AD________________

Award Number: DAMD17-03-2-0015

TITLE: REVEAL: Reconstruction, Enhancement, Visualization, and Ergonomic Assessment for Laparoscopic Surgery

PRINCIPAL INVESTIGATOR: W. Brent Seales

CONTRACTING ORGANIZATION: University of Kentucky Research Foundation
Lexington KY 40506-0057

REPORT DATE: February 2007

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



**1. REPORT DATE** *(DD-MM-YYYY)*
01-02-2007

**2. REPORT TYPE**
Annual

**3. DATES COVERED** *(From - To)*
06 Jan 06 – 05 Jan 07

**4. TITLE AND SUBTITLE**
Reconstruction, Enhancement, Visualization, and Ergonomic Assessment for Laparoscopic Surgery

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
DAMD17-03-2-0015

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
W. Brent Seales

E-Mail: seales@uky.edu

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

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

University of Kentucky Research Foundation
Lexington KY 40506-0057

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT:**
The goal of this work is to develop and test new technologies that will break down the barriers that block more surgeons from attaining and continuing to practice (without injury or pain) high levels of skill in minimally invasive surgery (MIS). This project will develop new technology by concentrating on three major research thrusts: Smart Image: the project will develop and evaluate new approaches for extracting, fusing, and presenting information cues from imagery and other data sources; Configurable Display: the project will develop new approaches for presenting existing data (video, CT data) and extracted cues (3D reconstruction, haptic cues, etc.) to the user within a flexible, configurable display environment; Ergonomic Assessment: the project will use existing technology and build new techniques as needed to acquire crucial ergonomic data relative to key factors of patient position, technology configuration, and instrument design.

**15. SUBJECT TERMS**
None provided

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** USAMRMC |
|---|---|---|---|---|---|
| **a. REPORT** U | **b. ABSTRACT** U | **c. THIS PAGE** U | UU | 28 | **19b. TELEPHONE NUMBER** *(include area code)* |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

## Table of Contents

**REVEAL: Reconstruction, Enhancement, Visualization, and Ergonomic Assessment for Laparoscopy**

**2006 Annual Report**

## *1. Introduction*

Information cues available in laparoscopy and other forms of minimally invasive surgery are impoverished relative to cues available in open surgery. Acquiring surgical skill in such an environment is extremely challenging. Even after mastery, continued practice can lead to problems for the surgeon as indicated by frequent incidence of pain and injury associated with laparoscopy. The long-term impact on the surgeon performing these procedures is largely unknown.

The goal of this work is to develop and test new technologies that will break down the barriers that block more surgeons from attaining and continuing to practice (without injury or pain) high levels of skill in MIS. This project will develop new technology by concentrating on three major research thrusts:

- **Smart Image**: the project will develop and evaluate new approaches for extracting, fusing, and presenting information cues from imagery and other data sources.
- **Configurable Display**: the project will develop new approaches for presenting existing data (video, CT data) and extracted cues (3D reconstruction, haptic cues, etc.) to the user within a flexible, configurable display environment
- **Ergonomic Assessment**: the project will use existing technology and build new techniques as needed to acquire crucial ergonomic data relative to key factors of patient position, technology configuration, and instrument design.

## *2. Major Accomplishments*

In this section we provide a functional view of major tasks accomplished during the 2006 project year. These include (1) support and upgrade of the REVEAL display system and tool suite in the University of Maryland Medical Center's Simulation Center, (2) stereo video display technology deployment, (3) stereo probe calibration benchmarks and support tools, (4) the production of research media, (5) baseline results from cognitive ergonomics experiments, and (6) continuing results from physical ergonomic experiments.

### 2.1 Support and Upgrade of REVEAL Display System at UMMC Sim Center

During the reporting period we concentrated our effort on the display system, making it more robust and improving the features we can offer with the goal of greater portability and flexibility. The primary technical improvements include

- Upgraded the compute cluster that controls the video rendering so that it now allows stand-alone net-boot of software and operating system from single head node. This allows the display code to be loaded dynamically at boot time from the head node, making the version management of the display cluster architecture very simple. This is now deployed and working in the REVEAL lab and in the Simulation Center at UMMC

- Improved the overall system latency by approximately 40 ms, which represents an improvement in performance (relative to end-to-end latency) of 25%.
- Improved throughput to achieve full frame rate and doubled the supported image resolution to 480 lines (from 240)
- Completed first implementation of stereo video application, which now runs at approximately 15 frames per second. Projectors are fully calibrated and filtered to support left/right polarized stereo with automatic calculation of left/right alignment
- Restructured video display application environment to a general and portable client/server model, which we call the smartServer, allowing future development with diverse devices and higher resolution
- Logged numerous bug fixes to calibration code, improving alpha mask generation, setting of default gamma values in the calibration process, and fixing settings to allow front, back, and ceiling-projected modes
- Augmented control interface so that non-expert researchers (e.g., physical and cognitive ergonomic researchers, who are using the system) can control the display smartServer and all of its various clients by adding a new dialog box packed with controls.

Given that this system is deployed and in use in the REVEAL lab and in the Simulation Center at the UMMC, we spent time studying its behavior and keeping it current as the development group produced the improvements above. Our analysis of the system has produced a paper for publication that measures the system performance, including the latencies of individual system components and the overall system latency. This paper, currently titled "Analysis of Display Architecture for Laparoscopy", will be submitted for peer review during the next review period (it is included in the appendix for reference).

We provided support to a new deployment of the VIBE display system in the Laboratory for Advanced Networking at the University of Kentucky. This test field deployment allowed us to tune the installation process and monitor how the current system can be set up from scratch, run and maintained in a generic production environment.
Finally, the technical team performed various administrative tasks to prepare for the final studies involving the display architecture. This included the development of specifications for an upgraded computer system capable of driving a 12-projector display system at full frame rate supporting two high-definition (HD) video streams. These specifications will be implemented in order to support full HD scopes in stereo. Other administrative tasks included cabling, network and operating system upgrades, and network analysis and tuning to improve throughput and to minimize packet loss.

## 2.2 Stereo Video Display Deployment

We continued development on casually-aligned passive (polarized) stereo displays. The system now robustly calibrates the left and right images from sets of projectors, aligns these sets through warping in software (or in hardware on the graphics cards), and runs two image video streams from a capture device in real time. This system is connected to our Vista stereo probe to provide a large scale real-time casually aligned stereo system. We have experimented with this system in the lab but have not yet deployed a version in Maryland for use in the protocol. The goal would be testing to confirm that the stereo cue, when available, provides substantial help in accomplishing benchmark tasks. The polarized system is portable, scalable, and the code is freely available from the REVEAL team.

## 2.3 Stereo Probe Calibration

A Stereoscopic Endoscope is an endoscope with two optical paths, either separate or shared, creating two images related to one another by a measurable disparity shift. Such an endoscope can be used to generate a stereoscopic view for a surgeon, as with the DaVinci robot in use today. In order to use such an endoscope for metric measurement of structures in the operative field, it is necessary to calibrate the dual optical paths according to a camera model. Once calibrated, it is possible to use stereo reconstruction in order to recover Euclidean metric measurements from the endoscopic images.

Bench calibration of stereoscopic endoscopes can provide a valuable way to make in-the-image instantaneous measurements from a single stereo pair with enough accuracy to save time in certain procedures where metric measurements are necessary for making decisions and recording anomalies. Errors in reconstruction are large enough that it warrants continued work on calibration methods and integration of second-order measurement equations (e.g., structure from motion, structured light) in order to narrow the error profile.

In support of the goal of building through-the-lens measurement tools for laparoscopes, we continued work during this period on calibration experimentation to quantify our ability to calibrate and measure with a stereo laparoscope. In particular, we have a complete bench calibration protocol for our scopes, with improvements in the lens distortion model (four-parameter model – two for radial and two for tangential), which yields re-projection errors of less than 0.25 pixels for camera intrinsic parameters. An example bench calibration result with calculated lens distortion appears in Fig. 1 below.

The stereo scope calibration gives our toolset the ability to perform through-the-lens, on-demand measurement within a working volume of 60 mm to sub-millimeter accuracy. As a result it is possible to measure anatomical features (e.g., defects, structures) within a predictable error model when using the stereo scope during training or surgery. We are working on the user interface issues of incorporating this measurement capability into the standard set of tools during scope use, and in structuring a set of tasks around the use of through-the-scope measurement in order to determine how this tool can affect efficiency. In particular we are interested in the time it takes to complete a hernia repair given that the scope-based measurement can replace the standard practice of viewing a measuring tape to estimate patch repair size.

In support of the technology needed to acquire data for the cognitive studies, we implemented the video effects including image rotation, change in scale, and simulation of blood masks and focus distractions. Additionally we performed a number of software and hardware related improvements to make the experimental environment more robust, including the removal of hard-coded data acquisition paths, a cleanup of the CVS (revision control) code base, and the deployment of a set of de-interlacing methods to improve the resolution and performance of the real-time video display environment. These enhancements to the experimental environment run

