the camera acquires an image sequence that is mapped directly onto the display device as shown in Figure 2-1 (left). Flexibility and power can be gained by separating and decoupling this connection and inserting an intermediate distributed processing environment, which enables a number of new ways to enhance, manipulate and work with the data before sending it to the display system. Such flexibility to manipulate and transform the data is impossible in architectures with sensor-to-display direct mappings. Figure 2-1 (right) shows the intermediate step we propose.

Since the early days of simulation technology, a number of strategies, approaches, and techniques have been proposed to facilitate the development of distributed systems to support simulations. Surgical simulation environments have mirrored this trend. Systems and software like the *TGS's amiraVR.Cluster*[\[63\]](#) use state-of-the-art techniques to create 3D medical visualization using the power of a number of

distributed computers. The cluster produces real-time visualization and simulation data that increases the comprehensive analysis and study of 3D data.

Some features of our proposed intermediate distributed processing enable new 2D and 3D real-time data and bring new levels of cost-effective performance. The distributed environment is created with inexpensive and commodity hardware, which can easily be upgraded. In addition the distributed environment is scalable, making it easy to add and remove computers from it.

In the case of medical simulation and more specifically laparoscopy, the stated goal is to build a system that can reconcile imagery from the scope with preoperative imagery captured by CT scans and MRI as well as any other method that might provide some assistance during the laparoscopic procedure. Also, we use the imagery from the scope to estimate the 3D position and orientation of the surgical instruments.

The integration of advanced algorithms and visualization techniques can enable a number of new methods to benefit the surgeon during MIS procedures. For example, comparison or recognition of anatomy with respect to a large database of procedures might be helpful. Some other options include: a 3D volume from a CT scan showing the orientation and position of the scope with respect to the anatomy, a 3D geometry of the surgical work field, and an overlay image on top of the real-time video. It is clear that these goals cannot be accomplished without a processing module and a processing step.

Parallel distributed computing environments may be the answer, and it may be the case that surgical simulation and, eventually, operating rooms will become driven by massively parallel clusters of computers designed specifically to manage distributed sensors and preoperative, potentially collaborative patient databases. We propose that the next-generation of surgical environments should include a cluster of computers capable of running medical 3D reconstruction, overlaying information on top of real-time video, and introducing the capability of tracking the instrument

location and orientation at any time during the procedure. With such a distributed system and flexible display, it is possible to bring more technology to the operating room to assist surgeons during minimally invasive procedures.

The processing part of the proposed next-generation of surgical environments is scalable and designed with the flexibility that any algorithm that generates valuable information for the surgeon can be included and used. It is our purpose to show that minimally invasive surgeries are improved in a number of ways by the proposed architecture. In the following sections, we will discuss how 3D tracking and video enhancement can be applied and used for the next-generation of surgical environments.

2.2 Instrument Tracking

Instrument tracking is an important element currently missing in laparoscopic surgeries. If we can track the position of the instrument at any time during the procedure, we can better control the surgeon’s field of view by manipulating the camera automatically. We can provide the surgeon with information regarding distance between the instrument’s tip and the patient’s anatomy, distance between the instrument and the scope, and information about the 3D orientation of the surgical tool.

This section presents a method for estimating complete 3D information about scope and instrument positioning from monocular imagery. These measurements can be used as the basis for deriving and presenting additional cues during procedures. In this work, we present a method to extract and acquire explicit information that is implicitly confounded in the imagery. Such information, though valuable as a direct cue, is usually subtle, especially in monocular imagery. Extracting an explicit representation can provide a ready cue or an analytical tool that otherwise would remain subtle and far less useful. In particular, we concentrate on the problem of recovering the 3D position and orientation of instruments within the endoscope’s view

as well as the distance of these instruments from the scope, from each other, and from the anatomy.

Providing 3D information is crucial in addressing one of the primary technical and visual obstacles in conducting MIS procedures, which is the lack of an explicit depth cue. Experts become very good at understanding 3D relationships from monocular imagery, which does not make depth explicit but, rather contains a number of subtle depth cues, such as perspective distortion and scale, expert knowledge of instrument size, shape and relative positioning, and narrow depth of field, providing a focus cue.

We believe that the ability to extract precise depth measurements, including the position and orientation of instruments, scopes, and anatomy, can substantially enhance laparoscopic environments of the future. In particular, we envision two immediate uses when depth information can be made explicit for tracked instruments and anatomy: *enhanced visualization for the surgical team* and *objective performance measures, given video of training and simulation cases*.

2.2.1 Previous Work

Due to the minimally invasive nature and the small incisions of laparoscopic surgeries, a visual tracking is the most logistical approach to estimate the location and orientation of a surgical instrument.

In the last few years, tracking of the surgical tools has received some attention, but is still an open problem and active research area due to the difficulties of visual-tracking algorithms applied to minimally invasive surgery scenarios and the necessity of a truly flexible method to do tracking.

Some of the work that can be applied to the instrument tracking in minimally invasive surgeries are Kim's[6] studies to track the instrument by analyzing the distribution and condensation of colors, and Casals'[7] work that suggests to do tracking of medical instruments by shape and edges.

Few researchers have approached the problem of tracking surgical instruments in laparoscopic images from monocular images. Zhang[4] proposed a way to track by using markers at the instrument, Wang[5] proposed a way to track surgical tools by using a statistical color classification approach, and Wei[8] proposed a way to track by analyzing both markers and colors.

Numerous assumptions exist in each of the previous work. Some assumptions are a static camera, that the instrument is always visible and present at the field-of-view, and that is possible to mark surgical instruments just before the procedure.

Due to the diversity of the instruments used during a minimally invasive surgery and the number of instruments that are used only one time, it is not possible to always mark the instruments before the procedure to accurately track the instrument. Our approach is similar to Wei’s approach, but using computer vision tracking techniques to track specific features of the instruments from where we can compute the orientation and 3D position of the instrument.[9] One of the unique features of our approach is that we do not assume a static camera. The surgeon can freely move the endoscope and the surgical tool without getting any uncertainty in the 3D estimation results. The estimation of the 3D position of a surgical tool is not affected by the position of the instrument with respect to the endoscope. We accomplish that flexibility and robustness by taking advantages of the camera model and the intrinsic parameters of the endoscope.

2.2.2 Endoscope Calibration

In order to formally model the geometry of the endoscope, we assume that the imaging system can be modeled as a pinhole camera (i.e. perspective projection). Using this camera model, we apply computer vision methods and algorithms in order to calculate its characteristic, geometry and distortion parameters.

We use the pinhole camera model to compute the endoscope parameters and characteristics through a calibration process. According to the pinhole model, the

relationship between a 3D-point X and its corresponding 2D-point x at the image plane is given by

$$x = PX \quad (2.1)$$

where $P = KR[I] - C$. K are the intrinsic parameters of the camera, R is a 3x3 rotation matrix representing the orientation of the camera coordinate frame, and $[I] - C$ represents a matrix divided up to into a 3x3 block (identity matrix) plus a column vector, the coordinates of the camera center.

Intrinsic Parameters

The intrinsic parameters are the coefficients needed to link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame.

The intrinsic camera parameters are:

- focal length of the camera
- aspect ratio
- principal point or image center
- radial distortion coefficient

The intrinsics of a pinhole camera can be defined with the following matrix where f is the focal length of the camera, m_x and m_y are pixel size in x and y directions, s is the skew angle, and p_x and p_y are the coordinates of the principal point in x and y directions.

$$K = \begin{bmatrix} fm_x & s & m_x p_x \\ 0 & fm_y & m_y p_y \\ 0 & 0 & 1 \end{bmatrix}$$

Extrinsic Parameters

The extrinsic parameters define the location and orientation of the camera reference frame with respect to a known world reference frame.

- a 3 x 3 rotation matrix (r_{ij})

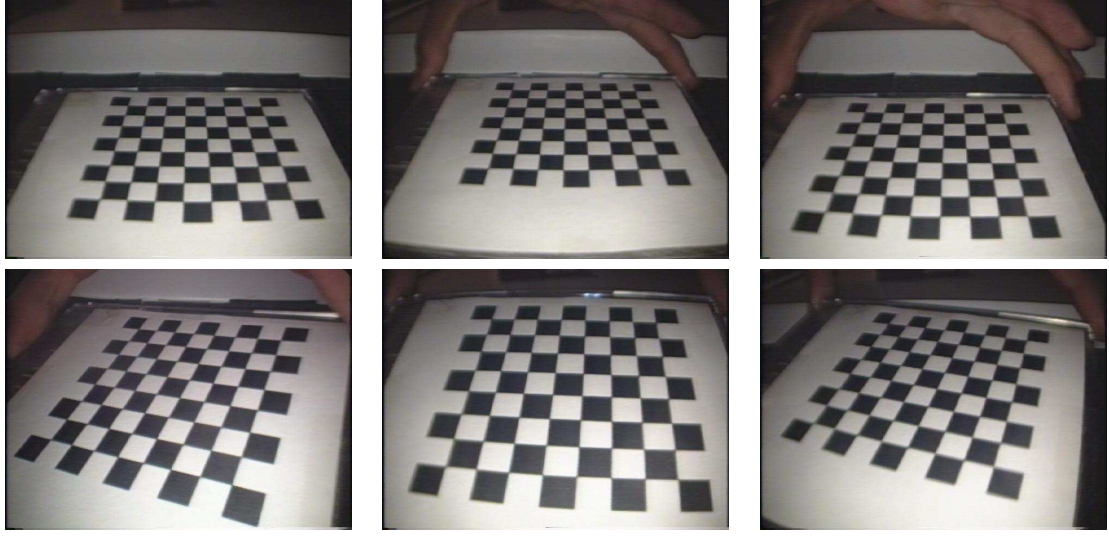


Figure 2-2: Images used for the endoscope's calibration process

- a 3D translation vector (T)

The extrinsic parameters can be defined with the following matrix:

$$M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & -R_1^T T \\ r_{21} & r_{22} & r_{23} & -R_1^T T \\ r_{31} & r_{32} & r_{33} & -R_1^T T \end{bmatrix}$$

Calibration Technique

For the calibration step, we use a two-stage calibration technique introduced by Tsai[10] and implemented in Matlab software[11]. For the calibration process, we captured a number of images of a calibration target through the endoscope, then we run the *Camera Calibration Toolbox for Matlab*[12], from which we obtain the parameters required to know the camera properties.

The used calibration pattern was a four inches 9x9 chessboard printed in high-quality paper. The corner points of the squares were detected with subpixel accuracy and treated as calibration points from where we obtained the camera geometry we need.

Figure 2-2 shows multiple images of the calibration target taken from a 30° monocular endoscope. The images were captured from the endoscope at a resolution

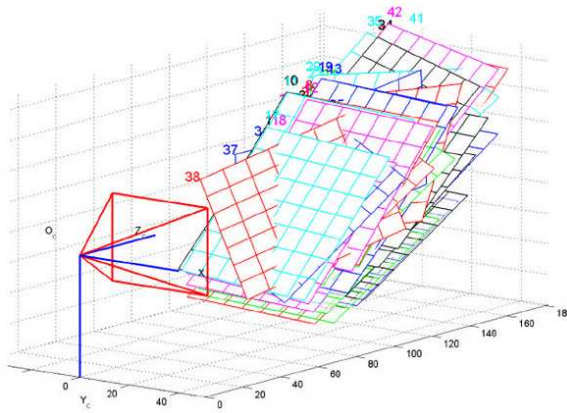


Figure 2–3: Visualization of the extrinsic results after the camera calibration step of 720x480. The different positions of the calibration target were realized by moving and rotating the calibration target to different positions inside the field-of-view of the endoscope. Those images were used as calibration images to obtain the intrinsic and extrinsic parameters of the endoscope.

Figure 2–3 shows the visualization of the extrinsic parameters that *Matlab* generates allowing us to check the calibration results. From the extrinsic representation, we can see the position of the pinhole camera and the 3D position of the chessboard with respect to the camera origin in each of the images we used for the calibration process.

After the calibration process, we have the coefficients and matrices from which we can compute the fundamental matrix which allows us to compute the 3D position and orientation of a surgical tool.

Image Distortion

As a result of the radial curvature of camera lens elements, there is distortion in the images. There is no real lens system that can produce perfect pinhole images. In the case of endoscopes, different viewing angle scopes and scopes with wide-angle

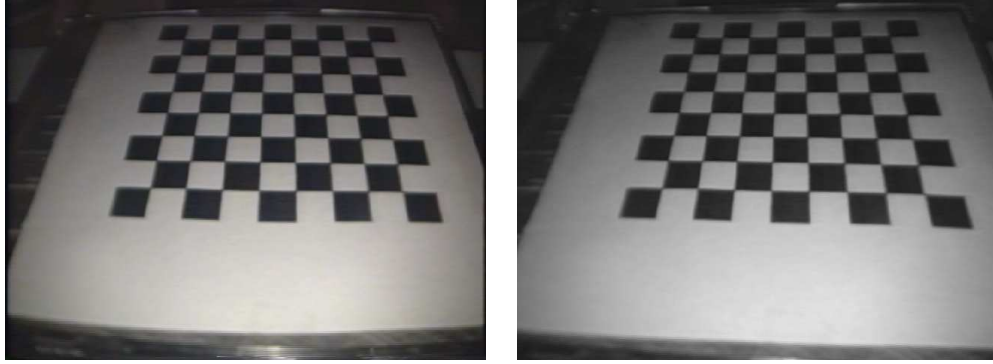


Figure 2-4: (left) Scope's image before removing distortion. (right) Scope's image after removing distortion

lenses, which enlarge the field of view, cause significant distortion. The lens distortions can be removed by calculating the distortion parameters through optical calibration.

