

# Remote 3D Medical Consultation

Greg Welch\*, Diane H. Sonnenwald<sup>†</sup>, Ketan Mayer-Patel\*, Ruigang Yang<sup>‡</sup>, Andrei State\*,  
Herman Towles\*, Bruce Cairns, M.D.<sup>§</sup>, and Henry Fuchs\*

\*The Department of Computer Science  
The University of North Carolina at Chapel Hill  
Chapel Hill, NC 27599-3175  
Email: {welch, kmp, andrei, herman}@cs.unc.edu

<sup>†</sup>The Swedish School of Library and Information Science  
Göteborg University and the University College of Borås  
501 90 Borås, Sweden  
Email: diane.sonnenwald@hb.se

<sup>‡</sup>Department of Computer Science  
The University of Kentucky  
Lexington, KY 40507-1464  
Email: ryang@cs.uky.edu

<sup>§</sup>Department of Surgery  
The University of North Carolina at Chapel Hill  
Chapel Hill, NC 27599  
Email: bruce\_cairns@med.unc.edu

**Abstract**—Two-dimensional (2D) video-based telemedical consultation has been explored widely in the past 15–20 years. Two issues that seem to arise in most relevant case studies are the difficulty associated with obtaining the desired 2D camera views, and poor depth perception. To address these problems we are exploring the use of a small array of cameras to reconstruct a real-time, on-line 3D computer model of the real environment and events. We call this 3D medical consultation (3DMC). The idea is to give remote users what amounts to an infinite set of stereoscopic viewpoints, simultaneously addressing the visibility and depth perception problems associated with 2D video. Here we describe our current prototype system, some of the methods we use, and some early experimental results.

## I. INTRODUCTION

We report here on the progress of a multi-year effort to develop, apply, and evaluate technology for remote three-dimensional medical consultation (3DMC). Our primary aim is to enhance and expand medical diagnoses and treatment in a variety of life-critical trauma scenarios. Our long-term goal is to provide both an expert consulting health care provider and a distant medical advisee with a high-fidelity visual and aural sense of 3D presence with each other. Our approach is to develop computer vision, graphics, and networking technologies that allow advisors and advisees to see “live” stereoscopic imagery of each other, to be able to move around to obtain a continuum of different viewpoints, and to constantly be aware of patient vital signs via integrated virtual monitors.

In this article we describe the motivation for the work, the current status of our research and prototype efforts, and a set of controlled experiments to evaluate whether a 3D approach can

overcome the limitations found in 2D medical consultation, while also being compatible with existing work practices.

### A. 2D Medical Consultation

Two-dimensional (2D) video-based medical consultation has been explored widely in the past 15–20 years. 2D tele-video has been demonstrated to be acceptable for face-to-face consultation—supplementing the telephone, and useful for visual examinations of wounds, abrasions, etc. However with the latter use in particular, the two issues that seem to arise in most relevant case studies are (a) the difficulty associated with obtaining the desired 2D camera views, and (b) depth perception.

For example, camera view difficulties were mentioned in multiple places in the final report for NLM’s National Laboratory for the Study of Rural Telemedicine [1]. One example is in the discussion of the use of the a 2D televideo system to observe children with swallowing disorders. The report states

“Limitations of telemedicine services for management of feeding and growth issues include the need to rely on the interpretations of others during physical exam. At times the camera angles were not ideal to allow for clear pictures of the mouth during feeding.”

Similarly, included in the concerns identified with the university’s “Clinical Studio” are the need for periodic movement of cameras and improper camera locations.

“Full-motion video and audio of the neurological examination is a reliable means of visualizing the patient between remote locations. This technology is not difficult and can be done by ER staff. However the images are in two dimension hence certain aspects of the exam could be enhanced by more than one camera angle.”