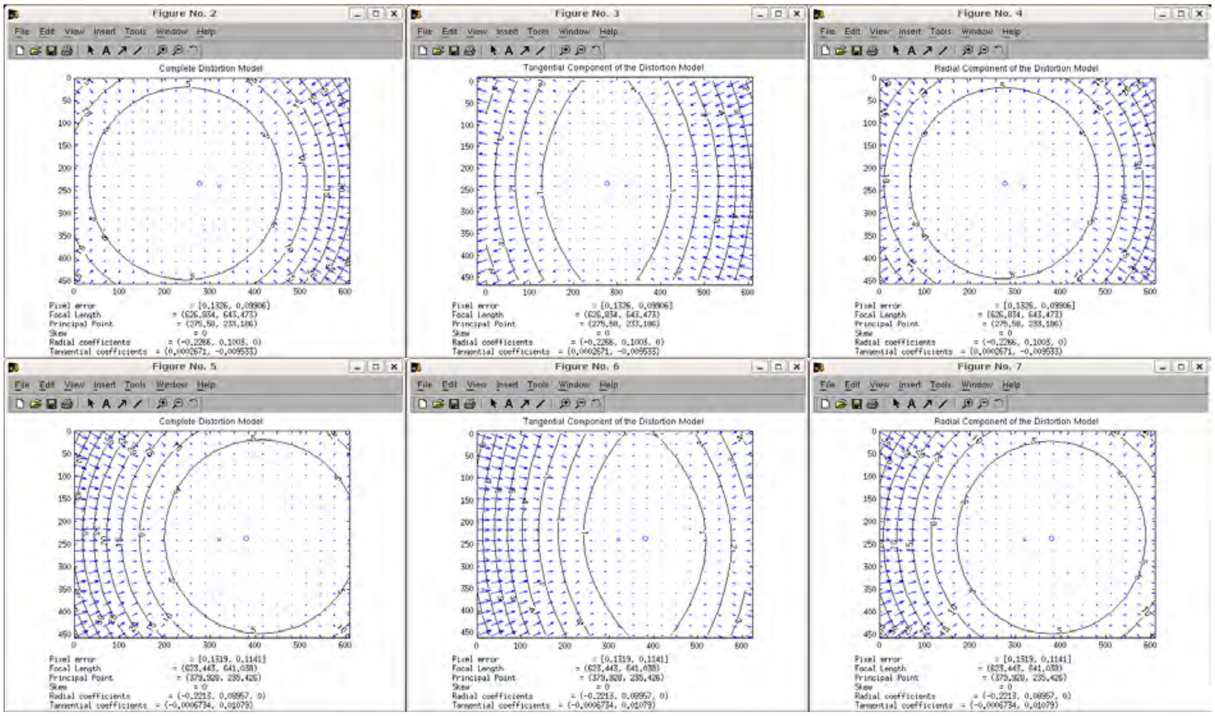


**Fig. 1**: Composite, tangential, and radial components of the distortion model for the left (top) camera and the right (bottom)

over the general simulation system in real time and can be toggled through a command interface as necessary.

## 2.4 Production of Research Media

We continued to pursue our goal of media production and documentation for all aspects of the project. In order to promote the research results and methods we produced a half-hour piece for the Research Channel, which will be delivered and reviewed for airing during spring 2007. We attended the Research Channel's annual meeting in Chicago, where we made a presentation detailing our methods and results from the media efforts within the REVEAL project.


**Fig. 2:** Screen shots from media produced to air on the Research Channel about the goals and accomplishments of the

In support of the documentation of REVEAL research, we shot video and interview footage at UMMC to record work being done there by Gyusung Lee in the Simulation Center related to physical ergonomic assessment. This provides documentation of process and goals, results, interviews with principal researchers, and descriptions of equipment setup and renovations necessary to support the research.

We continued to document the work in cognitive ergonomics being conducted by Melody Carswell at the University of Kentucky. We documented the controlled experiments being done in the REVEAL laboratory, including interviews with students, who explain in detail the experimental design. We have continued to experiment with stereo video acquisition for our research with stereo displays, including the capture of stereo still images taken with the REVEAL camera and a special lens.

We attended the "Operating Room of the Future" conference in Baltimore, MD, and recorded talks and interviews by all REVEAL team members, including Adrian Park, W. Brent Seales, Melody Carswell, Gyusung Lee, and Ivan George. We also interviewed invited speakers and TATRC personnel (Gerald Moses and Amy Nyswaner).

During this review period we have edited footage from the ORF meeting and other shoots in order to produce the piece to be aired on the Research Channel (see Fig. 2 for example screen shots from this piece).

## 2.5 Baseline Results from Cognitive Ergonomics Experiments

During 2006, we have made progress toward our goal of establishing a set of cognitive metrics that are suitable for assessments of our new visualization tools both in the lab and, eventually, directly in the surgical context. In 2005, our efforts were directed primarily at finding an appropriate secondary task measure of mental workload and comparing it to a well-established subjective measure of workload in terms of their respective sensitivities to a gross manipulation of task demand. In 2006, we have extended this work to the evaluation of more subtle changes in task demands in situations more similar to the training environment for laparoscopic surgeons. In addition, we have explored the use of a performance-based measure of situation awareness suitable for laparoscopic training environments, and representative of the type of measurement that could be used during high fidelity simulations or actual surgeries. We also continued to collect baseline data using a short, subjective measure of stress and an additional workload measure that provides a microanalysis of workload components (i.e., types of demand).

Turning first to our development of a secondary task workload measure, we followed up our initial encouraging work from 2005 using a time estimation procedure by modifying it in ways suggested by our earlier data, and by using the revised measure to 1) track the effects of practice on workload, and 2) determine its sensitivity to the effects of two ecologically valid distracters ( i.e., changes in image orientation and clarity).

We have completed intensive testing of 19 participants, each for approximately five hours (across two sessions). Data were collected using a simulation task environment supported by the REVEAL team which includes surgical instruments, imagery generated from laparoscopes, and a large-scale, tiled display. We found that even in the earliest phases of training, even when there were no reliable indicators of performance increments, reliable decreases in workload were detectable both in our subjective measure (NASA-TLX) and in our secondary task measure. Although changes in subjective workload could be attributable to participants' expectations, it is more difficult to explain changes in secondary task performance because participants are unaware of how their performance is predicted to change with decreases in task demand. In addition, the secondary task measure showed a greater tendency to correlate with evidence of participants having good situation awareness, yet another converging piece of evidence that time estimation is a particularly useful metric.

With respect to situation awareness, we have moved away from pursuing more traditional but disruptive measures such as screen "blankings" followed by probe questions. Instead, we have pursued the development of a simple Global Implicit Measure (GIM) technique which assesses the extent to which participants notice and use information in the immediate task environment to improve their performance, even though they have received no warning about the potential presence or use of this information. Specifically, participants were asked to position foam rings on a pegboard, with the goal of placing as many as possible during a series of two minute trials. The foam rings differed from one another in color, but in every other respect were identical. After several trials, the original rings were exchanged for rings that differed in terms of the diameters of their interior holes. Now, the color of the ring indicated the size of the hole, and when participants noticed this correlation, they could begin selectively retrieving the "easiest" rings to thread, boosting their overall scores. Although the change seemed obvious to onlookers, many of the participants never detected or made use of the color cue. The use of contextual information to aid positioning strategies has been emphasized in interviews with surgeons. Thus, although vastly simplified, this global implicit measure serves as a demonstration of how such measures, if made more realistic, could aid in both assessing and training the observational and inferential skills needed to establish and maintain situation awareness. The demonstration also once again indicates that even the simplest positioning tasks, when performed using the indirect views of laparoscopic surgery, create cognitive demands that, in this case, make it difficult to perform the "secondary task" of noticing and using additional task-relevant information.

Finally, with respect to our newer measures of stress and subjective workload, we have collected data that provide evidence that both add non-redundant information to our battery of cognitive measures. Our first look at results from the SSSQ (Short Stress State Questionnaire) indicates that the measure reflects different types of stress responses to different manipulations of task difficulty. Although the different manipulations may affect workload in similar ways, they affect our participants' emotional reactions to workload in different ways. For example, some manipulations (such as being given less practice) decrease engagement and increase worry, while changes in precision demands may influence general negative affect (e.g., irritation, depression). Thus, training and redesign interventions in surgery may have differing qualitative effects on the surgeon's response to his or her task. There are also likely to be large individual differences in these responses, and different individual stress profiles could be used to suggest individually-tailored training interventions, for example.

Just as the SSSQ makes several distinctions in the emotional reactions to changes in task difficulty, the Multiple Resource Questionnaire (MRQ) provides a finer parsing of the types of workload that are created by different task manipulations. Rather than measuring the gross amount of cognitive capacity used by a task, the measure attempts to determine which cognitive resources are specifically being tapped. This information is particularly useful in designing human-technology interaction modes that create minimal disruptions with the surgeon's primary tasks (e.g., do not compete for the same mental resources). Analysis and processing of the MRQ data are still in progress.

## 2.6 Ergonomics Assessment for Laparoscopy

**Equipment**

Nexus™ software (Fig. 3) that provides improved real-time processing of data collection and presentation has been added to our ViconPeak™ (Lake Forest, CA) motion capture system, which consists of twelve cameras, two AMTI™ (Advanced Mechanical Technology, Inc., Watertown, MA) force plates, and 16 channel Delsys™ (Boston, MA) electromyography systems. A common difficulty that we had been encountering during research was that a marker could be missed if anything, such as furniture, blocked camera view is located between markers and cameras. With a previous ViconPeak software, Workstation™, this problem would be detected only after data processing, and most trajectories of missing markers are not easily repairable. Using real-time marker display and stick diagram monitoring, we have been able to minimize data loss due to missing markers and thus improve the accurate capture of experimental data. Also the Nexus™ software provides us with more customizing options allowing us to capture more details. Workstation program, we previously used, was capable of handling only one DV input to which we added an external image capture board to record both the surgeons' body movement and the endoscope image. Because Nexus™ can handle multiple

DV inputs and synchronize movie data with motion capture data, better assessment is now possible.

Also, a ProMIS$^{TM}$ (Boston, MA) surgical simulator has been recently added to our experimental setup. (Fig. 4) This system permits the use of real laparoscopic instruments and uses camera-based tracking technology to measure instrument movements, which can then be analyzed in terms of performance time, path length, and smoothness.