Radial distortion is modeled as:

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = L(\tilde{r}) \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}$$

where

- (\tilde{x}, \tilde{y}) is the ideal image position
- (x_d, y_d) is the actual image position, after radial distortion
- \tilde{r} is the radial distance $\sqrt{\tilde{x}^2 + \tilde{y}^2}$ from the center for radial distortion
- $L(\tilde{r})$ is a distortion factor, which is a function of the radius \tilde{r} only

After a camera has been calibrated, and its distortion factors are computed. It is possible to use the camera parameters to resample any image taken by that camera so that its lens distortion is removed from the image.

Figure 2-4 shows lens distortion in an image captured from a 30° endoscope compared with a distortion-free image generated after the calibration process. Clearly we can see from the distorted image that a line that should look like a straight line, has a lot of curvature. After computing the distortion factor we can see how the bottom of the image looks more like a straight line.

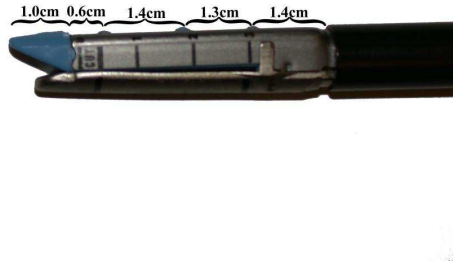


Figure 2–5: Surgical stapler with identifiable marks and known size distance between them.

2.2.3 Reconstruction Method

The key assumption that enables depth reconstruction of instruments visible in monocular sequences is knowledge of the shape, size and the metric measurements of visually identifiable fiducials or marks on the instrument. We can select specific features of each instrument as features of the surgical tool to be tracked, or use the technique mentioned before of marking the instrument with features easier to track.

Based on shape and size information, it is possible to track features at the imagery and recover the 3D position of each tracked point. From these points, with a priori information about the instrument, it is possible to compute the 3D position and orientation of the tip of the instrument. Figure 2–5 shows a stapler instrument with identifiable marks and known distances between each of the points.

Usually, the shape of MIS instruments are almost linear so that the instrument can be smoothly inserted into ports and manipulated through small incisions. We exploit this fact and as a result simplify the problem of tracking and estimating 3D points that lie on the instrument. Figure 2–6 shows a diagram of our method that by tracking at least three points that lie in the same plane of the surgical instrument, we can estimate the instrument’s position and orientation.

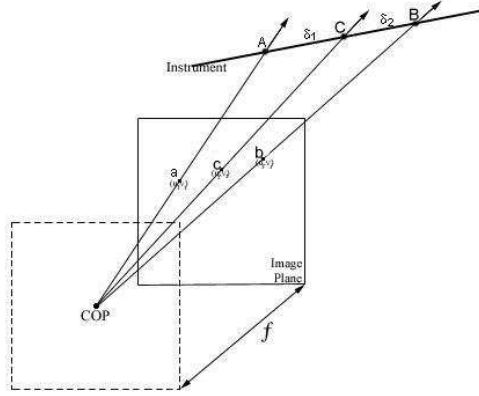


Figure 2-6: Diagram showing how we tracking surgical tools

From figure 2-6 we know the position of C with respect to A and B.

$$C = \delta_A A + \delta_B B \quad (2.2)$$

where δ_A and δ_B are known.

Also we know that the distance from A to B

$$\| B - A \| = \delta_A + \delta_B \quad (2.3)$$

By using the projection a, b and c of the 3D space points A, B and C at the image plane, we can compute the 3D position of each independent point and then estimate the orientation of the medical instrument in 3D space.

Because we exploit the fact that MIS instruments are usually linear, we simplify the problem to track points in a line, which is a 1D object. That is, by selection $R=I$ in equation 2.1, we can compute the depth of three unknown points A , B , C by:

$$A = z_A K^{-1} a$$

$$B = z_B K^{-1} b$$

$$C = z_C K^{-1} c$$

where z_A ; z_B and z_C are the unknown depths of the A , B , C points.

Using equation 2.2 we have that

$$z_C c = z_A \delta_A a + z_B \delta_B b \quad (2.4)$$

after eliminating K^{-1} from both sides. After a cross-product on both sides of the above equation, with c , we have

$$z_A \delta_A (a \times c) + z_B \delta_B (b \times c) = 0 \quad (2.5)$$

$$z_B = -z_A \frac{\delta_A (a \times c)(b \times c)}{\delta_B (b \times c)(b \times c)} \quad (2.6)$$

From equation 2.3 we have that

$$\| K^{-1}(z_B b - z_A a) \| = \delta_A + \delta_B \quad (2.7)$$

from where we can compute and conclude that

$$z_A = \frac{\delta_A + \delta_B}{K^{-1}\left(\frac{\delta_A (a \times c)(b \times c)}{\delta_B (b \times c)(b \times c)} b + a\right)} \quad (2.8)$$

$$z_B = -z_A \frac{\delta_A (a \times c)(b \times c)}{\delta_B (b \times c)(b \times c)} \quad (2.9)$$

$$C = [x_C, y_C, z_C]^T = \delta_A A + \delta_B B \quad (2.10)$$

From these equations and constraints we can reach the solution for unknown depth for each of the points as shown by Zhang[13]. This is, by just using the instrument coordinates as appears in the imagery, we can find the 3D position of each of the points. These depth values are derived based on the assumption that the instruments are linear, the camera is calibrated (i.e., the projection matrix is available from the off-line calibration process), and the distances between points on the instrument are known a priori.

2.2.4 Tracking Experiments

After observing the theory behind endoscope calibration and the way to reconstruct 3D-points from a monocular image, we implemented two different softwares to prove that we can use our tracking technique to track surgical instruments during laparoscopic procedures.

First, we developed a computer simulated medical tool in C and OpenGL. Then we implemented a program in Matlab that given a number of consecutive frames (i.e. video), the instrument is tracked, the 3D-point of the 2D features is found at the image plane, and the 3D coordinates of the tracked points are output. In addition, the 3D coordinates of the tip of the instrument are output. By using this information, we compute the 3D orientation of the instrument.

Computer Simulation

We created a simulated surgical tool by using OpenGL and the C programming language to show, test, and prove the formulas, algorithms and methods behind the tracking and 3D estimation of laparoscopic instrument. With the computer simulation we proved that it is possible to estimate the position of 3D points from monocular images by applying the formulas described above.

Figure 2-7 shows an OpenGL program displaying a simulated calibration pattern and a simulated surgical tool with three points to be tracked. Our study with synthetic data was done the following way. We first displayed the simulated surgical tool in the OpenGL window, then we captured the output of the simulation as an image. That image was the input to a Matlab software that by tracking the three points and by using the corresponding camera parameters, computed the 3D point for each of the three points of the line. After computing the depth of each of the three points, the software estimated the orientation and position of the instrument. After the 3D estimation process, we compared the 3D position of each of the points

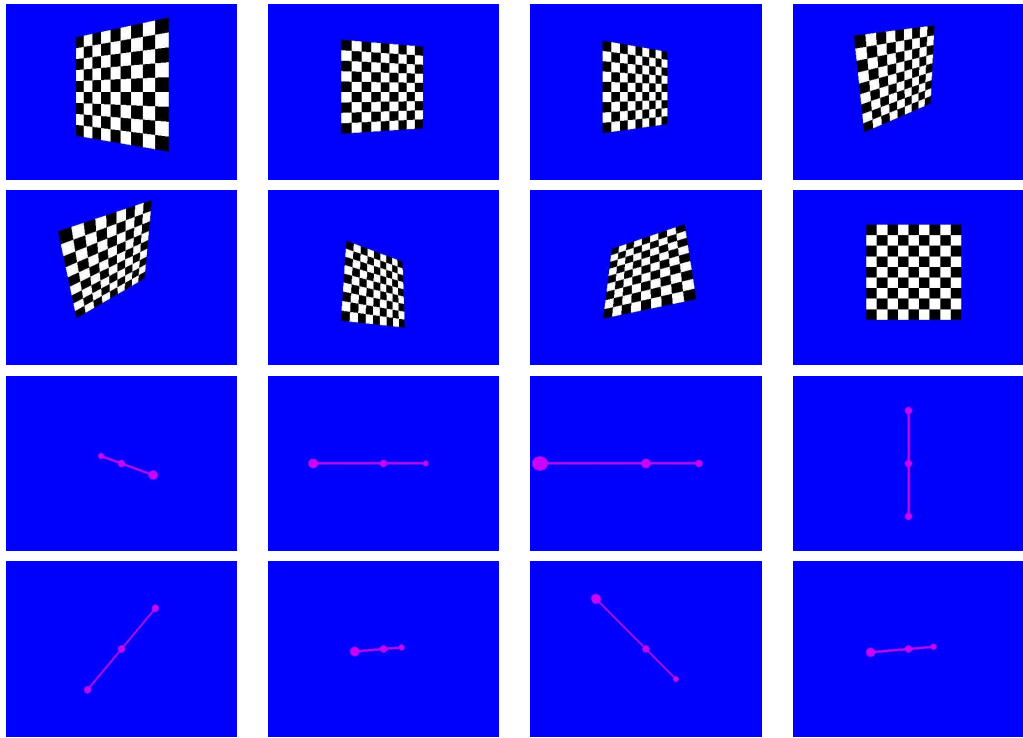


Figure 2-7: A computer simulation of a surgical tool used to test and prove our algorithms for tracking and computing the 3D position and orientation of laparoscopic instruments



Figure 2–8: The setup to test our method for estimating the 3D orientation of surgical instruments

from the Matlab output with the OpenGL coordinates where the points were drawn to calculate the accuracy of the system.

After running the simulation and after an analysis and comparison of the estimated 3D position versus the actual 3D location, we concluded that the accuracy of the technique depends on the precision of locating and tracking the features of the surgical tool. On average, the estimated 3D points were 0.9 units off from the actual 3D point, which corresponds to a 0.43% error. The main cause of this estimated error was the accuracy in detecting the centroid of the points. Small errors in computing the centroid of each of the tracked features results in some error in estimating the instrument location and 3D orientation.

Real Experiments

After showing that the formulas and algorithms worked in a computer simulated program, we decided to make experiments with real laparoscopes and surgical tools. We calibrated various endoscopes (e.g., 0° and 30° lenses). After the calibration

step, we used recovered lens distortion estimates to remove lens distortion from the images. By using the stapler presented in figure 2–5 with identifiable marks and known size distance between them, we captured several frames from the calibrated endoscope. Then by using the distortion-corrected images and our Matlab software we tracked the shaft of a stapler instrument in order to recover estimates of the 3D coordinates of points on the instrument. To have a rough estimate of the orientation of the instrument we had a protractor and a second camera perpendicular to the protractor. Figure 2–8 shows our setup while doing the experiments.

By acquiring images from the endoscope and images from the second external camera we were able to compute the 3D orientation of the instrument by using the calibrated endoscope. Then we were able to compare the orientation results with the images from the second camera with the protractor in the background.

Our experiments and results show that it is possible to track surgical instruments and estimate their 3D position and orientation, but a more accurate experiment has to be done to estimate the error in our 3D tracking technique. Zhang [4] used a mechanical instrument called the pcBird [57] that allowed him to put the instrument in a specific position and orientation, then compare that position and orientation from the vision-based tracking algorithm with the position and orientation returned by the sensors of the pcBird instrument.

Figure 2–9 shows some of the experiments we did to track a laparoscopic stapler. For each of the images the program detected the center of the tracked features, created a line between them and by using that line and the (u,v) position of each feature at the image plane, we were able to estimate the 3D orientation of the surgical tool.

We have found these methods to be very promising as a way to recover 3D cues from monocular data. We learned that:

- It is possible to calibrate an endoscope and model its properties as a pinhole camera.

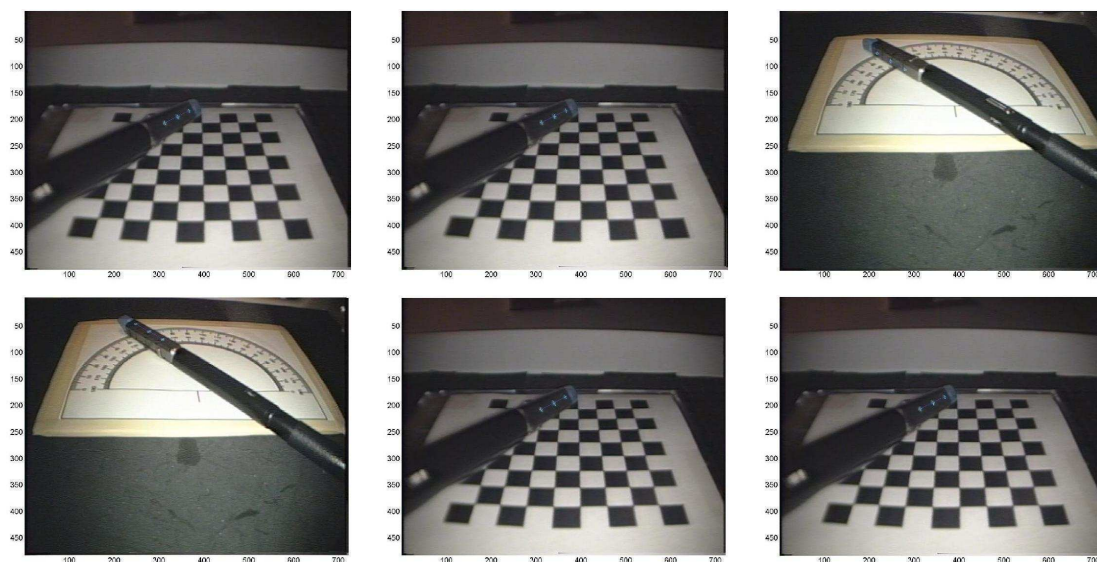


Figure 2-9: Tracking of a surgical stapler

- Due to the different lenses used during minimally invasive procedures, in order to use a calibrated endoscope in an operating room, we need to calibrate any possible lenses that might be used during the procedure before the actual surgery.
- Due to the small field-of-view and the wide-angle used by laparoscope cameras, the radial distortion is significant. Sometimes 25 or 30 pixels off. An accurate algorithm to remove radial distortion is needed to be able to better estimate the 3D orientation of the instrument.
- 3D tracking and 3D estimation are possible for surgical instruments if we know the shape, distances, and can find visible features to track.
- The key element for an accurate estimation of the instrument is to have an accurate tracking algorithm.