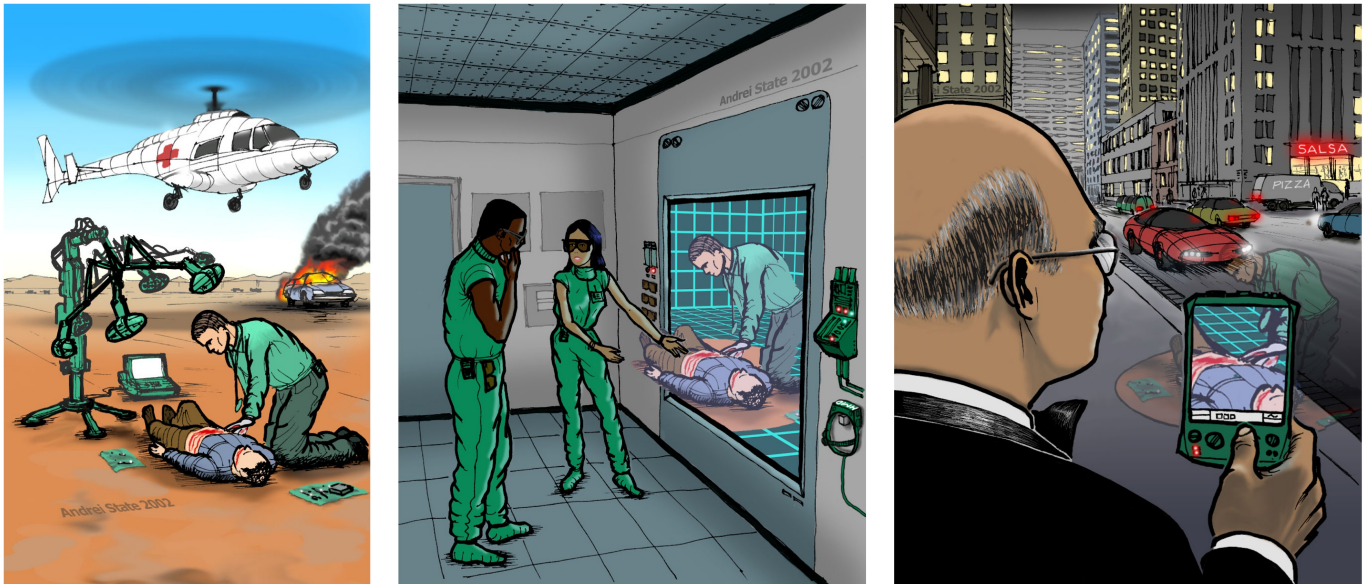


Fig. 1. Conceptual sketches of some three-dimensional (3D) medical consultation paradigms. Left: a portable multi-camera unit. Center: hospital-based 3D viewing. Right: mobile 3D viewing. (Sketches by Andrei State.)

The problem was also identified in [2] where they describe work using a computer-based telemedicine system for semi- and non-urgent complaints at a short-term correctional facility.

“The lack of remote control on the patient care camera at the remote site by the examining emergency medical physicians requires the nurse to spend considerable time operating the camera and responding to technical instructions. This problem has been resolved in a recent system upgrade, but it was another important reason for nonuse.”

Beyond obtaining the desired 2D view of a remote patient, [3] points out that “Impaired depth perception is a significant problem in telemedicine.” and notes that “The most important cue of depth is due to binocular disparity.” The author describes several “coping strategies” that can be used to overcome the inherent limitation of the 2D imagery. Chief among the coping strategies is the practice of “Rotating the camera in the transverse plane about 30° at a time....” This is not surprising given that object occlusion and motion parallax are two of the most powerful depth cues. The author surmises that this controlled camera rotation “...enables the consultant to build a three-dimensional mental image of the object by briefly storing a range of two-dimensional views.”

There are two primary problems illustrated by the above examples. First, for *any* chosen configuration of remote 2D cameras, it is unlikely that the available views will always match the consulting physician’s desired views. One could try and address the visibility problem using multiple cameras. But switching between numerous disjoint views as a security guard might with a surveillance system is not very natural. With a very large number of cameras and user head tracking, one could imagine automatic switching based on view position and orientation. But the quantity and configuration of cameras necessary to achieve smooth and appropriate switching over an operating room, as well as the 2D video storage and bandwidth needs, would be impractical. While pan-tilt-zoom cameras can

help address this problem, they require additional technical skills, impose an additional cognitive load, and require additional time to adjust (difficult in a trauma situation).

Second, in cases where depth perception would indirectly or directly aid in the consultation, users must resort to secondary visual cues or verbal clarification (from another remote collaborator), which both impose additional cognitive loads compared to the very natural views afforded if the consulting physician were able to “be there” with the patient or the collaborating medical personnel.

### B. 3D Medical Consultation

To address the 2D problems outlined in Section I-A above, we are working on systems for *three-dimensional* medical consultation (3DMC). The basic idea is to use a relatively small number of cameras to “extract” (estimate) a time-varying 3D computer model of the remote environment and events. When coupled with head (or handheld viewer) position and orientation tracking, this should offer a remote consultant a continuum of dynamic views of the remote scene, with both direct and indirect depth cues through binocular stereo and head-motion parallax. See Figure 1 for example scenarios.

We believe that such a system will provide the consulting physician with an increased *sense of presence* with the remote patient and medical personnel, thus improving communication and trust between the participants. There is evidence to support the increased sense of presence. In [4], the authors report on the results of three studies where they vary display parameters and attempt to assess the users’ sense of presence. The authors report that the results of the first and second studies indicated that the reported level of presence was significantly higher when head tracking and stereoscopic cues were provided. They report that the third study indicated that the level of presence increased with the visual field of view.