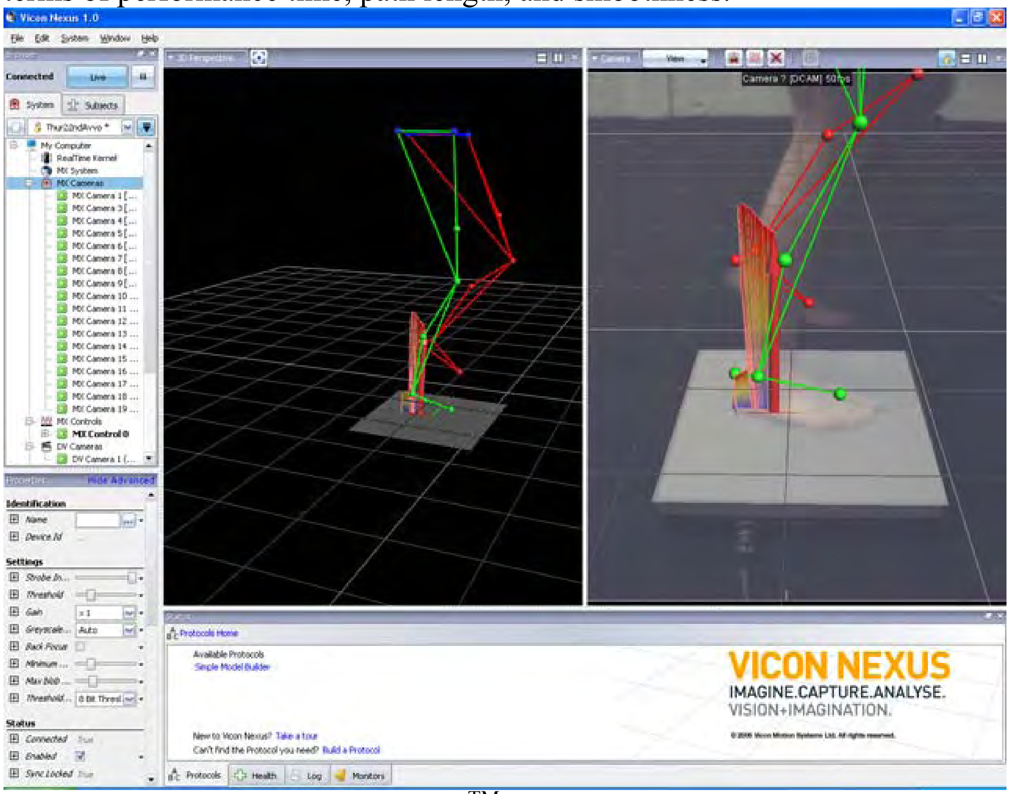



**Fig. 3. Screen shot of Vicon Nexus$^{TM}$ program.**

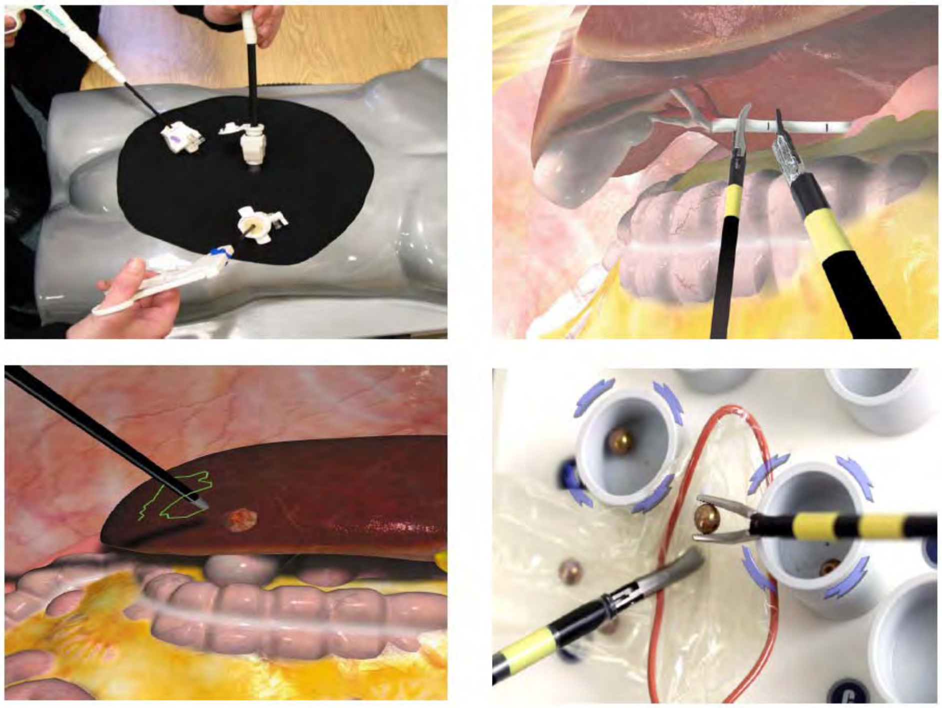


**Fig. 4. ProMIS™ Surgical Simulator**

**Research Outcomes**

As specified at the end of this report, we published one article in the peer-reviewed journal "Surgical Endoscopy" and two abstracts presented at Society for Neuroscience and Medicine Meets Virtual Reality (MMVR) respectively.

Additionally, the following three abstracts have been accepted for presentation:

- Study 1: Development of a Novel Tool to More Precisely Analyze Postural Stability of Laparoscopic Surgeons.[1]

Background: The physical difficulties experienced by surgeons performing MIS procedures are being given extensive attention by ergonomic researchers. We take postural stability, which has been insufficiently addressed, as one of our prime focuses. The few studies already existing in this area have used the Center of Pressure (COP) excursion range alone. Using COP, we previously correlated postural stability to instrument type and task difficulty in addition to subject skill level. This study extends the investigative scope, particularly in terms of skill level, by broadening analysis to include Center of Mass (COM) and what we uniquely term Postural Stability Demand (PSD).

Methods: Six subjects with different levels of surgical experience were recruited to complete 3 FLS tasks – circle-cutting, endo-loop placement, and pegboard transfer. Participants performed each task while standing on two force plates while a motion capture system recorded body movements. COP—at the bottom of the feet—is the point where ground reaction force is located. COM is the point at which the mass of the body is concentrated. Principal Component Analysis was used to create an ellipse covering 95% of COP and COM excursions for the calculation of COP and COM sway area. PSD we characterize as the mean distance between COP and COM locations in anterior-posterior (A-P) or medial-lateral (M-L) directions. Correlations between these postural parameters and performance time were analyzed.

*Results:* During circle-cutting, less skilled participants required longer times to complete the task and showed larger sway areas both in COM and COP (r=.858, p<.05; r=.779, p=.06). During endo-loop placement, sway areas of COM and COP were smaller for less skilled participants (r=-.899, p<.05; r=-.890, p<.05). These results indicate postural control differences between more and less experienced lap surgeons. No significant sway area correlation was found during pegboard transfer. Importantly, during all three tasks PSD in the A-P direction was strongly correlated with performance time (r=.744, p<.05; r=.913, p<.05; r=.772, p<.05). This indicates that less skilled participants experienced increased postural demand equated with higher postural instability.
Conclusions: This study demonstrated that variance in postural adjustments could be correlated to skill level and individual task. Strong correlation between PSD and performance time shows potential as an indirect predictive measure of surgical skill levels. Combining COM, COP, and PSD postural data results in a more robust analytical tool for identifying postural adjustments with skill level.

- Study 2: Assessment of table height change with laparoscopic instrument change[2]

Background: Surgeons seldom change the set height of an operating table once beginning a case and never do so to accommodate difference in instrument handles. Previous studies sought to determine optimal table height without taking into account the influence of different instrument handles.
Methods: We gave surgeons different styles of instrument handles and free range to choose optimal table height, based on comfort. Board-eligible, Board-certified general surgeons were recruited to complete two FLS tasks: peg board transfer (task 1) as well as intracorporeal suturing and knot tying (task 2). All tasks were conducted on a training stand with adjustable operating table and monitor height (Stryker). Subjects for task 1 were given two disposable pistol grip (PG) dissectors (USSC) and two inline (IL) needle drivers (Ethicon) for task 2. Nineteen reflective markers were placed on each subject's upper body, and 4 markers were placed on each instrument. A motion capture system (Vicon) used these markers to calculate upper-body joint angles and instrument shaft angles. For both PG and IL instruments, the table height was adjusted until maximum comfort was achieved. Kinematic measurements were made while instrument tips were in the center of the operative field.
Results: When using PG instruments, optimal table height averaged 98.1cm. When using IL instruments, a significant change was found as the table height lowered by 6.4cm to average of 91.7cm (p<0.005). Multiple changes in joint kinematics were observed when surgeons changed to IL instruments. Notable changes were in shoulder and wrist joint excursions while there was no significant change in elbow angle. With the IL instruments, elevation angle decreased from 45 to 33 degrees (p<0.005).
Conclusions: Optimal table height differs when surgeons work with PG versus IL instruments. A table height change based on instrument change faces limitations, such as drape and stand rearrangement and time consumption. Given such difficulties, ergonomic factors warrant further analysis to determine if a standard optimal table height for different instrument handles exists or if an ergonomic redesign of handles is warranted. Additionally, our data suggests that wrist position in addition to elbow position significantly impacts surgeon comfort and optimal table height. Further investigation will be conducted as part of our comprehensive research, including imaging, display, surgical ergonomics using whole body analysis, and human factors.

- Study 3: Subtask analysis of joint angles to characterizing surgical movement[3]

Introduction: Joint angle analysis using variables calculated from task beginning to end is the primary means in surgical ergonomics for understanding joint control. These traditional variables, e.g. range of motion, provide general joint movement information; however, they cannot differentiate specific patterns required to achieve certain goals of laparoscopic tasks. We propose a novel method of data analysis that extends our previous findings about characteristic joint kinematics determined from mean joint angle. [4] Here we divide a laparoscopic task into functional subtasks and analyze joint movement within the subtask's time frame to extract characteristic movement patterns associated with particular surgical maneuvers.
Methods: Nine right-handed, experienced laparoscopic surgeons were recruited to perform the standard FLS pegboard transfer task. A motion analysis system captured participants' upper body movements while simultaneously recording endoscopic images that allowed task performance to be partitioned. For each subject, left- and right-side data were captured in three rotations—flexion/extension, abduction/adduction, and internal/external—at three joints—shoulder, elbow, and wrist. The transfer of a single disk was divided into four subtasks: pickup, medial transfer (side to middle), lateral transfer (middle to side before dropping), and the actual drop. Joint angles were analyzed within the time windows of each subtask. Data from a representing subject suggests a possible rubric.
Results: Each subtask was composed of a unique set of joint kinematics (Table 1). These joint controls were relatively consistent through six repetitions. While left-right comparison in three subtasks showed no significant control strategy difference, during the pickup subtask left side movement involved fewer joints, well known as joint freeze on the non-dominant side.
Conclusions: Our study showed that detailed characteristic movement patterns, not fully depictive with traditional analysis, can be extracted when a laparoscopic task is partitioned into functional subtasks. This new approach will enable our investigation of intra- and inter-subject variability of joint kinematics as part of the development of standard joint control strategy matrices for optimal surgical performance.