2.3 Image Enhancement

Minimally invasive surgeries depend in the visual imagery that the surgeon can see through the LCD monitor. Due to hardware limitations at the CCD level of the camera (endoscope) or due to optics problems with the lenses, the image shown

to the doctor is usually not the best image we can display. We can easily improve and enhance the image quality and distortion effect of laparoscopic video to provide the surgeon with a better image that might help him to better understand of the operative field.

For example, by just applying some image processing techniques to the parallel and distributed environment, we can remove the distortion, enhance the colors, and compute extra information that might assist the surgeon. We believe that image enhancement is a key element of the next-generation of surgical environments, and it can result in several advantages for the surgeon, as well as advantages for the surgeon's potential to understand the surgical scene.

2.3.1 Distortion

It has been noted that the images obtained from an endoscope shows severe radial distortion and barrel-type spatial distortion due to wide-angle configuration of the camera lens.[15]

Barrel distortion introduces nonlinear changes in the image, causing image areas near the distortion center to be compressed less, while areas farther from the center to be compressed more. Because of this, the areas near the edge of the image look significantly smaller than their actual size. This inhomogeneous image compression introduces significant errors in the results obtained during feature extraction and 3D tracking. Unless the lens distortion is corrected, the estimation errors could be very large.

Several researchers have presented various mathematical models of the image distortion and techniques to find the parameters to complete the distortion-correction procedure.[15, 16, 17] By applying any of these distortion correction methods, we can compute, enhance and correct the imagery presented to the surgeon on-the-fly.

In our experiments we used Brown's model[14] to compute the distortion coefficients and its Matlab implementation to undistort images. Figure 2-10 shows a

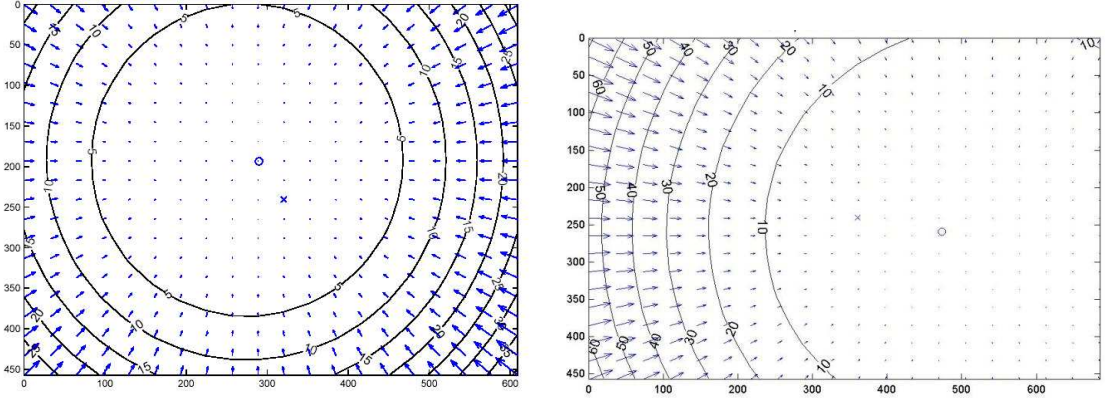


Figure 2-10: (left) Distortion model of a 30 degrees Stryker endoscope (right) Distortion model of a Vista stereo endoscope

distortion model for two endoscopes we have calibrated. The left image is the distortion model for a Stryker monocular endoscope. The image to the right of figure 2-10 is the distortion model of a Vista stereo endoscope. In each of the distortion models we can see that there is only a small portion of the image that is distortion-free. The majority of the image, particularly near the edges of the images, have distortion which can be as big as 20 or 30 pixels off.

2.3.2 Edges

We believe that both enhancing video image and extracting valuable information – such as estimated 3D surgical tool position – are of great benefit for the surgeon. To this end, we have also tested an edge detecting algorithm to further support our proposed architecture for the next-generation of surgical environments. By applying the Canny edge detection algorithms, it is possible to find edges in the image that can be presented to the surgeon in an independent window or even inserted as an overlay image on top of the surgical video to better assist the surgeon.

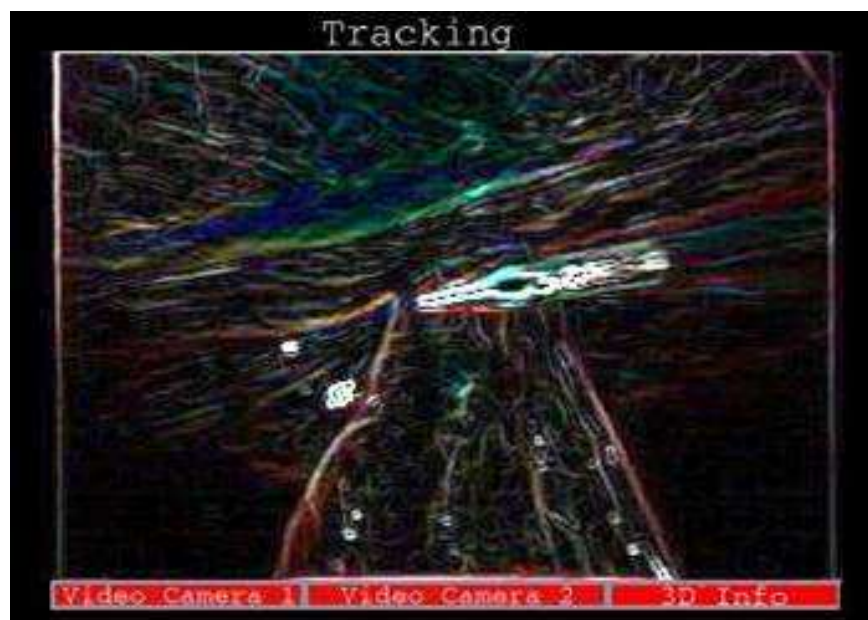


Figure 2-11: Computed edges of the surgical video

CHAPTER 3

DISPLAY

3.1 Introduction

The resolution, degree of complexity, and visual capabilities that surgeons desire in order to meaningfully explore, study, and analyze complex datasets and video pose an important technical challenge to researchers developing visualization technologies. Although new visualization methods can be used within a traditional desktop environment and even in some operating rooms, advances in multiple-view medical visualization and simulations, in conjunction with the continued increases in resolution, scale, and complexity of datasets, themselves, is now making the back-end display technology into a crucial informational bottleneck. The same visualization bottleneck surgeons experience in the operating room and the visualization equipment provided during minimally invasive surgeries.

This work focuses on how to build and deploy scalable and flexible display systems for medical applications. We believe that flexible display systems are a key element of the next-generation of surgical environments and they can greatly assist surgeons during laparoscopic procedures. These systems enable and encourage the development of visualization strategies that exploit high resolution, multiple data, stereo cues, adaptive and non-homogeneous display resolutions, and rapid display configurability. With this enabling technology it becomes possible to more readily match the functional capabilities of the end-display to the requirements of the data

and to the visualization strategy that most naturally supports the kind of analysis surgeons need and desire.

To make a flexible display system for medial visualization and a possible display system to assist surgeons during minimally invasive surgeries, we need to cover and solve the main requirements surgeons have. The three primary elements on which medical visualization applications are based and will succeed are:

- support for various modes of data and imagery formats
- tools, methods, and algorithms to manipulate and transform the data
- a system that meaningfully renders the results

Advances in the first two elements have continued with the international acceptance of image formats, federal standards for software allowed at the operating room, and tools and visual representations internationally approved to be used by doctors at hospitals or in operating rooms. However, multi-modal data and image resolution are out-pace with the capabilities of the end-user display systems. Given this trend, the next-generation of surgical environments demands we narrow that gap. Our approach centers on scalable projector-based display systems that are intended to enable meaningful and effective visualization in the face of current problems.

Some of the current problems in the medical visualization field include:

- large datasets that contain relatively subtle effects to be explored
- high-resolution displays capable of showing the data in detail
- computationally expensive visualization algorithms
- requirement of refinement transformations
- flexible display configurations

It is clear that data acquisition devices, medical imaging devices, sensors and data simulation environments are producing raw data in unprecedented volumes. Likewise, computational environments and algorithms to refine, enhance, and transform this data have continued to advance. These trends have heightened the mismatch between

the massive scale and complexities in refinement of the data to be examined and the capabilities of the devices that form the end-user environment. Display systems, for example, continue to support relatively low data resolutions and are inflexible in their configuration and operation. This exposes doctors and radiologists in the medical domain to a more challenging task: *the analysis, examination, interpretation and understanding of such complex data and visual results in low resolution devices.*

The motivation for this work is the desire to eliminate such constraints as: *resolution limits, configuration inflexibility, and the strong logical divide created at the framebuffer between the data and the display.*

Visualization applications deal with device resolution limits by providing the user with the ability to control the data and refine the view. In laparoscopic surgeries, one common operation is drilling down the problems of resolution and limited field-of-view through scale by zooming or moving the endoscope around in laparoscopic procedures. While this provides a way for the user to focus the available resources on the data of interest, it does not address the more fundamental mismatch between display resolution and data resolution. We address this issue through a scalable projector-based system that can provide a space of resolution options based on a set of projectors that cooperatively render data. A projector at its widest zoom setting (short focal length) yields fewer pixels per inch (PPI) on the display surface than one set to its narrowest zoom. Control of projector zoom alone can provide a way to vary the PPI of the display. We show how the control of PPI via multi-projector display systems can be managed automatically and achieves resolution scalability.

Most medical visualization applications accept the inflexible nature of the display environment and search for ways to lessen their effect. The same way, minimally invasive surgeons accept the inflexible LCD monitor by mounting them in flexible arms where they can move the LCD monitors to different locations in order to see the images from where they are doing the surgery. This inflexibility normally leads either

to expensive, monolithic, single-application systems, or to systems that simply cannot provide an adequate visualization or visual experience. We address this difficulty by building a flexible system from casually positioned projectors.

When a display is built from many potentially overlapping projectors, the management of the geometric relationship between the projectors becomes complex. There is where the necessity of having a flexible software that allows us to create a seamless, high-resolution display system from overlapping projectors arise.

3.1.1 Flexible Display

Since the relative geometry of member projectors is very loosely constrained, it allows a large number of configurations. Flat walls, completely immersive rooms, high-resolution display on irregular surfaces, and back-projected applications are all possible with only the cost of mounting projectors and sensors in desired locations. By providing a much higher degree of flexibility, we enable new visualization techniques to optimize the users' display configuration in ways that were previously impossible.

The framebuffer, as the interface between the data and the display environment, does not directly support logical abstractions that may be desirable such as data layering and multi-view simulation. Applications that facilitate rapid and seamless switching between logical, functional, and spatial views of the data must collapse and composite these views at the level of the single, common framebuffer. This implies that the framebuffer itself as the abstraction can become an information bottleneck.

In fact it may be more desirable to extend the multi-layered abstraction beyond the framebuffer, all the way down to the display. [41] For example, rapid transitions between views, where each data view is a complex, disjoint distillation of a large dataset, may be best accomplished more efficiently by dedicated devices, each with access to its particular relevant data. Using projection, the display surface becomes an optical framebuffer, where a number of multiple layers, represented as separate framebuffers, can be combined optically.

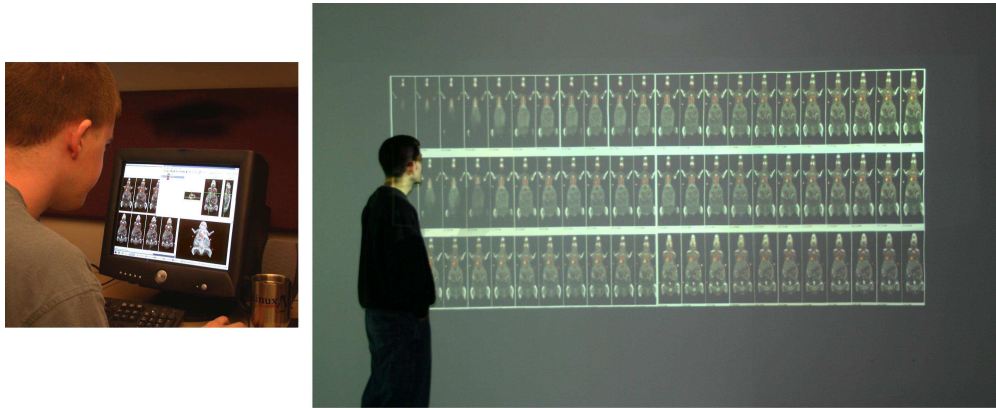


Figure 3-1: (left) Multiple views of medical visualization in a limited display space with limited resolution. (right) Multiple views of medical visualization in a multi-wall high-resolution display

We show how we can support the partitioning of projectors into sets that can function together, each set assigned to manage a single logical framebuffer, with the sets together forming a number of framebuffers that combine optically into a single display.

Stereo displays, for example, can be implemented by assigning one projector set to the framebuffer for the right eye, and a second set for the left eye. In the same way, framebuffers can be mapped to logical or functional data views, or can simply provide auxiliary detail available on command. We believe that support for a set of framebuffers assignable to sets of projectors can enable a number of interesting visualization scenarios.