Table 1. Significant joint movements observed in each subtask

| **Significant joint movements observed in each subtask** | | | | |
|---|---|---|---|---|
| **Subtask** | **Sides** | **Shoulder** | **Elbow** | **Wrist** |
| Pickup | Left | extended | — | extended & radial deviation |
| | Right | extended | extended | extended, radial deviation & supinated |
| Medial transfer | Both | flexed, abducted & externally rotated | extended | flexed, radial deviation & supinated |
| Lateral transfer | Both | extended & adducted | flexed | extended & ulnar deviation |
| Drop | Both | flexed | — | flexed & ulnar deviation |

## *3. Project Milestones 2006*

The 2006 project plan led us to near completion of visualization and display technology development, moving the technical component of the project toward a support role for remaining ergonomic assessment activities at the University of Maryland medical center. We have been granted a no-cost extension to complete the ergonomic assessment, with the technical team providing support for the equipment as the protocol is executed. The accomplished milestones and our progress in reaching each for the year are assessed in the sections that follow.

## Primary Milestones: Visualization Technology

These milestones will drive research by the IT team at the University of Kentucky:

a. Deploy and support ergonomic assessment environment based on stereo scope acquisition, flexibly hybrid (stereo/mono) display configurations, and integrated smart image cue support (months 1-3)

b. Design and test protocol support for integrated device standard within distributed architecture (months 1-5)

c. Gather requirements for technology review, re-design and re-implementation to support evolving surgical requirements as a result of baseline studies (months 3-6)

d. Re-design display position, resolution, layout and hybrid features for surgical simulation setting based on requirements (months 6-12)

The display system has been redesigned and upgraded in order to support more projectors per machine using multiple video cards within a single computer chassis (4 video cards per computer, each card capable of supporting 4 projectors). This upgrade allows an immersive distributed display system to scale to 80 projectors with just 5 machines (each machine supporting 16 projectors) at a cost of just $5,000 per machine. This new architecture has been specified, tested and deployed in the lab, with benchmarks indicating that it is capable of support HD scopes at frame rate with latency below 120ms.

As detailed in the technical sections above, we have made substantial progress toward completing each milestone, with the exception of milestone (b). We have put off work on protocol support for standardization in favor of supporting and executing the cognitive and physical ergonomic studies, which have required most of our attention and resources. We intend to address these issues in the follow-on STITCH project, currently underway.

## Primary Milestones: Ergonomic Assessment Group

These milestones will drive research by the Ergonomic Assessment group led by Adrian Park, M.D. and funded via subcontract to the University of Maryland:

a. Assist in requirements analysis and technology review based on prior baseline studies (months 1-3)
b. Conduct human skills pre-study in simulator environment with re-designed technology components (months 2-8)
c. Upgrade hardware/software environments based on revisions from IT group (month 9)
d. Assess and summarize all study results to date and report key findings (months 9-10)
e. Plan final study sequence based on technology re-design and summarized study results (months 10-12)

The ergonomic studies reported above (Section 2.6) show significant progress in completing these milestones.

## Project Year 3 Deliverables

The project has made available all source code detailed for the technical environments and algorithms above, including technical specifications for hardware and configuration information where necessary. In addition, we have published papers and made presentations as detailed in the reference list. One pre-publication paper is included in the appendix for reference.

## Appendix A: Project Personnel

| Name | Role | Location | 2006 FTE |
|---|---|---|---|
| W. Brent Seales, PhD | Principal Investigator | UK College of Engineering | 30% |
| Adrian Park, MD | Co-Principal Investigator | UM School of Medicine | 10% |
| Steve Bailey | Media Specialist | UK Department of Computer Science | 100% |
| Tsegay Baraki | Administrative Support (Budget, Reporting) | UMMS Department of General Surgery | 40% |
| C. Melody Carswell | Senior Researcher | UK Department of Psychology | 25% |
| Duncan Clarke, PhD | Technical Project Lead | Fremont Associates, LLC | 60% |
| Ryan Davis | Student Programmer | UK Center for Visualization | 25% |
| Praveen Devabhaktuni, MS | Program Systems Analyst | UK Center for Visualization | 25% |
| Matt Field | Program Systems Analyst | UK Center for Visualization | 50% |
| Ivan George | Technical Support | UMMS Department of General Surgery | 20% |
| Kimberly Hall | Administrative Support (Budget, Clerical) | UK Center for Visualization | 20% |
| Stephen Kavic, MD | Senior Researcher | UM School of Medicine | 20% |
| George Landon | Research Assistant | UK Center for Visualization | 50% |
| Gyusung Lee, PhD | Senior Researcher | UM School of Medicine | 50% |
| C. Andy Martin, MS | Program Systems Analyst | UK Center for Visualization | 50% |
| Linda Rice, RN, CCRC | Administrative Support (Research Protocols) | UK Medical Center | 20% |
| Ross Segan, MD | Senior Researcher | UM School of Medicine | 10% |
| Robert Shapiro, PhD | Senior Researcher | UK Department of Kinesiology and Health | 10% |
| Donald Witzke, PhD | Senior Researcher | UK Department of Pathology | 5% |

## Appendix B: Laboratory Facilities

### B.1 UK Software Development Laboratory

- Project Staff Office
  - Location: 1 Quality Street, 801 Kentucky Utilities Building
  - Purpose: Working environment for day-to-day activities of software developers.
  - Equipment:
    - Two developer workstations
    - Media workstation
    - Twin-processor Dell PowerEdge server with DLT tape drive
    - Cisco firewall/router

- High-Performance Multi-Projector Display Laboratory
  - Location: 871 KU Building
  - Purpose: Test environment for MIS video image processing techniques and large-scale projected displays.
  - Equipment:
    - RackSaver 22 Processor cluster computer with gigabit backplane
    - Dell workstation
    - Gigabit network switch
    - 12 DLP projectors with overhead mounts
    - Custom heat- and vibration-tolerant filters w/ mounts for stereo projection
    - 7.5’ x 10’ back-projected polarity-preserving screen with mounting frame
    - General purpose Canon video camera
    - Stryker trainer stand with auxiliary high-definition LCD display
    - 2 Stryker 888 high resolution cameras, controllers and lens probes
    - Stryker light source
    - Viking Systems stereo camera probe, controller and light source
    - Assorted MIS surgical instruments

- Project Office
  - Location: 883 KU Building
  - Purpose: Working environment for project management and small team meetings.
  - Equipment: One general purpose computer.

- Web site
  - Location: http://halsted.vis.uky.edu
  - Purpose: Provide general overview of project activities, distribute project documents and software, and serve as repository for project images (still and video).

### B.2 UMMC Simulation Center

- High-Performance Ergonomic Assessment Laboratory (within Sim Center)
  - Location: University of Maryland Medical Center, Baltimore Maryland

- Purpose: Test environment for ergonomic assessment of MIS tasks and assessment of large-scale projected displays.
- Equipment:
  - RackSaver 22 Processor cluster computer with gigabit backplane
  - Dell workstation
  - Gigabit network switch
  - 6 DLP projectors with overhead mounts
  - 6 x 8’ back-projected polarity-preserving screen with mounting frame
  - General purpose Canon video camera
  - Stryker trainer stand with auxiliary high-definition LCD display
  - Multiplicity of  Stryker equipment, including HD cameras, controllers, lens probes, and light sources
  - Assorted MIS surgical instruments and training stands
  - Vicon 12-camera beacon tracking system
  - Nexus™ software
  - two AMTI™ (Advanced Mechanical Technology, Inc., Watertown, MA) force plates
  - 16 channel Delsys™ (Boston, MA) electromyography systems.

## *Appendix C: Publications (2006)*


1. Lee G and Park AE. Development of a novel tool to more precisely analyze postural stability of laparoscopic surgeons, abstract accepted for oral presentation, SAGES 2007 annual conference
2. Lee G, Dexter DJ, Lee TH, Roth JS, Turner P, Kavic SM, Park AE. Assessment of table height change with laparoscopic instrument change, abstract accepted for poster, SAGES 2007 annual conference
3. Lee G, Dexter DJ, Lee TH, Park AE. Subtask analysis of joint angles is key to characterizing surgical movement, abstract accepted for poster, DDW 2007 annual conference
4. Carswell, C.M., Lio, C., Seales, W.B. and Clarke, D. (in submission). Situation awaareness during the performance of laparoscopic training tasks. Proceedings of the Human Factors and Ergonomics Society.
5. Carswell, C.M., Clarke, D., Lio, C., Kurs, Y., and Seales, B. (2007). Measuring subjective stress profiles during surgical skills training. MMVR, Feb. 2007, Long Beach, CA.
6. Lio, C. Carswell, C.M., Bailey, K., Clarke, D., Seales, B., and Payton, M. (2006). Time estimation as a measure of mental workload during the training of laparoscopic skill. *Proceedings of the 50$^{th}$ Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA: The Human Factors and Ergonomics Society.
7. Lee G, Kavic SM, George IM and Park AE (2007) Postural instability does not necessarily correlate to poor performance: Case in point, Surg Endosc. 21:471-474
8. Lee G, Kavic SM, Shapiro R, George IM, Park AE. Analysis of systematic changes in posture and joint kinematics demonstrated by minimally invasive surgeons. Society for Neuroscience, Atlanta, GA, October 14-18, 2006
9. Lee G, Kavic SM, George IM, Park AE. MIS Surgical Ergonomics: Future Trends. Medicine Meets Virtual Reality (MMVR), Long Beach, CA, February 6-9, 2007
10. C. Andy Martin, Duncan Clarke, C. Melody Carswell, Adrian Park, and W. Brent Seales. Analysis of Immersive Display Architecture for Laparoscopy, in preparation, to be submitted to IEEE Trans. On Information Technology in Biomedicine.
11. W. Brent Seales, Duncan Clarke, George Landon, Matthew Field, and Adrian Park, MD. Evaluation of Stereo-Endoscopy as a 3-D Measurement Tool, in preparation, to be submitted to IEEE Trans on Medical Imaging.

# Analysis of Immersive Display Architecture for Laparoscopy

C. Andy Martin, *Member, IEEE,* Duncan Clarke, C. Melody Carswell, Adrian Park, and W. Brent Seales

***Abstract*— DRAFT** `Revision: 1.9` **Construction of a scalable, flexible, information-rich laparoscopic surgical environment is being enabled by advances in general-purpose computing power and consumer video projection devices. The accepted practice of laparoscopic environments is a tightly coupled camera/monitor system. We have constructed a system in which the input sensors (such as cameras) are loosely coupled with the outputs (such as displays) by inserting a low-latency, high-powered computing platform between the traditionally tightly-coupled inputs and outputs of the laparoscopic surgical environment. Furthermore, the inputs/outputs of the system are easily reconfigured. For instance, the display architecture utilizes** *uses?* **casually-aligned projector arrays registered and calibrated using automated computer vision techniques. The architecture of this system has been deployed in the Surgical Simulation and Technology Center at the University of Maryland Medical Center, and its application to the laparoscopic environment is analyzed based upon its performance.**

***Index Terms*— laparoscopic procedures, immersive environments, casually-aligned projectors, technology infrastructure**

## I. Introduction

THE current laparoscopic surgical environment consists of limited, disparate data sources which must be manually integrated by the surgical team. The laparoscopic surgeon uses a constrained, tightly-coupled camera/monitor system for visual feedback as he manipulates surgical tools [1]. The technical limitations of the current camera/monitor system create an information bottleneck and offer less visual information to the surgeon than traditional open surgery. Also, there are monitors displaying real-time sensory data such as pulse and heartbeat and other sources of visual data such as perioperative 2D and 3D imagery (e.g., CAT scans, MRIs, X-rays) [2]. These difficulties are compounded by the physically complex and demanding laparoscopic operative procedures. These factors cause minimally invasive surgery to be cumbersome to the surgeon and require an exacting level of skill [3], [4]. It would be helpful to the surgical team to remove the information bottleneck present in current input/output surgical systems (specifically the laparoscopic camera/monitor system) while providing as much data integration as possible by using a consolidated and unified immersive platform.

Manuscript received January 1, 2007; revised January 30, 2007. This work was supported by project REVEAL, funded by DOD TATRC grant DAMD17-03-2-0015.

C. A. Martin, D. Clarke, C. M. Carswell, and W. B. Seales are with the University of Kentucky Center for Visualization and Virtual Environments. Email: andy.martin@vis.uky.edu, dclarke@freemontassociates.com, cmcars00@uky.edu, seales@uky.edu

A. Park is director of the Surgical Simulation and Technology Center at the University of Maryland Medical Center. Email: apark@smail.umaryland.edu

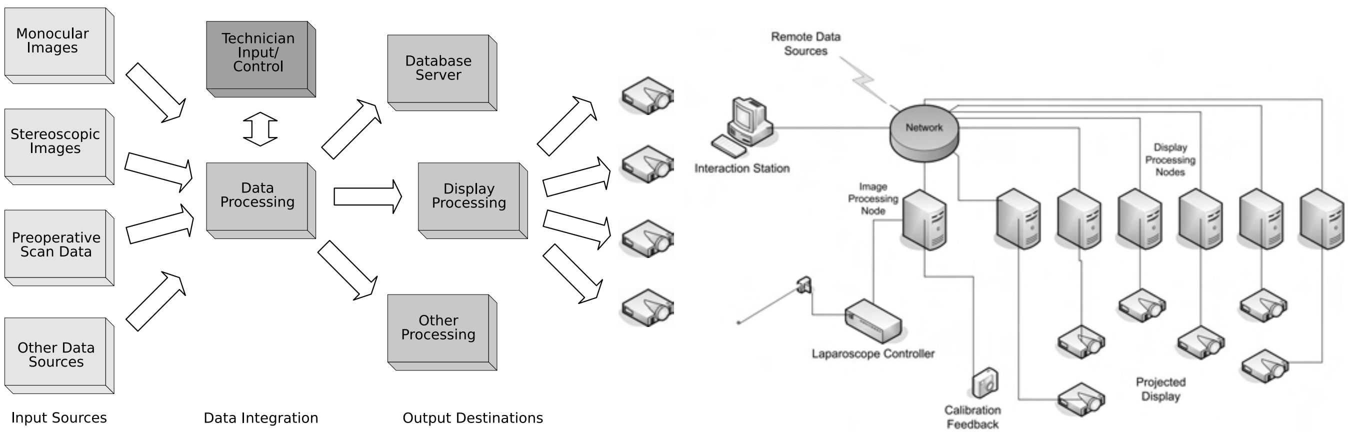
(a) Current Surgical Trainer (b) Immersive Display Trainer

Fig. 1. Views of (a) the current laparoscopic training environment and (b) an immersive display training environment

Because of these limitations, there is ongoing research into developing what has been termed "the OR of the future" [5]–[7]. The areas of research include plug-and-play "smart instruments," augmenting image-guided techniques, intelligent display and storage of data, and data processing. The Center for Visualization and Virtual Environments at The University of Kentucky and The University of Maryland Medical Center have been collaborating on project REVEAL: Reconstruction, Enhancement, Visualization and Ergonomic Assessment for Laparoscopy, which builds on the OR of the future concept by combining large displays, visualization tools, image enhancement and ergonomic analysis for the laparoscopic environment [8]. Specifically, the REVEAL team has created an architecture which is designed to have a high enough bandwidth to unlock future high-density data sources such as high-definition video and large 3D datasets, and is able to integrate disparate data sources and display them on a scalable display infrastructure.

A key component of the REVEAL project is the use of large, scalable, highly-configurable display technologies. The difference in usable display area achievable can be seen by comparing the current surgical training environment to an immersive display surgical training environment (Fig. 1). The benefits of a large working display area for performing virtual-reality and spatial task performance is well documented. [9]–[11]. Also, there is a large body of research documenting the technical aspects of such casually aligned projector systems [12]–[16]. However there is little published documentation of the actual implementation and performance of such systems in a production environment (such as the OR) and analysis of the system's performance implications on various applications including surgical tasks.

In this paper we analyze the performance impacts of an implementation of such a scalable, information-rich display system in the Surgical Simulation and Technology Center at




(a) System Architecture (b) System Infrastructure and Hardware

Fig. 2. REVEAL System Diagrams

University of Maryland Medical Center, specifically using the architecture outlined by Seales and Clarke for project REVEAL as a starting point[8]. Our goal is to describe the system as implemented, measure its performance characteristics, such as overall latency and throughput, and to analyze the impact of the results on its usefulness in various applications, focusing on the minimally invasive surgical setting.

## II. System Architecture

The system architecture is based on the architecture described by Seales and Clarke [8]. The architecture consists of data (including image data) processed by a data processor and controlled through a real-time software interface by a technician. The data is then integrated according to the technician's commands and processed for archival, display on the projector array, or other output (Fig. 2). Data includes preoperative image data, monocular or stereoscopic video images captured from laparoscopes, and other data sources. When outputting aggregate visual data, a seamless display is created from scalable, loosely-aligned projectors based on past research efforts [13]–[16].

### A. Infrastructure

An outline of the required hardware infrastructure for deploying the REVEAL system can be found in Fig. 2. The heart of the system is the data processing node. This node must have the hardware capabilities to process real-time video feeds from the laparoscope, calibration camera and other data input sources, process the data and then send them via network technologies to the various output processing nodes and devices. The output devices include a data archival server, a display processing node, and other output device processors. The display processing node is where the raw imagery output from the data processing node will be further processed to provide a seamless and integrated display to the surgical team via an array of casually-aligned projectors. In our implementation, the data processing node is a hand-assembled server computer which has some necessary features such as a PCIe bus architecture, a fast AMD Athlon 64-bit processor, two gigabit network interfaces, and a video capture card. PCIe is a necessary technology for the data processing computer because it is a point-to-point serial bus topology that can be scaled in parallel to increase throughput for a single expansion slot. Since it is a point-to-point serial technology, the bus does not suffer from the sharing of bus resources as its predecessor PCI. Therefore, the various data sources and the network interfaces do not interfere with one another on the bus.

The data processing node is connected via a gigabit switched IP network to a cluster of display processing nodes (also called render nodes as they render the images into a final processed scene). These render nodes are housed in a server rack outside of the OR to conserve valuable OR space. Each node has a gigabit network interface, a 3.2 GHz Intel Xeon processor and a high-end NVidia graphics card. The graphics card is attached to a Dell DLP multimedia projector with a native resolution of 1024x768. The projectors are mounted on standard Dell ceiling mounts on a truss structure in the OR and pointed at a rear-projected, large-area screen. The screen and the output of the render nodes via the projectors prior to calibration of the projector array are shown in Fig. 3.

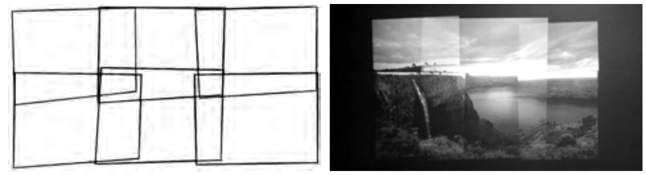
The laparoscope used in this installation is a Stryker 888, a 3-chip CCD camera which only has analog NTSC video outputs. The calibration camera is also an NTSC video camera which includes controls for exposure, aperture, shutter speed and focus— necessary controls for a successful calibration of the projector array.

### B. Software

The multi-projector display system and the interaction station have a variety of software that enables the entire system to apply digital imagery, video and other data sources to various output devices, including the large-area, seamless display constructed from the projector array. For this project, Ubuntu linux was used as the primary operating system. For simplicity of management, the display processing nodes (render nodes) were configured to netboot from the data processing node (elsewhere called the "head" node for this and other reasons).