Frequently, radiologists and doctors need to study and analyze different image modalities such as computed tomography (CT), magnetic resonance imaging (MRI) and X-ray. Furthermore, radiologists need to compare images from different medical tests, compare image changes over time, and conduct analysis of multiple images side-by-side. We believe that the array-of-projectors architecture is a flexible way to create a high-resolution display system that may assist the study and interpretation of medical images. Since there are very few positioning constraints, the projector array can be positioned to completely overlap other projectors to create a number of

coincident layers on the display surface. Such configurations allow experimentation with visualization systems that support new ideas such as smoothly-blended high-resolution insets, continuous shadow removal in front projection displays, and 3D stereo graphics display systems.[35, 32]

With respect to visualization, most approaches accept the constraint of low spatial resolution. Given the limitations, there are efforts to pack more information into the available display real estate. For example, multidimensional software like TGA's Amira[63], which is important in many medical applications, become challenging as the number of dimensions increases. One approach to managing increased dimensionality is to display multiple graphics at the time instant. Coordinated data such as the slices of a CT scan, and axial, sagittal, and coronal views of a 3D data set give a sense of how complex data are and glimpse the requirements of medical visualization for surgeons or radiologists. Multiple window coordinations[26] offer a number of benefits, such as improved user performance over other exploration methods, discovery of unseen relationships, and unification of a desktop environment. The obvious problem is display real estate as figure 3-1(left) shows. As the number of dimensions and coordinated views increase, the number of independent windows grows. When the display cannot grow in resolution, multi-form and coordinated multiple-view data are severely limited.

By increasing the size, brightness, resolution, and flexibility of the display, it is possible to facilitate the data exploration in medical imaging and medical multiple view data sets. Figure 3-1(right) shows a high-resolution, multi-projector display system used to analyze a number of transverse, coronal, and sagittal images. If we compare figure 3-1(left) and figure 3-1(right), we can conclude that the limiting resolution and display space is an important factor while doing medical data analysis and exploration.

Multi-layered visualization is another important technique in medical imaging and can be used in visualization and data exploration in single or multi-user collaboration display systems. In applications such as virtual collaborative environments, people can display local and collaborative simulations simultaneously. In collaborative multi-layered environments, the layers have logical, dedicated semantics and are based on radically different datasets. Furthermore, multi-layered displays can be used to display stereo graphics and stereo visualization where two layers, one for each eye, create the impression of 3D-stereo graphics.

In this chapter we briefly review some related research and discuss a technique we have developed to enable display scalability, flexibility, and multiple layered framebuffers. The high-resolution visualization system we discuss supports flexible, parallel, multi-layered, multi-form, and adaptive-resolution visualization.

We believe that by reducing or removing the display constraints of limited resolution, rigid inflexibility, and single framebuffer architecture, we can narrow the gap between large raw/refined datasets and the end-user display system. Finally, we present examples of how such flexible high-resolution display systems can be used to enhance the visual capabilities and assist surgeons during minimally invasive surgeries.

3.2 Related Work

A primary purpose in building scalable and flexible displays is to facilitate new and emerging visualization techniques. Attention recently has been given to scalable, flexible displays using projectors[42, 43, 33, 20] with a few researchers addressing layered displays.

Using a cluster of computers, we distribute the computational load of rendering to create a scalable system than fits the requirements demanded by particular visualization applications. We believe that a commodity computer cluster or a distributed rendering system can be deployed in research settings in a more cost-effective way and

that this will be useful for display environments intended for visualization, medical data explorations and multi-dimension data simulators.

Clustered computing requires distributed processing and distributed rendering algorithms in order to spread the data and process it in parallel.

There are several benefits and unique points to emphasize about our work.

- We rely on commodity hardware to support applications that cannot afford the cost of more expensive hardware.
- We emphasize flexibility by accommodating unknown display surfaces and arbitrary projector positioning.
- We support out-of-the-box OpenGL applications by leveraging the Chromium distributed rendering project. [\[49\]](#)
- We use Open Source libraries distributed through the GNU license, making our software free and easy to distribute.[\[54\]](#)
- We exercise sub-millimeter accuracy and deterministic methods to geometrically calibrate the display.
- We address the problem of how high-resolution display systems can be used in surgical environments.
- We support passive stereo visualization.

Parallel and distributed rendering has been around for decades. Early methods used high-performance computers and supercomputers to distribute the computation amongst a number of processors and distribute the rendering load to a number of different graphics pipes. For example, the SGI Onyx 2 family[\[65\]](#) is a shared memory supercomputer scalable in CPUs, memory and graphics pipes and often used to visualize complex simulations and data-sets. Figure 3–2(left) shows an Onyx supercomputer being used for visualization and active stereo visualization. During the last five years, parallel rendering and the use of cluster computing to render distributed



Figure 3–2: (left) An SGI Onyx 2 supercomputer used for visualization and for active stereo visualization (right) Cluster of computers connected by a high-speed network creating an scalable distributed computing system by using commodity hardware

graphics have taken popularity and there are a number of companies selling solutions involving distributed, parallel rendering but using a cluster of rack-mountable computers.[65, 66, 67, 68] For example, the SGI Prism is a Linux based visualization system created from a number of rack-mountable computers. This is an example of the movement from supercomputer hardware to scalable commodity hardware. Figure 3–2 (right) shows our distributed Linux cluster built from commodity and common hardware.

There are a number of different rendering methods used by clusters to distribute the geometry and speed-up the process by using multiple video cards. Rendering methods like *sort-first* divide the display space into a number of regions, which can vary in size and shape.[30] A portion of the display is assigned to each rendering process, which is responsible for rendering its portion of the display in parallel. The *sort-last* approach, also know as image composition, assigns a rendering process to perform both geometric processing and rasterization in a way that is independent of all other rendering processes. Local images rendered by the individual processes are composed together to form the final image. [29, 28, 31]

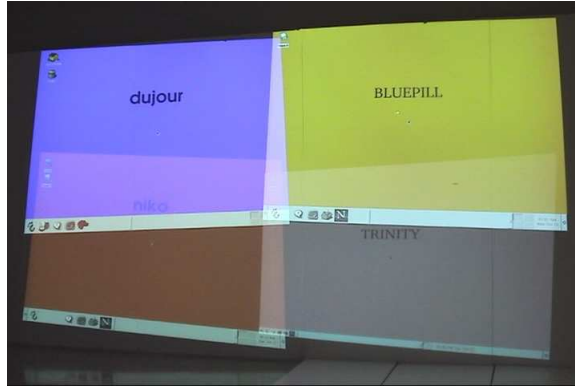


Figure 3-3: A set of casually aligned projectors used to create a seamless high-resolution display system

In either of these cases, the array-of-projectors architecture allows side-by-side positioning to creating a scalable display. Such displays play an important role in multi-views applications because the size and the resolution of the display allow the user to display a set of different views or angles in the same display area without sacrificing the resolution of each of the views. In the medical domain, the accuracy of some results is closely related to the quality of the obtained image and the quality of the display image.

3.3 Implementation

Although large-scale, high-resolution displays may help solve issues for medical visualization and minimal invasive surgeries, it is challenging to build them. Issues of cost, flexibility, setup and maintenance all play a role in making use of the technology for visualization. Our approach is to build multi-projector display systems from commodity hardware (projectors, PCs, and graphics cards). The support of a large number of projectors arranged in any geometric configuration leads to the scalability and flexibility we wish to provide.

We address the primary problems of flexible, scalable deployment and cost-effective use through a commodity hardware-based design. We assume projectors are arranged in a tiled configuration, where several projectors are positioned together

to create a seamless display area. This way, the resolution (pixels per inch) of the surface can vary by changing both individual projector settings (zoom and position relative to the display surface) and collective projector geometries. Figure 3-3 shows a set of casually aligned and positioned projectors. Such setups and configurations can be used to create a seamless high-resolution display system by utilizing a number of casually positioned projectors. The high-resolution feature comes from the combination and addition of the resolution of each independent projector that is part of the system.

Standard distributed rendering approaches make it possible to coordinate the operation of this projector set[43, 42]. We focus here on the primary issues of *geometric and photometric correction*, which we solve through a camera based monitoring system. This system makes very few assumptions about the display surface shape and the projector locations. We address geometric and photometric correction here, and later we explain how these high-resolution display systems can lead to new possibilities for medical visualization and surgical trainings.

3.3.1 Chromium

To distribute graphics across multiple computers and multiple projections, we use Chromium[29, 49]. Chromium is an open source software for interactive rendering and manipulating streams of graphics API commands on clusters of workstations. Chromium is derived from Stanford’s WireGL project[44].

The main reason why we picked Chromium as the underlying system to distribute OpenGL and graphics across multiple computers is that it allows the modification, deletion or replacement of graphics commands on-the-fly from programs written in the OpenGL programming language without the necessity of recompiling the software. That is, unmodified off-the-shelf OpenGL applications can be run through Chromium to distribute the rendering load between a number of different computers.[29] The

system uses stream packages to move geometry and imagery across a network as required by a distributed application.

Stream transformations are performed by OpenGL "Stream Processing Units", or SPUs. SPUs are implemented as dynamically loadable libraries that provide the OpenGL interface, so each node can load the required libraries at run time and build an OpenGL dispatch table based on the transformation to the stream required of the specific SPU. The SPU takes a single stream of OpenGL commands as input, and produces zero or more streams of OpenGL commands as output.

A node's stream transformation does not need to be performed by only a single SPU; hosts can load a linear chain of SPUs at run time. SPUs can be chained together to achieve more complex results. Using this feature, an SPU might intercept and modify calls to one particular OpenGL function and pass the rest untouched to its downstream SPU.

3.3.2 Geometric Calibration

Tiled display systems face the physical alignment problem with the recognition that aligning the projectors manually is very challenging. It is possible to build systems through precise physical alignment, but is a time-consuming process and will require frequent realignment to ensure each projector generates an exact rectangular image necessary to align with neighbor projectors. Planar surfaces are easier for manual alignment of projectors than arbitrary surfaces. For example, in curved display surfaces, it is hard to generate a rectangular image which can be aligned with a neighbor projector.

With the vision that high-resolution, multi-projector display systems can be used to assist surgeons during minimally invasive surgeries, we have created a calibration mechanism for planar surfaces as well as a calibration mechanism for arbitrary surfaces that within seconds can generate a seamless display system using a number of casually aligned projectors. Figure 3-4 shows a diagram of four projectors that are

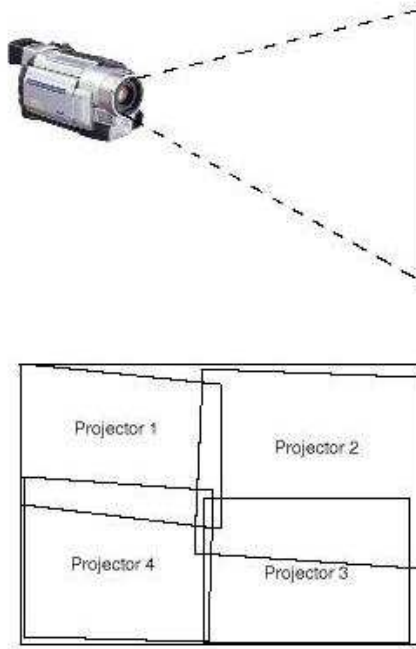


Figure 3-4: Diagram of our camera-based multi-projector calibration process

casually positioned. Using a single camera that can view the whole display, we can calibrate the system to create a seamless display.

When a precise manual alignment is used to calibrate the projectors, it is difficult to achieve correct alignment, and it is rare to change the configuration once an alignment is obtained. Most often, precise physical mounting devices and restricted geometric configurations are used to assist the alignment process. Even with these aids, vibration, weight, and lamp-changing all necessitate frequent re-calibration, which may lead to hours spent maintaining the system.

”Geometrically correct” means that geometric primitives in the displayed imagery, such as lines, triangles, polygons and texture, appear correct to the viewer, regardless of the individual projector positions, their relative geometry, and the underlying display surface. To create a flexible high-resolution display system, we need

to guarantee the capability of displaying on arbitrary display surfaces and with any number of projectors.

Planar Surface

Our planar calibration software is based on the work of Raskar et al[47]. This method calibrates casually aligned projectors with a vision-based approach. From images taken of the display system through a video camera, we can compute homography matrices between projector space, camera space, and display space. During rendering time, we can pre-warp the images to compensate for the oblique projection of the projectors. After applying the homography transformation to each of the projectors, we can display a seamless, geometrically correct image throughout the whole display system.

The planar calibration and the pixel mapping between uncalibrated projectors involves computing the camera-to-projector, projector-to-projector, and display-to-projector homographies. To obtain these homographies, each projector displays a chessboard calibration pattern. Four or more point correspondences are automatically detected and then used to compute the homography between the projector image and the camera image plane. The homographies are extracted using pattern recognition techniques of OpenCV[51]. A simple chessboard pattern is projected for each display and captured by the camera. The feature points of the pattern are extracted using FindChessBoardCornerGuess() function and homographies are computed with the FindHomography() function.

It has been shown that the location of the corners of the chessboard can be obtained with sub-pixel accuracy by calculating the center of mass of the responses.[38] This makes the calibration result a sub-pixel accuracy calibration technique.

The theory behind the homography matrices and the chessboard calibration approach is the following[47]. In computer vision cameras and projectors are often equated due to their similar characteristics. Given two cameras (i.e. video camera

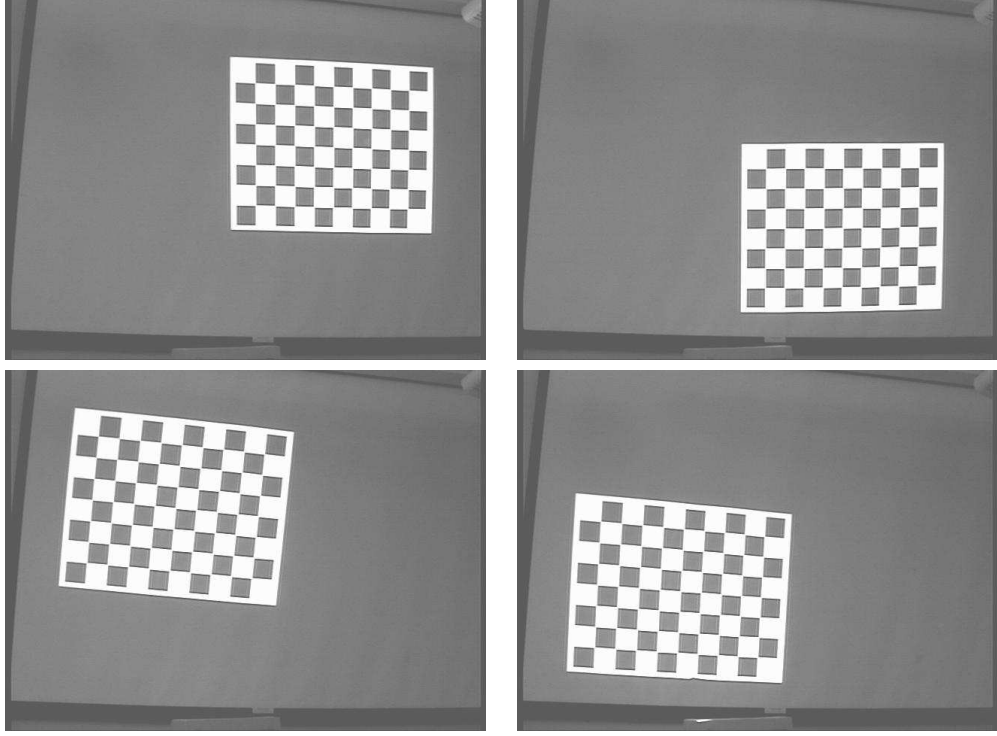


Figure 3-5: Chessboard calibration patterns

and a projector), viewing points on the same 3D plane ψ , the positions of a single point in the two images are related by a 3×3 homography matrix H , defined up to scale. That is, if m_1 and m_2 are projections of a 3D point M which belongs to ψ , then

$$m_2 \cong Hm_1$$

where m_1 and m_2 are homogeneous coordinates and \cong means equality up to scale.

Knowing about this relation, we use one single camera ζ to record all the projected images. We first project the chessboard pattern from each projector sequentially and capture the projected image on the display surface by a single camera. By extracting the feature points from the 2D camera image corresponding to known 2D points from the projector pattern, we can determine the 3x3 homography between the static camera and each projector based on

$$u_i = H_{c_i}x_c$$

where (x_c = camera coordinates, u_i = projector coordinates and H_{c_i} = a homography matrix). With four or more correspondences between camera image and projector pattern, the 8 unknown parameters of $H_{3 \times 3}$ can be computed using least-square method.

Figure 3-5 shows a set of four casually aligned projectors calibrated using the chessboard calibration technique. After computing the corresponding homography matrices, we were able to create a seamless display surface.

Arbitrary Surface

To support and create a flexible display system, we cannot assume that the display surface will always be planar; we need to support and been able to create a high-resolution display system in planar and non-planar surfaces. Because the typical human field-of-view is around 160 degrees, curved display surfaces create the feeling and perception that the person is immersed into the graphics been displayed.

To calibrate casually aligned projectors for arbitrary surfaces, we use a visual, camera-based approach which helps us to correspond pixels from the projector image to the camera image plane and reach our goal of a seamless, geometrically correct display system. Ideally, we would like to correspond every single pixel of the display system to a pixel in the camera, but due to the limits in the camera resolution and the desire of creating a scalable display system, we need to modify the problem to correspond a number of pixel of the display to a pixels at the CCD of the camera. To do so, we display known patterns which can give us an accurate mapping from the display system to the camera image plane.

If the display surface is completely arbitrary, we project a number of equally spaced fiducials onto the display surface from each projector involved in the system. We implemented the system to support different types of fiducials. Depending which fiducials are used, the calibration results can increase or decrease in accuracy and the computational time to locate them can increase or decrease. Our software allows

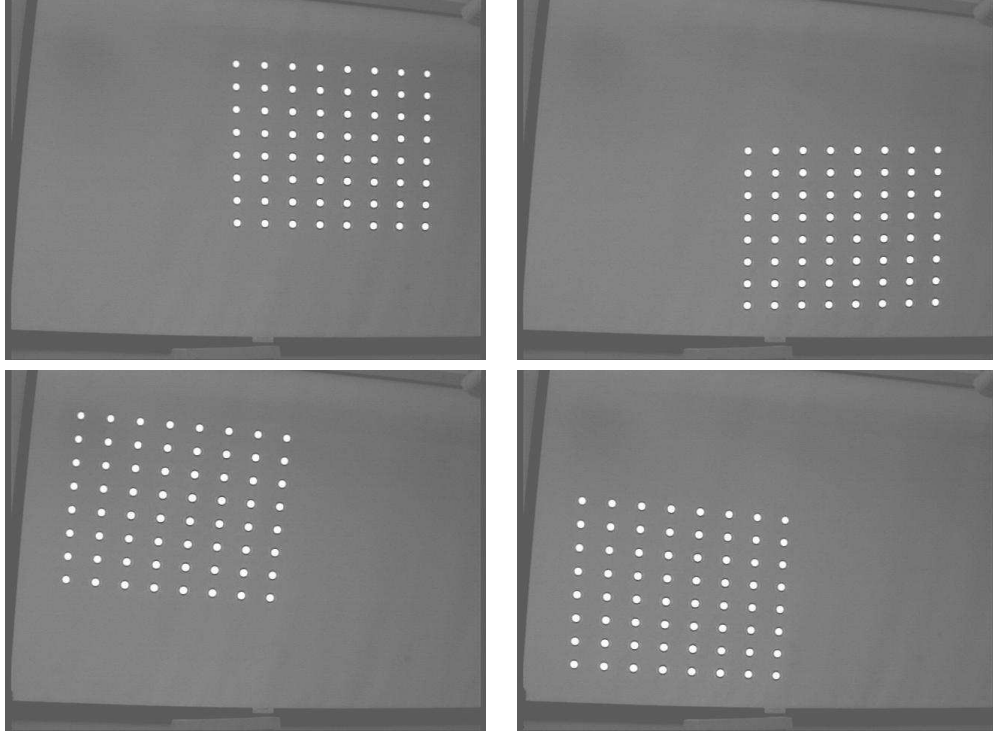


Figure 3-6: Blobs calibration patterns

the user to pick between circles, squares or Gaussian blobs as the fiducials used for the calibration step. From those fiducials, the Gaussian blob grants the best results because the centroid of the blob can be detected with sub-pixel accuracy as proved by Yang[42]. Figure 3-6 shows the calibration of a display system of 4 projectors using the arbitrary surface calibration technique by displaying Gaussian blobs. Each of the projectors display the same number of blobs from which we compute the tessellation that creates a seamless display.

This approach uses a stationary camera, positioned where the viewer would be seated when the display is in operation. The camera allows us to determine the appropriate projector warping function to create geometrically correct imagery for a given viewing position.

The fiducials can be logically connected to form a tessellated grid. The display surface illuminated by the projected fiducial is observed by the camera. The tessellated grid is determined in the camera's image plane.

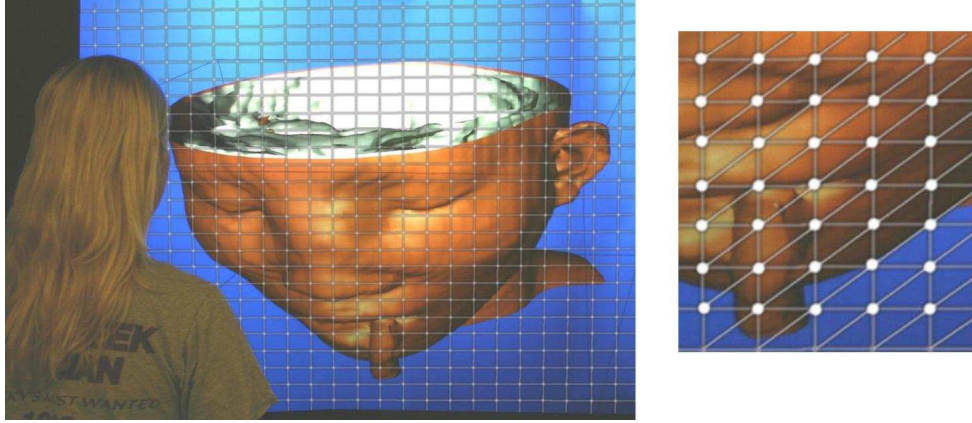


Figure 3–7: (left) By overlapping an image of the calibration grid, of a six-projectors display system, with a medical visualization in the same display we can see how the surface is approximated by the tessellation process (right) A zoom to a portion of the display where multiple projectors overlap

Because we don't assume any orientation of the projectors, we use a binary-encoding scheme proposed by Raskar et al[20] to assign a unique ID to each of the displayed fiducials. That way, each of the blobs located at the image plane can get a unique ID, from which we can generate a tessellation grid at the image plane without any uncertainty. Figure 3–7 shows an image we obtained by overlapping the calibration result and a medical visualization. By taking a picture of the tessellation grid displayed after the calibration process and an image of a visualization from the same camera position, we can overlay the two images to show how the distributed system is warping the image. Each of the triangles are pieces of the framebuffer that the calibration and display technique is warping. In the end this creates a seamless display image.

We can talk about the accuracy of the calibration in two different ways: *calibration accuracy between projectors* and *surface estimation accuracy*. The calibration between projectors is a sub-pixel calibration when using Gaussian blobs. To talk about the surface estimation accuracy, we need to talk about the surface variation between the points of the tessellation grid. From figure 3–7 we can see that the size of each block is around three inches. In curved displays, those blocks or triangles are displayed as

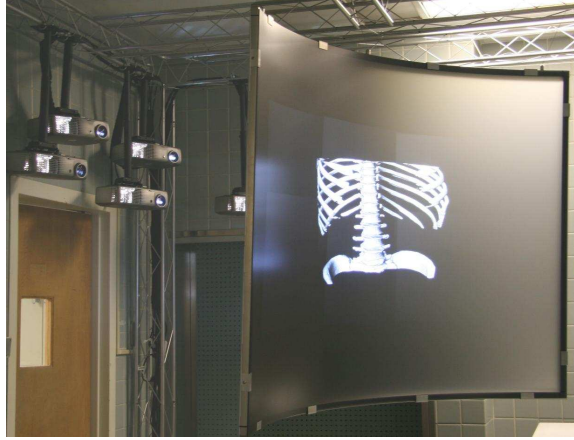


Figure 3–8: Seamless display system in a curved screen

blocks or triangles, respectively. That is, if there is a lot of curvature or surface change inside one of those blocks, the final visual representation will not look seamless. In surfaces with a lot of curvature or variation, if we increase the numbers of displayed blobs during the calibration step, then we can increase the accuracy of the surface estimation. On the other hand, in planar surfaces, only few points can be used because there is no surface variation.

From the distributed rendering point of view, after the fiducials are detected, they are logically connected, forming vertices of a triangulated grid, which we use to define a piecewise warping function. This warping function defines how to transform the framebuffer of each rendering computer before display such that the display will appear geometrically unified as a part of the complete projector set. By using the OpenGL call *glCopyTexSubimage2d*, each rendering computer captures the content of the framebuffer, apply the triangulation mesh to that image, then puts that image back to the framebuffer and displays it.

Figure 3–8 shows a medical CT reconstruction been displayed in a curved, rear-projection screen. After the calibration process, we end up with a unified, seamless and high-resolution display.

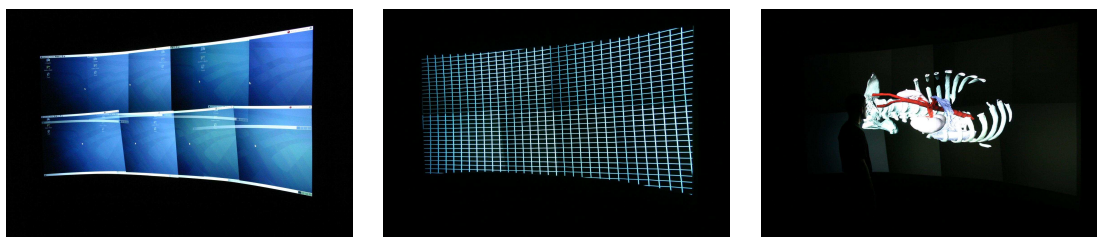


Figure 3-9: (left) A high-resolution display system created by casually aligning 8 projectors in a curved screen. (center) A grid displayed through the display system to show the calibration results. (right) Medical visualization is possible after the calibration results

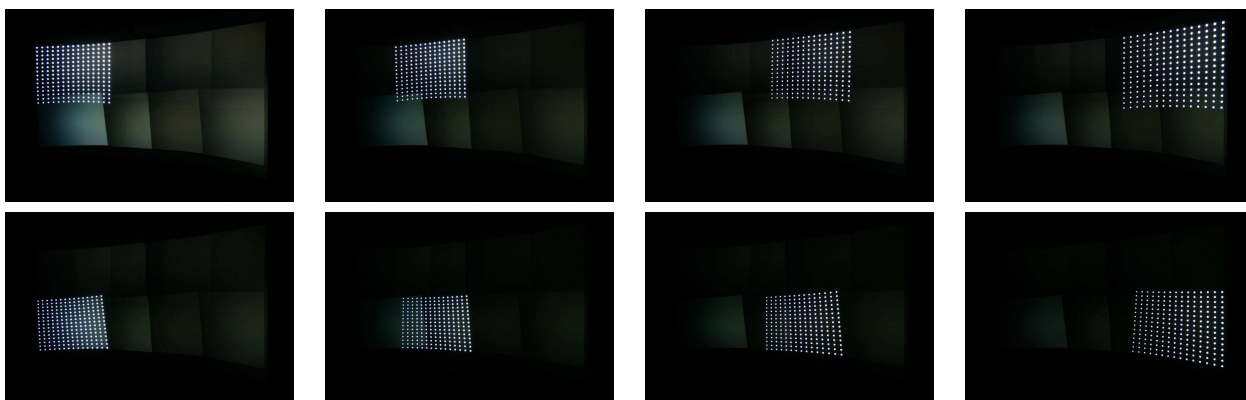


Figure 3-10: Calibration step of 8 casually-aligned projectors. By projecting equally spaced fiducials to each of the projectors that are part of the display we can generate the warping required to create a seamless images

Figure 3–9 shows an eight projector display system with a curved screen being calibrated using the tessellation approach: displaying equally spaced fiducials from each of the projectors involved in the display system. Figure 3–10 shows the projection of equally spaced fiducials in each of the projectors that are part of the system so we can calibrate the eight projectors to form a seamless display. As we have mentioned before, the projectors are not physically aligned, but the alignment is dynamically calculated through displayed fiducials detected by the camera.