(a) Precalibrated Display (b) Projector Array

Fig. 3. Precalibrated projector array output. This shows the raw output of each of the display processing nodes. The background (root window) of each X.org server is set to a pre-rendered background containing the node's hostname. The X cursors are set to blank to prevent seeing a cursor overlay on the display.



(a) Projector Geometry (b) No Warp

(c) Warp Only (d) Warp and Blend

Fig. 4. Blending and Warping: (a) shows the raw projector relative geometry, (b) shows an image being displayed on the projectors with both the warp and the blend disabled, (c) shows the same image being displayed with just the warp applied, and (d) shows both warping and blending being applied to generate a seamless, large-area display.

To easily network and distribute image rendering across the nodes, the Chromium project is used [17]. Chromium is an OpenGL wrapper library which allows OpenGL applications to be distributed across a set of computing nodes by modifiying the OpenGL stream with **Stream Processing Units** (SPUs). OpenGL is a standard API for describing 2D and 3D rendering of visual data. Chromium comes with some standard SPUs, and the SPU API is well documented so custom SPUs may be implemented. Our Chromium configuration consists of a standard `tilesort` SPU which divides OpenGL calls from the data processing node up by geometric tiles and sends the separate streams to their corresponding render nodes. The render nodes then use the standard `render` SPU combined with a custom `warp` SPU (described in detail in Section II-B.2) which provides the warping/blending functionality to build a seamless display from the projector array.

The REVEAL team has developed a software platform that runs on top of the Chromium installation detailed above called VIBE which stands for Visually Immersive Blended Environment [18], and is an extension of an earlier system described by Brown and Seales [12]. The VIBE code base consists of four main parts, an OpenGL calibration program which calculates the relative projector geometries, the Chromium `warp` SPU which provides warping and blending based upon the calibration data provided by the previous program, an OpenGL, client/server image processing framework called SmartImage, and a Java based GUI to enable simple configuration and interaction with the whole system from the interaction station. Each of these components is described in detail below.

*1) Calibration:* The calibration program is an OpenGL application that uses Chromium to control each projector and project a structured light grid using a binary coding scheme along with a video camera with manual settings as the feedback element. In sequence, it projects a structured light grid of circular fiducials on each projector to be calibrated. The structured light calibration procedure is described in detail by Seales and Brown [12]. One frame of the calibration pattern can be seen in the VIBE GUI shown in Fig. 5. Blending alpha masks are then generated for each projector from the observed mesh of fidicuals according to Raskar *et al.* [13]. The observed fidicuals and the computed alpha masks are saved to disk for use by the other VIBE tools.

The calibration system technician must ensure that the calibration camera views the entire scene and must set the aperture, shutter speed and focus to appropriate values for the lighting conditions and the brightnesses of the projector. It is essential for accurate calibration that the camera captures a clear image of the entire scene. The aperture is usually adjusted and fixed to full-open and the shutter speed fixed to 1/60 to prevent color separation from the color-wheel inside the DLP projector (which runs at 120Hz). The video images must be consistent in aperture and shutter speed as an initial background image is sampled and averaged from the first few frames of video and is then subtracted from each image to allow for other static objects in the camera's view from confusing the computer vision algorithms.

*2) Warping and Blending:* The `tilesort` SPU is configured in logical tiles which contain a bounding box of the scene data that is applicable to each render node. The scene data in the tiles is a clipped OpenGL stream for that particular bounded tile. The `warp` SPU injects the OpenGL commands necessary to warp the final rendered scene onto the calibrated projected mesh by reading the saved calibration information on initialization [12]. Finally, the `warp` SPU adds OpenGL commands to apply the alpha mask generated for the particular projector in the calibration step. This process yields a real-time, seamless image from the perspective of the calibration camera.

The `warp` SPU also listens for commands on a UDP broadcast address so the interaction station may disable the blend and/or warp at any time, which is useful for demonstration purposes. Fig. 4 shows a projected geometry with no warping or blending correction applied, with only the warp correction applied and with both the warping and blending applied to provide a simple visual of the correction process.

*3) Image Framework:* The imaging framework uses a client/server model. The two communicate using an established API with both standard TCP/IP networking and shared memory for efficient image transfer when the client and server reside on the same machine. Several clients have been implemented: one which displays a test grid pattern, a still

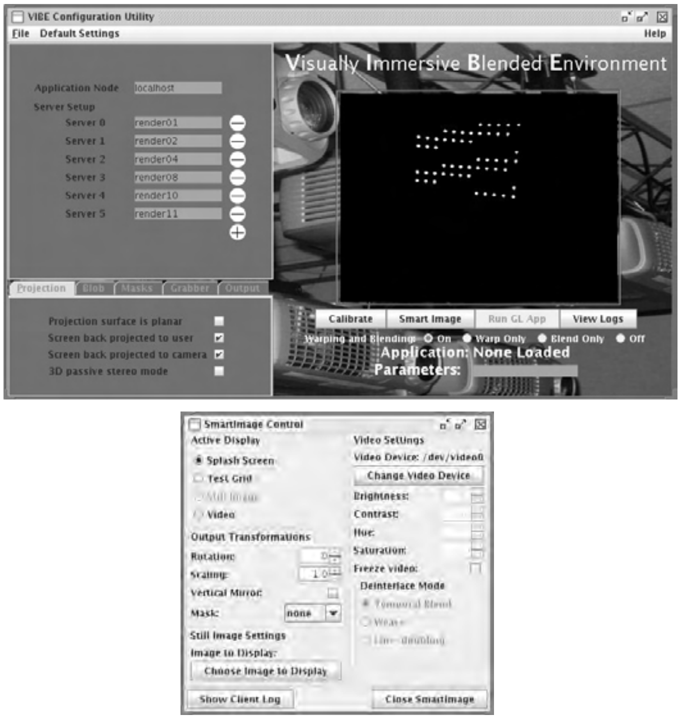



Fig. 5. VIBE interface. A user at the interaction station can control all aspects of the system. A calibration is running in this image, and the observed structured light pattern is visible.

image viewer, a video file viewer and a real-time video-capture viewer. The framework is design to be flexible and general to allow integration of various image-data sources. The test pattern is used for testing and demonstration purposes. The image file and video file viewer are used to visualize high-resolution perioperative data. The video capture viewer is used to display video sources, such as a laparoscopic camera, in real-time over the projector array. The video-capture client includes an efficient software deinterlacer that uses a linear-blending mechanism to generate each deiterlaced frame and runs in real-time at the linux capture API's maximum 29.97 frames per second.

The server receives each set of image data and then splits the image into chunks for efficient tilesorting of the image data. The Chromium `tilesort` SPU can only reject sending a texture to a particular node by examining its bounding box, and if the entire image were sent in one large chunk, the tilesorting mechanism would send all of the image data to all of the nodes (the `tilesort` SPU does not multicast data). Chunks are 64x64 pixel OpenGL textures which contain 62x62 useful pixels from the source image (the border pixels prevent chunk boundaries from being visible due to the anti-aliasing performed on the textures in the video card hardware). These chunks are then seamlessly rendered by the Chromium pipeline.

*4) Interface:* A screenshot of the Java VIBE graphical user interface can be seen in Fig. 5. The interface runs on the interaction station connected to the main data processing node via a network connection. The GUI allows full control of the VIBE system including running a calibration, launching generic OpenGL applications running over the projector array, launching the image processing server and clients, controlling the calibration parameters and controlling the image processing server and clients. It has controls to disable/enable warping and/or blending, to scale, rotate and mirror the output imagery, to select which real-time video stream to display, to alter video capture parameters such as brightness/contrast. It also displays the structured light feedback read from the calibration camera during a calibration run.

## III. System Performance

To analyze the system's usefulness in the surgical setting, its performance must first be measured. We analyze the throughput of the system as a whole, the latency as a whole and the latency of individual components to get an understanding of where the latency comes from.

### A. Image Throughput

End-to-end image throughput can be measured as the maximum sustainable frame-rate. The VIBE system as deployed in the University of Maryland Medical Center Surgical Simulation and Technology Center maintains a constant 29.97 frames per second throughput (the frame rate of NTSC video and the current maximum of the frame grabber). The frame rate is measured by the image server processing node which can monitor how many frames it was able to render through the Chromium pipeline per second. This is a very desirable frame rate as motion artifacts are very difficult to perceive. Therefore, the total raw image bandwidth available to the system is at least 221 Mbps (640x480 images at 24 bits per pixel at the frame rate of 29.97 Hz).

### B. Overall Latency

Overall latency was measured at the University of Maryland Medical Center Surgical Simulation and Technology Center by viewing a digital stop-watch through the laparoscopic camera and projecting the video feed though VIBE. This measurement will yield an overall system latency measurement that can serve as corboration to our other individual component latencies later on. Simply, the overall latency should equal the sum of the component latencies.

A high-shutter speed still camera was used to take 10 pictures of the stopwatch and the stopwatch projected through the system. The difference in the time reading between the actual watch and the view of the watch captured by the laparoscopic camera and seen through VIBE is the total system latency from image acquisition all the way through display. The analysis of the data depends on random sample theory. If a sample is a random sample of the total population and the population variance is estimated by using the sample data, then according to simple random sampling theory, the standard error, $SE$, is given by (2) where $s$ is the sample's standard deviation given by (1), $n$ is the number of observations in the sample, $x_i$ is the $i$th sample, and $\overline{x}$ is the sample mean [19].

$$s = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}} \qquad (1)$$

$$SE = \sqrt{s^2/n} \qquad (2)$$

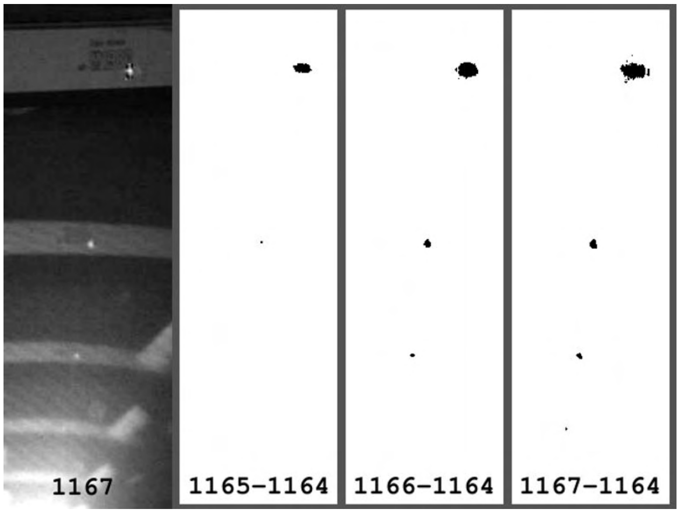



Fig. 6. Measurement of latency inherent to Stryker 888 camera and camera controller involved video recording three feedback loops of the camera through a medical-grade CRT monitor. A red laser beam was switched on and off such that it appeared in each feedback loop. This figure demonstrates one sample (from video frames 1165–1167) where the difference was two video frames. The image on the left is of the final frame when the laser is visible in all three loops (frame 1167). The next image is a frame difference from the first frame the laser appeared (1165) to the frame immediately before (1164), showing the appearance of the laser dot in the view of the world and also the appearance of the laser in the first feedback loop. The third image shows the next frame difference (1166-1164), highlighting the appearance of the laser in the second loop. The last image is the final frame difference (1167-1164), showing that the laser is now visible in all three feedback loops.