The system automatically calculates the piecewise warp, and supports the calibration of a system of projectors at a cost of approximately 12 seconds per projector. For example, we can modify and then calibrate an 8 projector system in a little more than a minute. When each projector is XGA resolution, a 8-projector display can display an image of about 8 mega-pixels.

After the piecewise warp has been calculated based on the detected fiducials, we display a uniform grid across the entire display in order to demonstrate that the warp function is correct. Figure 3–9 (center) shows a uniform grid across a curved display system with a resolution close to 8MP. Such a display system has been used for scientific visualization and immersive simulations.

3.3.3 Photometric Calibration

Since intensity varies among projectors, and since overlap regions are multiply-illuminated, there are areas in the display that are noticeably brighter to the user. In order for the complete display to be seamless, we need to attenuate brightness in overlap areas so that the user has the impression of one continuous display environment.

Photometric calibration achieves a correspondence between intended image luminance and the chrominance sent to the projector, and the actual luminance and chrominance of the display system.

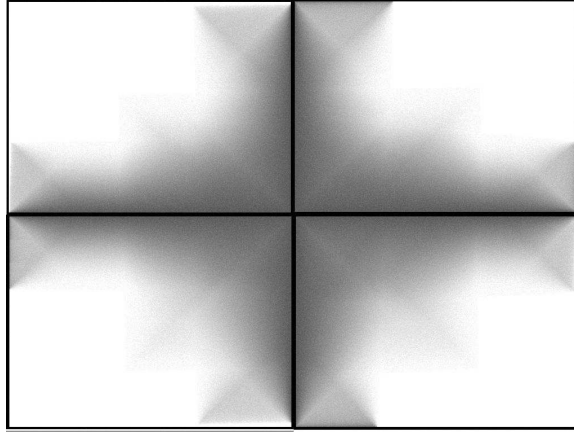


Figure 3-11: Alpha images used for the photometric calibration step

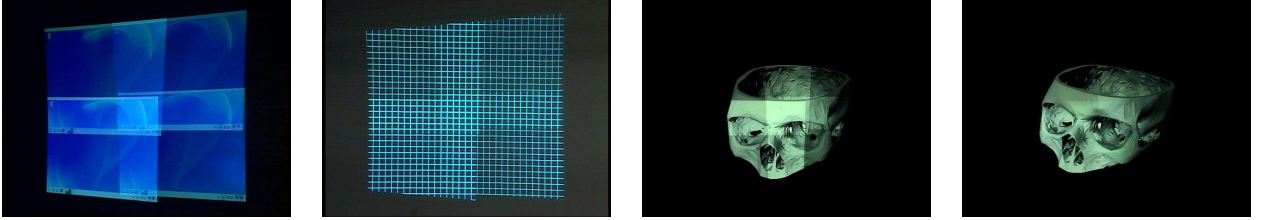


Figure 3-12: Photometric results. A set of casually aligned projectors can be dynamically calibrated, but without the correct photometric correction, the user will see brighter areas which can be distracting

Our system addresses the photometric problem by calculating a blending factor based on the geometry of overlapping projector regions. From the computed projector-to-projector matrices or from the tessellated grid, we can compute the overlap regions of the system, as well as the number of projectors that overlay an specific area. Once detected, the projected brightness in overlapping areas is attenuated based on the number of projectors contributing to the display area. This method creates an efficient first order approximation to the photometric issue that substantially supports the illusion that the user is viewing a single, continuous display. Figure 3-11 shows the alpha mask images computed from the tessellation grid from a four projector casually aligned display system.

Figure 3-12 illustrates the photometric problem and solution. With multiple, overlapping projectors, geometric warping corrects the structure but does not correct

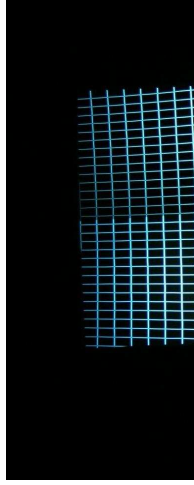


Figure 3–13: Calibration results and errors caused by the camera lens distortion

for brightness. The brighter areas are still noticeable even with correct geometry warping (figure 3–12(center-right)), and that usually causes distraction to the user. By using simple alpha blending presented in figure 3–11, with alpha values derived from the overlap structure detected in cameras, the display becomes structurally correct and photometrically blended (figure 3–12(right)). An extensive and in-depth study about photometric correction for multi-projectors display system has been done by A. Majumder. [39]

3.3.4 Improving the Calibration Accuracy

Our calibration technique uses commodity hardware to compute the warping required by each projectors to be able to create a seamless display system. Because the calibration software requires a camera that can see the complete display surface, we usually need to zoom-out the camera as much as we can. Because cameras, specifically, lenses are not perfect, and they have a lot of distortion as discussed in Section 2.3.1, we have seen that the distortion affects the calibration results.

From figure 3–13 we can see that in the edges of the calibrated display system, the lines that should look straight in the vertical direction, are slightly curved. That is caused by the lens distortion of the camera used for the calibration process.

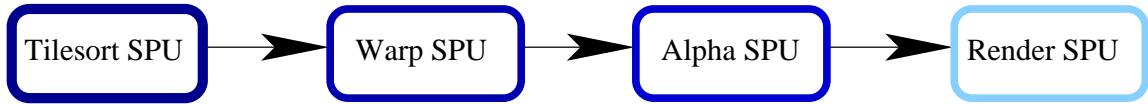


Figure 3-14: SPU diagram

By calibrating the intrinsic parameters of the camera (see section 2.2.2) before the multi-projector calibration process, we can undistort the images and obtain better calibration results, hence improving the calibration accuracy.

3.3.5 Color Correction

Because the color, brightness and illuminance of each projector is different despite the fact that we usually use the same manufacturer and model of projectors, we created a simple color correction algorithm that takes images of each projector and computes the white intensity of each projector. By using the minimum white intensity, we can decrease the overall intensity of the display system by that number, creating an intensity image that can be applied to each rendering computer before the rendering step to create a more consistent image throughout the display system.

The only problem with this approach is that in some surfaces (e.g. curved surfaces) the intensity changes according to the viewing angle. In such situations, this approach does not work or it will be really dependent on the position of the camera. A depth study and possible solution for projector's color difference has been done by Aditi at UNC. [40]

3.3.6 Implementation of Calibration in Chromium

We implemented the calibration software as an OpenGL application that can run through Chromium. By using the *video4linux* and *lib1394* libraries, we implemented our own class to capture images from a variety of different cameras and input devices.

We run the calibration through Chromium as any other OpenGL program will run. After the calibration process that outputs the matrices or triangulated mesh

that define a per-projector transformation required to create a seamless image using all the projectors, we need to apply the actual warp just before the rendering step. To apply the geometric warp, we implemented the *WarpSPU* as a stream processing unit. The *WarpSPU* runs in each end-node and captures everything that the main-node with the tilesort SPU is sending to that specific rendering computer. There, the *WarpSPU* applies the mesh corresponding to that computer and puts everything back into the framebuffer so it can be displayed.

The same way we send the tessellation information of each of the rendering computers. The *warpSPU* can also receive homography matrices used to warp the image to correct the oblique geometric problem of the casually aligned projectors.

The photometric calibration creates an image that is applied to the OpenGL stream chain before the rendering step. The head-node computer sends the alpha image to each of the rendering computer and the *AlphaSPU* applies that blending image to each of the rendering computer's framebuffer just before the rendering step, which at the end creates the seamless display.

Figure 3-14 shows a diagram of the SPU chain we use in Chromium to create a seamless display system from a casually aligned projectors.

3.4 Applications

This is the essence of the provision for flexibility: the camera-based system yields the flexibility by avoiding the need for physical alignment. The projected imagery is aligned through warping operations derived from the camera, which allows the user to deploy and experiment with a practical, flexible tiled display.

New display capabilities hold promise as an enabling technology for advanced medical visualization systems designed to exploit them. In particular, we anticipate in the following areas:

3.4.1 Scalable, Adaptive Resolution

Displays can function at lower and higher resolutions as required by an application. More projectors in various configurations gives a scalable and controllable way to improve and experiment with issues in brightness, pixels-per-inch on the display surface, and trade-offs such as ppi vs. brightness. An interesting capability that is now practical, which has not been widely explored, is the usefulness of the adaptive display.

3.4.2 Flexibility

The display package we have implemented solves the geometric and photometric problems together with software for distributed rendering. [49, 50] This makes possible the rapid deployment of a flexible high-resolution display system. We have conducted on-the-spot demonstrations with this system and proven its utility in collaborative efforts in both visualization centers and individual use (office and small labs). We can use the system to create, for example, a portable medical visualization display, or a portable laparoscopic display system, which can be set up and calibrated in very short order thanks to the camera-based solution for the geometric and photometric problems that would otherwise demand careful and rigid physical alignment. The required hardware does not involve anything other than PCs, graphics cards, a computer network and projectors. Clearly the projectors are the crucial high-dollar component, although we can operate with almost any model and can scale the system from a few to as many as is practical on a local area network (currently 12-16 projectors). Note that a 16-projector system where each projector is capable of a Mega-pixel of display resolution leads to a 16 Mega-pixel display device, which is well beyond the foreseeable capability of the desktop monitor. The flexible calibration of the system allows the projectors to be arranged as desired so that a number of configurations can be used depending on the application and subjective wishes of the user.

3.4.3 Medical Visualization

In an effort to push forward high-resolution display system technologies for medical applications and surgical environments, we have to develop different methods to support 3D data sets, DICOM imagery, X-ray images and other medical image formats in our display system. By harnessing the power of Chromium in being able to execute unmodified OpenGL programs in a distributed fashion, we have been able to run medical software like Amira, Amide and ivview in our high-resolution display systems. [63, 64]

Further, in an effort to facilitate the analysis, interpretation and understanding that radiologists need, we support the display of a number of time-variant images side-by-side so the radiologist and surgeons can move away from the "light-box" currently in use to make such analysis from a high-resolution display environment, preserving the resolution and details of each of the images.

3.4.4 Layering

The flexible positioning that our techniques support provide an interesting environment in which to experiment with multi-layered and multi-view visualization problems. The multi-layered system follows directly from the tiled-projector algorithms, and can be exploited to support dedicated, high-resolution insets, real-time video overlays, and even polarized stereo graphics. [53, 52] For example, we have demonstrated a stereo system where one layer of the display maps to the right-eye framebuffer, and another layer maps to the left-eye framebuffer. Figure 3–15 shows a calibrated multi-projector display system enable to display stereo images. Now that stereo imaging and stereo laparoscopes are becoming popular, such stereo displays can be used to present the surgeon with a flexible high-resolution stereo video. Figure 3–15 shows a test we were doing involving multiple-projectors. One set of projectors displays a specific color to the left eye and the other set of projectors displays only a specific color to the right eye.

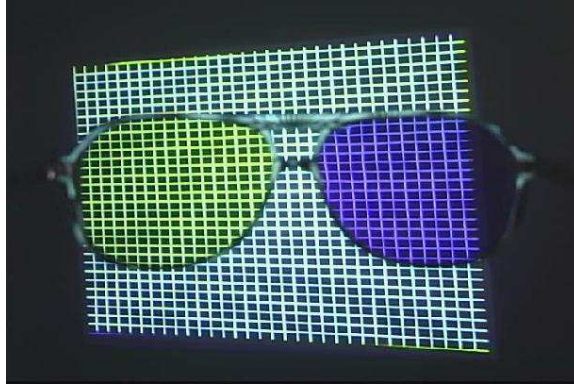


Figure 3–15: By using multiple projectors and our calibration technique, it is possible to create stereo display system that, by wearing the required glasses, the user can perceive depth

Another interesting application that follows the multi-layered system is auto-stereoscopic displays. An auto-stereoscopic display is a 3-D display that presents concurrent independent views of the imaged scene without special viewing aid. It has been proven that a set of projectors can be used to create an auto-stereoscopic display[34], and we believe that with the flexible positioning technique that our system provides, it is possible to create such a stereo system with even less constraints and more flexibility.

3.5 Software Distribution

The alpha version of the calibration software was released in January 2003. After that, a number of universities and research labs have expressed their interest in the project and source code. The project has continued its development by independent people as well as by the REVEAL project. [58] By using the source code and applications, REVEAL researchers are trying to create a software suite robust enough to be distributed and used at hospitals.

Some of the universities using or that at some point used our calibration software are University of Kentucky, University of Puerto Rico, Massachusetts Institute

of Technology, University of Maryland, University of Nottingham, and Zhejiang University. Some research institutes that have used our calibration software are Sandia National Laboratory and Argonne National Laboratory.

CHAPTER 4

STUDIES AND RESULTS

In this chapter we discuss, analyze and review some of the software, programs and projects we have done as part of the next-generation of surgical environments to support our idea of how information technology can help and assist surgeons during minimally invasive surgeries.

4.1 Performance Analysis and Evaluation

Minimally invasive surgeries demand greater skills from the surgeon. An objective way to analyze, compare and evaluate the surgeon's performance is of great benefit to the medical field. In interest of identifying bottlenecks in the surgical procedures, helping surgeons improve their skills and facilitating the comparison of different surgeons in an objective way, we have created a way to analyze surgeon's performance. Minimally invasive surgeries are created to benefit the patient, but the extensive motion of the surgical instruments, including the endoscope, during the procedure can cause more trauma to the patient than ideally necessary. To this end, we created a performance evaluation software based on motion. By tracking the 3D position and orientation of the surgical instrument over time, we can create a metric to measure surgeon's performance and compare surgeons in a given task.

For a given task, by tracking the 3D position and orientation of the instrument in a series of consecutive video images, we can compare the overall motion of the procedure. To test our motion-based performance analysis software, we gave the

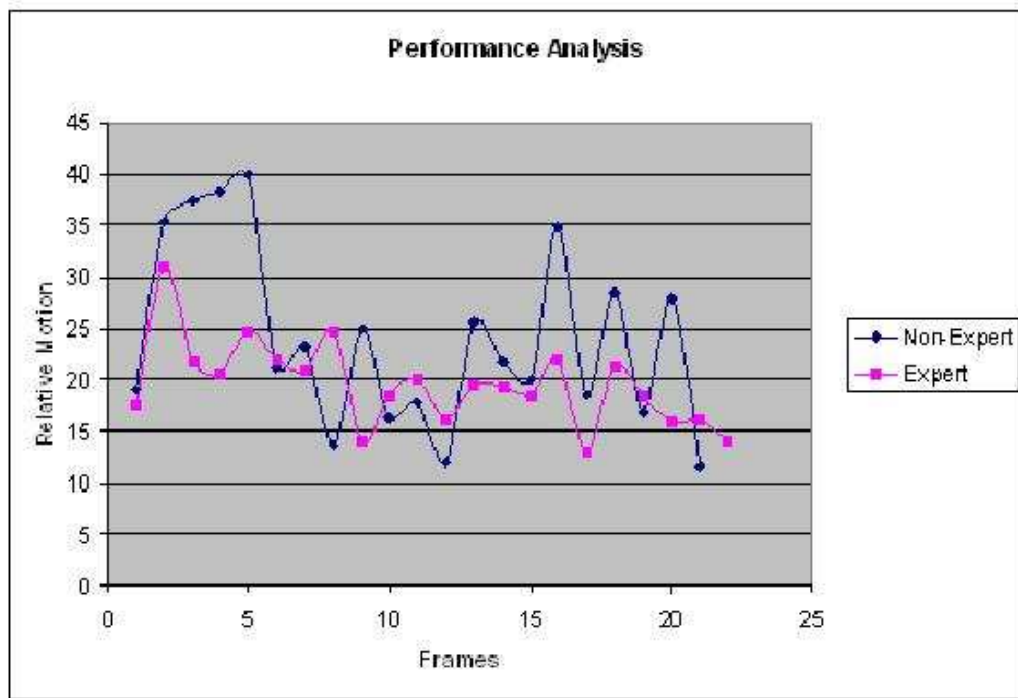


Figure 4–1: Plot of the performance analysis experiment

same task to two different persons: a novice and an expert. We tracked the 3D position of the instrument over the period they were doing the procedure and after they finished, we were able to compare the difference in their performance.

Consider figure 4–1. The two curves on this graph show 3D motion estimates for the stapler instrument over a set of frames. The value plotted as the height of the curve for each frame value is the 3D position of the instrument measured relative to a fixed point. The curve that corresponds to the expert performing the stapling action shows much less relative-motion variation than the curve corresponding to the novice. In this case, economy of motion over a set of frames, evaluated in 3D to capture movement toward and away from the camera, shows how an expert handles the instrument in a way that is measurably and objectively different from the novice.

This objective performance measurement method can be of great benefit to train surgeons in the basic skills of minimally invasive surgeries. By merging a motion-based performance analysis technology with MIS standard metrics, we believe that we can

create a system which gives trainees more feedback about how they can improve their skills.

4.2 Heterogeneous Display Systems

In our proposed architecture to support the next-generation of surgical environments, we need to take advantages of the high-resolution display system and the processing step to present enhanced video and any other imagery that might assist surgeons during surgical procedures. We believe that the display environment we have demonstrated will provide valuable insight into how best to move beyond the "in-the-box" display systems that have been only incrementally improved over the past 20 years. The display framework removes key constraints on display real-estate (resolution and configuration), embraces the ability to include seamless stereo regions, and still provides the ability to keep information available in a way that is tightly-coupled and potentially less distracting for the surgical team.

With the high-resolution display system we have shown how to break free from the display constraints currently present at the operating room by moving that technology forward with a hybrid, heterogeneous display framework that preserves key characteristics of current systems (low latency, specialized devices). We have engineered a hybrid display and currently we are using it to build a surgical simulation and training environment within which we can evaluate both the technology and the performance of subjects using the technology.

4.2.1 Multi-context Display System

To show that it is possible to integrate real-time video, 3D pre-operative data sets, external video and tracking information obtained from the distributed system of the proposed architecture of the next generation of surgical environments, we developed an OpenGL program that can display all of them. The software gives the user the flexibility to view or hide each of the windows, move the windows to any place in the

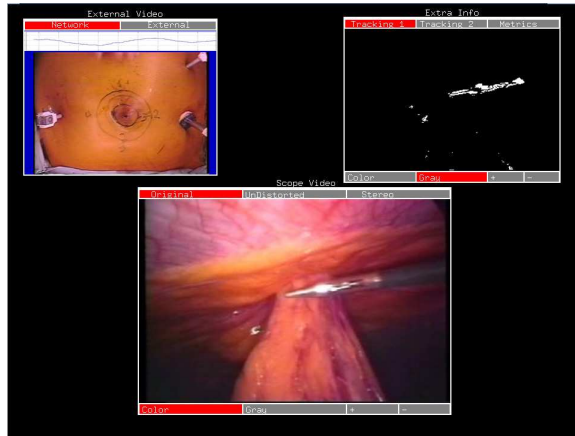


Figure 4–2: OpenGL software that creates a unified display that can receive multiple real-time video, tracking information, and other pre-operative imagery

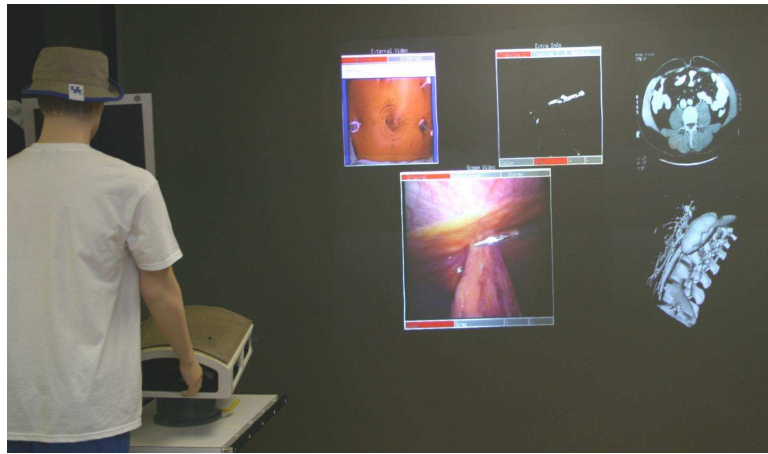


Figure 4–3: A multi-context high-resolution display system used to display multiple images and videos in a single display space.

display system, and increase or decrease the size of each of the windows as is his or her preference. Figure 4–2 shows an screenshot of the software.

After we demonstrated with our OpenGL software that it is possible to incorporate multiple image formats, videos and 3D data sets in a single OpenGL window, we extended our software to a high-resolution display system. Using an array of casually aligned projectors, we created a unified environment where minimally invasive surgeons can have access to the scope video, images, data and pre-operative information they might need during the procedure. Currently, minimally invasive surgeries (MIS) are performed by looking to a single LCD display. We believe that by presenting

more visual information to the surgeon, we can assist and fulfill their needs during the surgery. Figure 4–3 shows a seamless display system created out of 9 projectors where we extended the MIS setup from a single LCD display to a high-resolution coordinated multiple-views display. The display is capable of showing X-ray images, CT data, 3D reconstruction, real-time laparoscope video, real-time external video, and apply some "on-the-fly" tracking analysis to the video been displayed. We believe that such system can assist surgeons because they have access to real-time video, enhanced video, access previously taken X-ray images and other medical records by just keeping the focus of attention in the seamless display system.

4.2.2 Hybrid Display Systems

We have engineered a display system, with high-resolution, multi-context and hybrid mechanisms to facilitate surgical scenarios. Our unified display system can display real-time and enhanced video from a variety of scopes and cameras, 3D data, metrics, and tracking information. See figure 4–4. The normal 2D video and images can be seen in parts of the display, while other areas are enabled for passive, polarized stereo. Also, it is possible to smoothly-incorporated traditional displays devices such as LCD panels or plasma TVs, creating a hybrid, but unified display environment. Figure 4–4(left) shows a display system with an incorporated LCD panel. Cameras, hidden behind the screen, communicate images to the software and automatically configure the display layout. The computer cluster acts as a distributed platform for running simulation code (collision detection, for example) as well as processes that can enhance live video from scopes. The output from multiple scopes can be shown simultaneously without loss of resolution since the 3x3 projected grid has a total resolution of over 9 mega-pixels.

Using this environment, we can assemble 3D data, pre-operative CT-scan data, live scope video, procedure slides from a medical image database, metric overlay

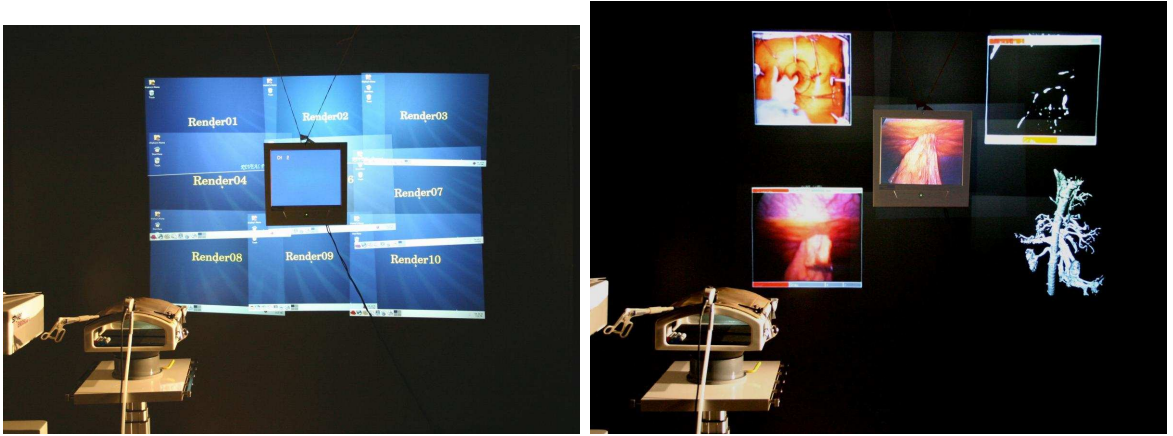


Figure 4-4: The hybrid display system

information, and other important custom data (e.g., stereo reconstruction and identification of anatomy) to create a fused, unified display.

We are experimenting with configurations that guarantee a reduction in distraction and streamlines the user’s focus of attention under specific constraints in order to better support particular procedures and tasks. Because the display system supports stereo, overlays, scalable resolution, and the potential for side-by-side views to overcome latency issues, we are able to study new configurations that have the potential to improve performance and reduce the onset of fatigue. Additionally, the scalable screen real-estate provides a substrate with which we can integrate features such as remote collaborative consultation and video conferencing on-demand.

Our working prototype consists of any configuration of 9 projectors and a core LCD panel display. Our software system runs on a tightly-coupled computer cluster and drives a rear-projected environment. We use a Stryker laparoscopic training stand in front of the display as a baseline configuration.

4.3 Distributed Video

Minimally invasive surgeries are dependent of the video obtained from the endoscope. We have developed a distributed video player and real-time video viewer based on NCSA’s pixel blaster software[55]. Now, with our software we can capture

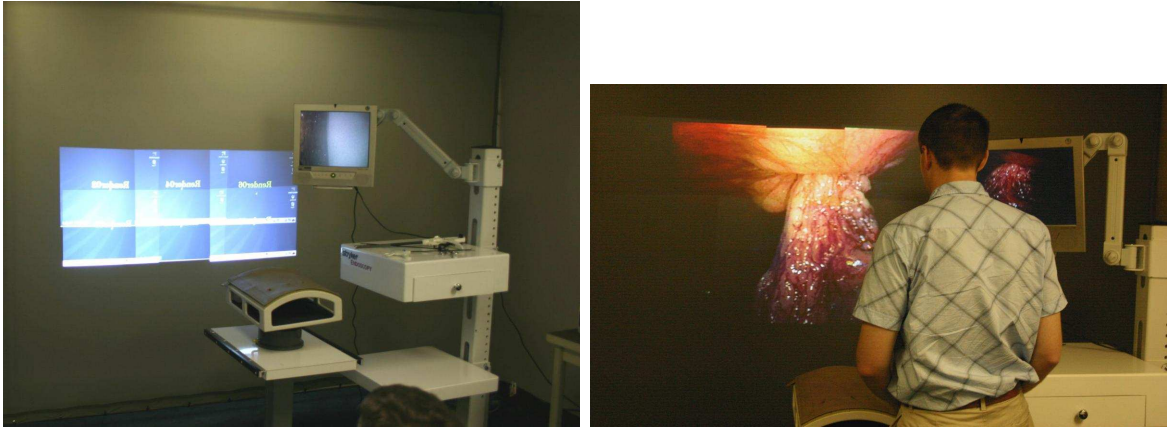


Figure 4-5: Distributed Video

real-time video from the endoscope, send it through the distributed processing part of the proposed architecture for the next generation of surgical environments, and then display the video in multiple projectors.

Figure 4-5 shows a 6 projector display system where, after calibration, we can display real-time content through the set of projectors in a distributed fashion.