Results will be written with the sample mean first, then a "±" symbol, followed by the standard error. Analyzing the 10 samples as a random sample population of all possible sample pictures using the above methodology yields a latency of 120±13.3ms.

### C. Component Latencies

Each component of the system contributes some amount of latency to the system as a whole. The five components which contribute to the latency from image acquisition to display are the laparoscopic camera and controller, the capture card and capture software, the client/server image processing software, the Chromium distributed rendering framework, and the projectors themselves. Each subsection below breaks down each latency component and measures it empirically either with camera sensor feedback or with software feedback.

*1) Camera Latency:* The latency of the Stryker 888 laparoscopic camera was measured by attaching the S-Video output of the camera (which is the output used in the VIBE system) directly to a high-end, medical-grade CRT monitor (Panasonic MT1980). The Stryker 888 camera was aimed at the monitor to create several feedback loops and then clamped in place. A video camera was mounted to a tripod and set to record the entire scene. A red laser was then shone on the top rim of the monitor such that it was visible in the real-world and at least three feedback loops of the laparoscopic camera/monitor system from the recording camera's perspective. The laser was switched on and off so the transitions could be observed on the recording from the monitoring camera. The CRT monitor should have a very low latency compared to the latency of the CCD analog-to-digital and NTSC digital-to-analog conversion that must take place in the laparoscopic camera. The video was analyzed and the delay between when the laser dot was visible in the real world to when it was visible in the third feedback loop was counted for 15 different on pulses giving 15 random samples. Fig. 6 shows one sample where there was a two frame difference between the first appearance of the laser and the appearance of the laser in the final frame. The data was analyzed and the average latency was measured to be 16.3±1.48ms (average of 1.46 frames over 3 feedback loops).


Fig. 7. Measurement of latency inherent to Dell 2300MP projector. As the VGA signal arrives at both the CRT (right) and Projector (left) the delay in the projector can be seen relative to the CRT. Since the CRT is driven directly by the analog signal, the latency of the projector (which is in the tens of milliseconds) is much greater than the latency of the CRT display. This is one of fifteen pictures taken to measure the average latency.

*2) Capture Latency:* The capture latency was measured in the frame grabber driver in the linux kernel. It was defined as the amount of delay from when a VSYNC was detected on the analog input indicating the end of a field to when the completed frame was actually returned to the user process which intitiated the frame grab. The analysis didn't begin until the frame grabber reached a steady state as it took a few frames before the card tuned into the signal and "settled down." The latency was measured over 30 frames to be 66.0±.37ms.

*3) Application Latency:* The application latency was measured as the time it took from image capture in the frame buffer card to when the image frame was handed off to the Chromium layer for processing (via a call to `glutSwapBuffers()`). This time difference was measured over 30 frames, again allowing for a settling period. The average latency was 6.652±.033ms.

*4) Chromium Latency:* The latency in the Chromium tile-sorting, distributed rendering and warping was measured by finding the amount of time spent in the OpenGL `glutSwapBuffers()` call which does not return until every display processing node has finished rendering the entire scene in the video graphics card. Again, 30 frames of data were collected once the application reached a steady state. The latency was measured to be 18.53±0.011ms.

*5) Projector Latency:* Measuring the projector latency required a more complex setup. A laptop was used as a digital clock signal. A large font digital clock was run on the laptop

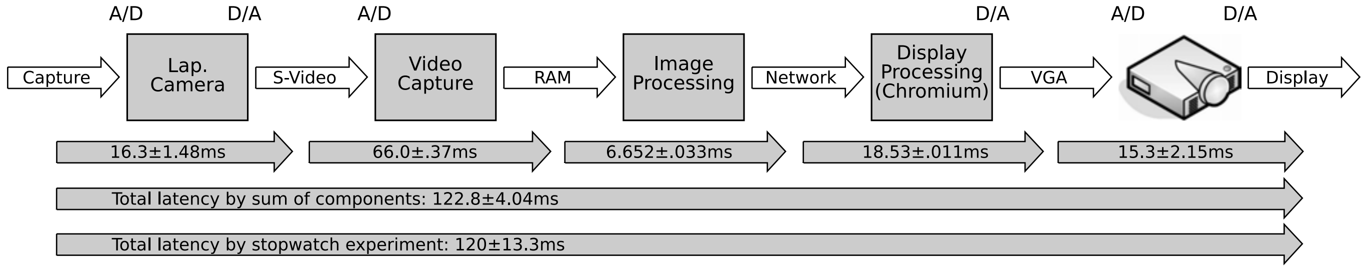



Fig. 8. Diagram showing the component latency measurements and overall latency measurement. Also, the diagram shows analog-to-digital (A/D) and digital-to-analog (D/A) conversion steps in the processing of one image frame.

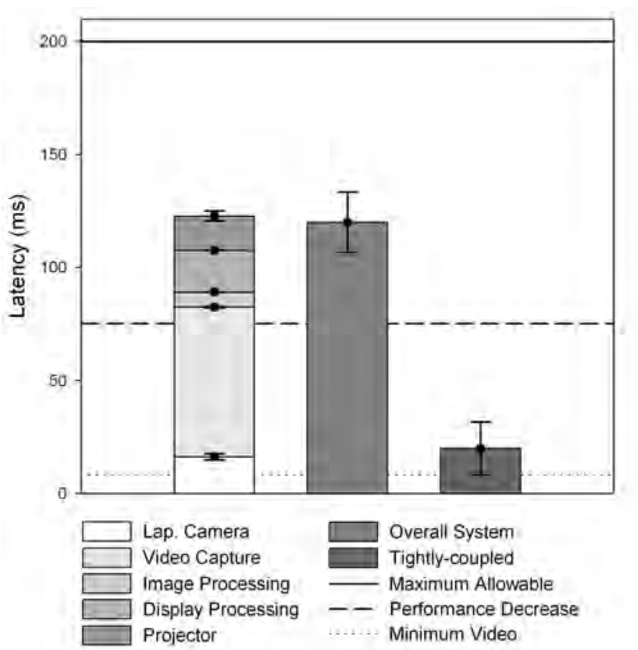



Fig. 9. Graph of latency results. The graph shows latency measurements and confidence intervals of the immersive environment's components, overall immersive environment, and a typical tightly-coupled laparoscopic surgical environment. These are compared to lines depicting the maximum allowable latency (200ms), the point at which significantly measurable performance degradation occurs (75ms) and the minimum latency possible in an NTSC video system (8.33ms).

and the VGA output attached to an active VGA duplicator. One output of the duplicator was sent to a high-end computer CRT monitor; the second output was sent to the Dell 2300MP projector. A digital camera with the shutter speed set to 1/125 was used to capture 15 random images of the projected image beside the CRT monitor. Fig. 7 shows one of the pictures captured. The shutter speed is chosen to be slightly more than half of the 60Hz refresh used by the VGA signal (the same vertical refresh used in the system as a whole). Since the latency of the CRT (where the analog signal directly drives the display) is much less than the latency of the projector (where the analog signal must be digitized, applied to the DLP chip and then displayed) the time difference between the two clocks shown should give the latency of one sample. Using this method, the latency of the projector was measured to be 15.3±2.15ms.

### D. Latency Summary

The overall latency was determined to be 120±13.3ms. The sum of the above component latencies yields a calculated total latency of 121.6±3.85ms. Fig. 8 shows the latencies of each component, their sum and the overall latency measured by the stopwatch experiment. Fig. 9 shows the results on a precise graph with error bars compared to the latency of current tightly-coupled systems with some key latency points indicated with vertical lines. The component latency measurements are validated by the overall latency as their sum falls neatly within the error range of the overall latency measurement.

Obviously, the overall system latency can be reduced by reducing the component latencies. The largest signal contributor of latency is the video capture component. Reducing this latency can be achieved by eliminating the unnecessary digital-to-analog conversion in the laparoscopic camera and the analog-to-digital conversion in the video capture card. This can be achieved by using digital camera technologies found in newer model laparoscopic cameras and by using digital video capture cards. Also, eliminating the digital-to-analog conversion on display output may also reduce the projector latency. This can be achieved by using a digital interface between the display processing component and the projector. Latencies can be further reduced by increasing the efficiency of the Chromium component, or by replacing Chromium with another technology that has lower latency.

## IV. Surgical Application

The large display size of the VIBE display does provide the surgeon a more immersive and flexible environment in which to perform laparoscopic surgery. However, for the surgeon to be able to use the VIBE system for surgery it must conform to certain performance expectations.

For the video aspect of the system to be usable in real-time it must have a latency less than 200ms and must have a frame rate of at least 10 frames per second. [20]–[23]. Pausch shows that low-latency is significantly more important than high resolution or stereoscopic vision [24]. Also, MacKenzie and Ware show that even a lag (latency) of 75ms will degrade performance and increase error rate slightly [21], so the lower the latency the better.