4.3.1 Performance Analysis

After analyzing the performance of our video player, we concluded that the main bottleneck of the real-time video player is the network. If we just take the image from the scope and distribute it to different computers and require each computer to display only the part for which it is responsible, the performance is bad because we are replicating data and sending data to computers that they do not need it. By dividing the image into multiple portions and distributing only those parts, the replication of packets is not as great and we have an increase in the speed we can display the video. That is, by dividing the image we have an improvement in the display system final performance.

4.4 Immersive Medical Visualization

In an effort to facilitate the analysis and interpretation required of radiologists and surgeons, we have engineered two distributed display systems that work as unified environments where surgeons and radiologists can examine their medical data. Figure 4–6 shows an immersive, auto-calibrated display system we designed that enables flexible access to visualization of complex medical data-sets and images. As mentioned before, radiologists frequently need to study, analyze and compare image changes over time, and complete an in-depth analysis of a number of images side-by-side. Currently, some of the image modalities are viewed in the "light-box", while other images and 3D reconstruction are visualized in specialized computers with the capacity to show 3D volume reconstruction. Figure 4–6(bottom-left) shows a multi-wall, high-resolution display system used to display a number of traverse images of the data set in one wall, and at the same time visualize the 3D volume in the other wall. While visualizing the 3D model, and with the rapid access to the traverse, coronal and sagittal images that form the 3D volume, we believe that we can help surgeon to understand and reach conclusions faster than if they were using single LCD monitor as presented in figure 3–1

4.5 Maryland Deployment

During the summer of 2005 we had the opportunity to deploy our software and proposed next-generation of surgical environments in a research Operating Room at the University of Maryland Medical Center(UMMC). We used a cluster of computers located in an external room, a curved, rear-projection screen, and 6 casually aligned projectors. After the calibration process, we were able to create a seamless display system. The display environment and all the architecture deployed at the University of Maryland's SimCenter will be use to test the performance and benefits of a high-resolution, multi-context and hybrid display system for surgeons.



Figure 4-6: High-resolution medical visualization

4.5.1 Maryland SimCenter

Figure 4-7 shows the setup in that Operating Room that is part of the UMMC SimCenter.



Figure 4-7: Maryland SimCenter

CHAPTER 5

CONCLUSION

Through the research, experiments, and deployment of the software created as part of the next-generation of surgical environments, we believe that the work done for this thesis accomplished a major step in incorporating information technology, computer vision and image processing with the operating room.

The proposed architecture for the next-generation of surgical environments is a scalable design that can be used for a number of new experiments, tests, and evaluation about how technology can be used to assist surgeons during minimally invasive procedures.

Clearly, the distributed processing part of the proposed architecture is an essential module that takes care of all the computation required to enhance images, create 3D volumes and access preoperative data without introducing latency to the overall system.

High-resolution display systems are a key element to take advantage of more pixels, brightness and size to present surgeon with enough detail so they can truly take advantages to the image quality.

REFERENCES

- [1] Davis CJ, Filipi CJ., "A history of endoscopic surgery." *Principles of Laparoscopic Surgery: Basic and Advanced Techniques*, New York, NY; Springer-Verlag NY Inc; 1995: 3-2
- [2] Liem MS, van der Graaf Y, van Steensel CJ, et al. "Comparison of conventional anterior surgery and laparoscopic surgery for inguinal-hernia repair." *The New England Journal of Medicine*, no. 336, pp. 1541-1547, 1997
- [3] R.Raskar et. all H. Fuchs, M. Livingston., "Augmented reality visualization for laparoscopic surgery." *MICCAI*, pp. 934–943, Oct 1998
- [4] Xiaoli Zhang and Shahram Payandeh., "Application of visual tracking for robot-assisted laparoscopic surgery." *Journal of Robotics Systems*, 315–328, Vol. 19, No. 7, 2002
- [5] Y. F. Wang, D. Uecker, and Y. Wang, "Choreographed Scope Maneuvering in Robotically-Assisted Laparoscopy with Active Vision Guidance" *IEEE Workshop on Applications of Computer Vision* Saratoga, FL, pp. 187-192, Dec, 8-10, 1996
- [6] M. S. Kim, J. S. Heo, J. J. Lee "Real-time visual tracking for laparoscopic surgery" *International Journal of Human-Friendly Welfare Robotic Systems* Vol. 5, No. 1, 2004, pp. 2-9
- [7] Casals, A., Amat, J., Laporte, E. "Automatic guidance of an assistant robot in laparoscopic surgery" *IEEE International Conference on Robotics and Automation* Vol. 1, pp 895-900, 1996
- [8] Guo-Quing Wei, Arbter, K., Hirzinger, G. "Real-time visual servoing for laparoscopic surgery. Controlling robot motion with color image segmentation" *IEEE Eng. in Med. and Bio. Magazine* Vol 16, Issue 1, pp. 40-45, 1997
- [9] Jesus Caban and W. Brent Seales "Reconstruction and Enhancement in Monocular Laparoscopic Imagery" *Proceedings of Medicine Meets Virtual Reality 12* 37-39, January 2004
- [10] R. Y. Tsai "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lens" *IEEE Journal of Robotics and Automation* 323-344, Vol. RA-3, No. 4 1987
- [11] "The MathWorks" <http://www.mathworks.com/>

- [12] "Camera Calibration Toolbox for Matlab"
www.vision.caltech.edu/bouquetj/calib_doc/
- [13] Z. Zhang. "Camera Calibration with One-Dimensional Objects." *Microsoft Technical Report MSR-TR-2001-120*, August 2002.
- [14] D.C. Brown "Lens Distortion for Close-Range Photogrammetry" *Photometric Engineering*, 855-866, Vol. 37, No. 8, 1971.
- [15] T. Yamaguchi et al. "Camera Model and Calibration Procedure for Oblique-Viewing Endoscope" *LNCS MICCAI*, 2003.
- [16] Haneishi H, et al. "A new method for distortion correction of electronic endoscope images" *IEEE Transactions on Medical Imaging*, vol. 14, pp. 548-555, 1995.
- [17] Vijayan K, et al, "A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation" *IEEE Transactions on Medical Imaging*, vol. 18, pp. 345-354, 1995.
- [18] "Stryker Corporation." *http://www.stryker.com/*
- [19] "Vista medical technologies INC." *http://www.vistamt.com/*.
- [20] R. Raskar, M. Brown, R. Yang, W. Chen, H. Towles, B. Seales, and H. Fuchs "Multi projector displays using camera based registration", *Proceedings of IEEE Visualization*, 161-168, 1999
- [21] Raskar, R., van Baar, J., Chai, J.X. "A Low-Cost Projector Mosaic with Fast Registration." *Asian Conference on Computer Vision (ACCV)* 161-168, 2002
- [22] R. Surati "A Scalable Self-Calibrating Technology for Seamless Large-Scale Displays" *PhD thesis, Department of Electrical Engineering and Computer Science* Massachusetts Institute of Technology, 1999
- [23] Y. Chen, D. Clark, A. Finkelstein, T. Housel, and K. Li "Avoiding Seams on High-Resolution Multi-Projector Displays using an Un-calibrated Camera" *Technical Report, Dept. of Computer Science, Princeton University*, 2000
- [24] Fredo Durand, Julie Dorsey "Fast bilateral filtering for the display of high-dynamic-range images" *ACM Transactions on Graphics* Volume 21 Issue 3, July 2002
- [25] Sanjini Jayaraman, Chris North "A Radial Focus Context Visualization for Multi-Dimensional Functions" *IEEE Visualization*, Oct 27-Nov 1, 2002
- [26] North, C., and Shneiderman, B. "A Taxonomy of Multiple Window Coordinations" *Technical Report CS-TR-3854 University of Maryland, Computer Science*, 1997

- [27] Sean Ziegeler, Robert J. Moorhead, Paul J. Croft, Duanjun Lu "The MetVR Case Study: Meteorological Visualization in an Immersive Virtual Environment" *12th IEEE Visualization 2001 Conference*, 2001
- [28] Rudrajit Samanta, Thomas Funkhouser, Kai Li, Jaswinder Pal Singh "Hybrid Sort-First and Sort-Last Parallel Rendering with a Cluster of PCs" *SIGGRAPH/Eurographics Workshop on Graphics Hardware*, August 2000
- [29] Greg Humphreys, Mike Houston, Ren Ng, Randall Frank, Sean Ahern, Peter D. Kirchner, James T. Klosowski "Chromium: A Stream-Processing Framework for Interactive Rendering on Clusters" *SIGGRAPH*, 2002
- [30] Carl Mueller "The sort-first rendering architecture for high-performance graphics" *Proceedings of the 1995 symposium on Interactive 3D graphics*, 1995
- [31] Rudrajit Samanta, Thomas Funkhouser, Kai Li "Sort-last parallel rendering: Parallel rendering with k-way replication", *Proceedings of the IEEE 2001 symposium on parallel and large-data visualization and graphics*, 2001
- [32] C. Jaynes, S. Webb, M. Steele, M. Brown, and B. Seales "Dynamic Shadow Removal from Front-Projection Displays", *IEEE Visualization*, 2001
- [33] Michael S. Brown, W. Brent Seales "Incorporating Geometric Registration with PC-Cluster Rendering for Flexible Tiled Display" *International Journal of Image and Graphics*, 2003
- [34] W Matusik, H Pfister 3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes *SIGGRAPH*, 2004
- [35] Desmond M.Germans, Hans.J.W.Spoelder, Tom van der Schaaf and Henri E. Bal "Realizing a Stereo Tiled Display Using Commodity Components", *ASCI*, 2003
- [36] Brent Seales and Jesus Caban "Visualization Trends: Application in Laparoscopy" *Seminars in Laparoscopic Surgery*, 2003
- [37] Staffan Bjork "Hierarchical flip zooming: enabling parallel exploration of hierarchical visualizations" *Proceedings of the Working Conference on Advanced Visual Interfaces*, May 2000
- [38] R Raskar, J van Baar, JX Chai "A Low-Cost Projector Mosaic with Fast Registration" *Proceedings of Asian Conference on Computer Vision*, 2002
- [39] A Majumder, R Stevens "LAM Luminance Attenuation Map for Photometric Uniformity in Projection Based Displays" *Proceedings of ACM Virtual Reality and Software Technology*, 2002

- [40] A Majumder, Z He, H Towles, G Welch "Achieving color uniformity across multi-projector displays" *IEEE Visualization*, 2000
- [41] Majumder, Aditi and Greg Welch "Computer Graphics Optique: Optical Superposition of Projected Computer Graphics" *Fifth Immersive Projection Technology Workshop, in conjunction with the Seventh Eurographics Workshop on Virtual Environments*, May 2001
- [42] Yang, Ruigang, David Gotz, Justin Hensley, Herman Towles and Mike Brown "PixelFlex: A Reconfigurable Multi-Projector Display System" *IEEE Visualization 2001*, October 2001
- [43] Han Chen, Grant Wallace, Anoop Gupta, Kai Li, Thomas Funkhouser, and Perry Cook "Experiences with Scalability of Display Walls" *Seventh Annual Immersive Projection Technology Symposium*, March 2002
- [44] G Humphreys, M Eldridge, I Buck, G Stoll "WireGL: a scalable graphics system for clusters" *Computer Graphics Annual Conference*, 2001
- [45] Tan K, Kobler J, Feris R, Dietz P, Raskar R. "Shape-Enhanced Surgical Visualizations and Medical Illustrations with Multi-Flash Imaging" *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2004)*, September 2004
- [46] Deidra L. Donald, Nick Andreou, Jeffrey Abell, Robert J. Schreiber "The new design: the changing role of industrial engineers in the design process through the use of simulation" *Proceedings of the 31st conference on Winter simulation*, 1999
- [47] R. Raskar. "Immersive planar display using roughly aligned projector." *IEEE Virtual Reality*, 2000
- [48] Jesus Caban, W. Brent Seales, Adrian Park "Heterogeneous Displays for Surgery and Surgical Simulation" *Medicine Meets Virtual Reality 13*, 2005
- [49] "The Chromium Software Project", <http://chromium.sourceforge.net>
- [50] "Distributed Multihead X", <http://dmx.sourceforge.net>
- [51] "Open Source Computer Vision Library" sourceforge.net/projects/opencv/
- [52] "Cyviz - A new dimension to 3D technology" <http://www.cyviz.com>
- [53] "AGAVE (Access Grid Autostereo Virtual Environment)" <http://www.evl.uic.edu/cavern/agave/>
- [54] "GNU General Public License" www.gnu.org/copyleft/gpl.html
- [55] "NCSA Pixel Blaster" <http://brighton.ncsa.uiuc.edu/prajlich/wall/ppb.html>

- [56] "Barco" <http://www.barco.com>
- [57] "Ascension Technology Corporation" <http://www.ascension-tech.com/>
- [58] "REVEAL" <http://halsted.vis.uky.edu/index.htm>
- [59] "Fakespace Systems", <http://www.fakespacesystems.com>
- [60] "TeraRecon Inc.", <http://www.terarecon.com>
- [61] "VisionMax International Ltd.", <http://www.visionmaxint.com>
- [62] "SunnyBrook Technologies Inc.", <http://www.sunnybrooktech.com>
- [63] "Mercury Computer Systems", <http://www.tgs.com>
- [64] "AMIDE: a Medical Image Data Analysis Tool"
<http://www.amide.sourceforge.net>
- [65] "Silicon Graphics Inc." <http://www.sgi.com>
- [66] "IBM" <http://www.ibm.com>
- [67] "Tungsten Graphics" <http://www.tungstengraphics.com/>
- [68] "GraphStream" <tp://www.graphstream.com/>

VITA

Place of Birth:

Aguadilla, Puerto Rico

Education:

Bachelor of Science in Computer Science,
University of Puerto Rico, 1998

Work Experience:

Research Assistant (2002 - 2005)
Digital Media Networks Lab
University of Kentucky

Research Assistant (1998 - 2002)
High Performance Computing Facility
University of Puerto Rico

System Engineer-Coop (2000)
Silicon Graphics Inc.
Orlando, FL