The VIBE system meets the minimum video requirements. It displays frames at the full NTSC frame rate of 29.97 frames per second and has a moderate latency of about 120ms. Also, it performs real-time deinterlacing of the NTSC video stream which improves picture quality significantly by removing interlacing artifacts that are readily visible on large, high-resolution video screens. It has precise color reproduction and can account for non-planar display surfaces. However, while the system does meet the minimum requirements, lowering the latency would make the system even more desirable for surgery as it would increase performance. The current tightly-coupled approach, while sacrificing flexibility and bandwidth, provides a low latency of 10-30ms. The difference in latency is due to the cost of separating the tightly-coupled input/outputs and adding a processing and collection step. As technology improves, this latency will lessen, but will never be less than a tightly-coupled approach. We think that the benefits of a flexible, integrated approach outweight the latency penalty.

### A. Perioperative Data

For high-resolution still images such as radiological imagery, the images need not appear in real-time. However, the delay between a request to view a high-resolution image should not be greater than a few seconds for the system to be responsive. Since the VIBE system has a total throughput of 9.2 megapixels a second, it can display even the highest resolution photography within that time requirement. Also, the VIBE system is capable of rendering OpenGL modeling streams in real time using the Chromium backbone. Since the entire system is connected to the network, by tying into a facilities electronic records system it is possible to display perioperative data such as radiological imagery during the surgery alongside the intraoperative video stream. Also, there is plenty of left-over bandwidth for such (comparatively) low-rate displays such as vital sign monitoring.

## V. Conclusion

The system as described promises to improve performance and provide new functionality to the laparoscopic environment by decoupling the laparoscopic camera and display and inserting a flexible architecture in-between that is capable of integrating disparate data sources and output mechanisms. The implementation of the system at the University of Maryland Medical Center Surgical Simulation and Technology Center was described. The architecture was shown through detailed analysis and experimentation to provide sufficient performance to display both real-time video from the laparoscopic camera and other high-resolution static imagery such as perioperative radiological scans on a large-area, casually-aligned, automatically-calibrated projector array. This will improve the surgeon's environment by providing a large-area display capable of high-bandwidth imagery and by centralizing all data output in one area improving context awareness by integrating disparate sources of information.

Future implementation could be improved in many ways. More data sources and more output mechanisms can be added to the architecture to increase spatial and multi-modal integration. Higher resolution video imagery can be leveraged to increase the quality of imagery. New hardware technologies can be leveraged such as faster internetworking topologies, digital video interfaces that avoid analog-to-digital and digital-to-analog conversions. The network software can be streamlined by broadcasting the digital images instead of using point-to-point communications.

## References

[1] "Surgical video systems used in laparoscopy," *Health Devices*, vol. 24, no. 1, pp. 28–33, January 1995.

[2] N. Bitterman, "Technologies and solutions for data display in the operating room," *Journal of Clinical Monitoring and Computing*, vol. 20, pp. 165–173, 2006.

[3] R. Berguer, "Surgical technology and the ergonomics of laparoscopic instruments," *Surgical Endoscopy*, vol. 12, pp. 458–462, 1998.

[4] R. Berguer, D. L. Forkey, and W. D. Smith, "Ergonomic problems associated with laparoscopic surgery," *Surgical Endoscopy*, vol. 13, pp. 466–468, 1999.

[5] D. W. Rattner and A. Park, "Advanced devices for the operating room of the future," *Seminars in Laparoscopic Surgery*, vol. 10, no. 2, pp. 85–89, June 2003.

[6] D. M. Herron, M. Gagner, T. L. Kenyon, and L. L. Swanström, "The minimally invasive surgical suite enters the $21^{st}$ century," *Surgical Endoscopy*, vol. 15, pp. 415–422, 2001.

[7] W. B. Seales and J. Caban, "Visualization trends: applications in the operating room," *Seminars in Laparoscopic Surgery*, vol. 10, no. 3, pp. 107–114, September 2003.

[8] W. B. Seales and D. Clarke, "Computing support for information-rich laparoscopy," *Surgical Innovation*, vol. 12, no. 4, pp. 357–363, December 2005.

[9] F. Tyndiuk, V. Lespinet-Najib, G. Thomas, and C. Schlick, "Impact of large displays on virtual reality task performance," in *3rd International conference on computer graphics, virutal reality, visualization and interaction*, 2004, pp. 61–65.

[10] D. S. Tan, D. Gergle, P. Scupelli, and R. Pausch, "Physically large displays improve performance on spatial tasks," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 13, no. 1, pp. 71–99, March 2006.

[11] P. Baudisch, N. Good, V. Bellotti, and P. Schraedley, "Keeping things in context: A comparative evaluation of focus plus context screens, overviews, and zooming," *CHI Letters*, vol. 4, no. 1, April 2002.

[12] M. S. Brown and W. B. Seales, "Incorporating geometric registration with PC-cluster rendering for flexible tiled displays," *International Journal of Image & Graphics*, vol. 4, no. 4, pp. 653–681),, October 2004.

[13] R. Raskar, M. S. Brown, R. Yang, W.-C. Chen, G. Welch, H. Towles, W. B. Seales, and H. Fuchs, "Multi-projector displays using camera-based registration," in *IEEE Visualization*, 1999, pp. 161–168.

[14] M. Brown, W. Seales, S. Webb, and C. Jaynes, "Building large format displays for digital libraries," *Communications of the ACM*, vol. 44, no. 5, 2001.

[15] C. Jaynes, W. B. Seales, K. Calvert, Z. Fei, and J. Griffioen, "The Metaverse–a collection of inexpensive, self-configuring, immersive environments," in *7th International Workshop on Immersive Projection Technology/Eurographics Workshop on Virtual Environments, Zurich, Switzerland*, May 2003.

[16] R. Raskar, J. van Baar, and J. X. Chai, "A low-cost projector mosaic with fast registration," in *Asian Conference on Computer Vision (ACCV)*, January 2002.

[17] "Chromium homepage." [Online]. Available: http://chromium.sourceforge.net/

[18] "VIBE: Visually Immersive Blended Environment." [Online]. Available: https://halsted.vis.uky.edu/downloads/download.php

[19]

[20] E. Hoffman, "Fitts' law with transmission delay," *Ergonomics*, vol. 35, no. 1, pp. 3–48.

[21] I. MacKenzie and C. Ware, "Lag as a determinant of human performance in interactive systems," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 1993, pp. 488–493.

[22] C. Smith, T. Farrell, S. McNatt, and R. Metreveli, "Assessing laparoscopic manipulative skills," *American journal of surgery*, vol. 181, pp. 547–550, 2001.

[23] C. Ware and R. Balakrishnan, "Reaching for objects in VR displays: lag and frame rate," *ACM Transactions on Computer-Human Interaction*, vol. 1, no. 4, pp. 331–356, 1994.

[24] R. Pausch, "Virtual reality on five dollars a day," in *Proc. of the CHI '91 Conference ono Human Factors in Computing Systems*.

**C. Andy Martin** received both the B.S. degree with honors in computer science and the B.S. degree with honors in electrical engineering from University of Kentucky, Lexington in 2003.

He is currently a Software Engineer for the REVEAL project at the Center for Visualization and Virtual Environments, University of Kentucky. From 2003–2006 he worked as the Senior Engineer at Lumenware, LLC, Lexington, Kentucky. His past experience also includes Software Engineering positions at the Department of Defense and Lexmark.

**Duncan Clarke** received the B.S. degree in computer science from Michigan State University, East Lansing, Michigan in 1986, the M.S in computer science from Rensselaer Polytechnic Institute, St. Troy, New York, in 1987 and the Ph.D. in computer and information science from the University of Pennsylvania, Philadelphia, Pennsylvania in 1996.

He is currently founder and software development consultant for Fremont Associates, a consultancy specializing in software project management and research on tools and techniques to support the development process for embedded systems. He has held positions in both industry and academia including five years as a Member of Technical Staff at AT&T Bell Laboratories, one year as a senior analyst at Proton World International, Brussels, Belgium, one year as an Ingenieur Expert with INRIA/IRISA, Rennes, France, and four years as a faculty member at the University of Kentucky and the University of South Carolina. His research interests include software testing, formal methods for real-time systems, formally based software development tools and visualization of specifications. He has published numerous research articles in these areas.

Dr. Clarke is a member of various professional societies including the IEEE Computer Society and the ACM, he has served on program committees for various workshops and conferences within his research area, and serves as a Program Evaluation Visitor for the Accreditation Board for Engineering and Technology.

**C. Melody Carswell** received the FIX degree at the University of Fixme, Fixme, and the FIX degree at the University of Fixme.

She is an Associate Professor in the Department of Psychology at the University of Kentucky, Lexington, Kentucky. Her major research interests fall within the fields of engineering psychology and human factors engineering.

**Adrian Park** received the B.S. degree with honors at the University of Guelph, Guelph, Ontario, Canada, the M.D degree from McMaster Universtiy, Hamilton, Ontario, Canada, residencies in general surgery at McMaster University and St. Joseph's Hospital, Ontario, Canada, and the fellowship in laparoscopic/endoscopic surgery at the University of Montreal, Quebec, Canada.

He is currently the Chief of General Surgery, University of Maryland Medical Center, the Campbell and Jeanette Plugge Professor of Surgery, University of Maryland School of Medicine, and the Director of the Surgical Simulation and Technology Center at the University of Maryland Medical Center. His clinical interests lie in the field of minimally invasive gastrointestinal surgery, including the treatment of gastroesophageal reflux, achalasia, hiatal hernia, intestinal disorders and ventral/incisional hernias. He also has extensive experience and interest in laparoscopic splenic and adrenal surgery. He is an active researcher in minimally invasive surgery and has been directly involved in initiating, developing, and applying new techniques and technologies in endoscopic surgery. He is the author of over 150 scientific articles and book chapters, and is internationally known as a leading authority in the field of minimally invasive surgery, "OR of the Future" technology and surgical education.

**W. Brent Seales** received the FIX degree at the University of Fixme, Fixme, and the FIX degree at the University of Fixme.

He is currently an Associate Professor in the Computer Science department at the University of Kentucky, Lexington, Kentucky.