There is also evidence to suggest that an immersive 3D display will increase a user’s *performance* on certain tasks. For example [5] reported a moderately positive relationship between perceived presence and task performance. In [6] the authors present the results of a study where they found that users performing a generic pattern search task decrease task performance time by roughly half when they change from a stationary 2D display to a head-mounted (and tracked) 3D display with identical properties. In [7] the authors present the results of a study where distant collaborators attempted to solve a Rubik’s cube type puzzle together. The authors compared face-to-face (real) task performance with networked performance using both a immersive 3D display and a conventional 2D desktop display. They found the task performance using the networked immersive 3D display and in the real scenario were very similar, whereas desktop performance was “much poorer.” Most recently, in [8] the authors describe a careful 46-person user study aimed at determining whether or not immersive 3D virtual reality technology demonstrated a measurable advantage over more conventional 2D display methods when visualizing and interpreting complex 3D geometry. The authors found that the head-tracked 3D system showed a statistically significant advantage over a joystick-controlled 2D display.

### C. Why is Obtaining 3D So Difficult?

The most common approach to 3D scene reconstruction is to use cameras and effectively “triangulate” points in the scene. This involves automatically picking some feature in one camera’s 2D image, finding the same feature in a second camera, and then mathematically extending lines from the cameras into the scene. The place where the lines intersect corresponds to the 3D location of the feature in the room. If one can do this reliably for a sufficient number of points in the scene, many times per second, then with a lot of compute power one can turn the dynamic collection of disjoint 3D points into a coherent dynamic 3D computer model that one can use like a flight simulator.

However there are at least three areas of fundamental difficulty associated with trying to reconstruct dynamic 3D models of real scenes: feature visibility, feature quality, and reconstruction algorithms. Features might not exist or might be confusing/ambiguous, they are hard to detect, isolate, resolve, and correlate, and automating the overall reconstruction process in light of these difficulties is a very hard problem. The state of the art is limited to static environments for large spaces, or dynamic events in relatively small controlled spaces. In Section II-A we describe a promising new approach we call *View-Dependent Pixel Coloring*.

### D. Project Status and This Article

The work we are reporting on here is primarily associated with a multi-year project sponsored by the NIH National Library of Medicine (Craig Locatis and Michael Ackerman). Currently the project includes three somewhat independent

areas of research: real-time computer vision/graphics, network adaptation strategies, and a formal evaluation of the (likely) effectiveness of 3D medical consultation. By the end of the project our goal is to have a working prototype that combines the first two areas (vision/graphics and networking), and complete results from our formal evaluation.

With respect to this article, in Section II we describe our research efforts—which include both the vision/graphics and networking aspects, and present some results, and in Section III we describe the formal evaluation.

## II. RESEARCH

Our current 3DMC prototype consists of a camera-lighting array, a transportable compute/rendering cluster, and a head-tracked display. See Figure 2.

The camera-lighting array (Figure 2, left) consists of eight  $640 \times 480$  resolution digital (IEEE 1394a) color cameras from Point Grey Research [17]. The cameras are currently mounted in two horizontal rows of four on a portable stand that can be positioned next to a patient. The cameras are positioned so their visual fields overlap the same region of interest on the patient. Mounted around the cameras are multiple Stocker-Yale high-frequency fluorescent fixtures for flicker-free illumination.

The compute/rendering cluster (Figure 2, right) consists of five dual-processor PCs mounted in a transportable, shock-resistant rack case. Four PCs function as camera servers, JPEG compressing the raw Bayer pattern images and forwarding the compressed video streams via dedicated gigabit Ethernet to the 5<sup>th</sup> PC.<sup>1</sup> The 5<sup>th</sup> PC then decompresses the video streams, loading the color images into texture memory of the graphics card for view-dependent 3D reconstruction as described in Section II-A. To achieve synchronized image capture and time stamping of the eight video streams, the four 1394 video networks also include Point Grey synchronization units.

Note that as indicated in Section I-D, our current prototype streams the compressed video over the dedicated gigabit ethernet *without* the network adaptation strategies described in Section II-C. By the end of the project we expect these strategies to be fully integrated.

### A. 3D Reconstruction

The 3D reconstruction process involves two major steps: the reconstruction of 3D *points* from 2D images and the reconstruction of 3D *surfaces* from the 3D points. To reconstruct 3D points from 2D images we use a novel approach called *View-dependent Pixel Coloring* [9]. VDPC is a hybrid image-based and geometric approach that estimates the *most likely color* for every pixel of an image that would be seen from some *desired viewpoint*, while simultaneously estimating a 3D model of the scene. By taking into account object occlusions, surface geometry and materials, and lighting effects, VDPC can produce results where other methods fail—in the presence

<sup>1</sup>Each camera server can optionally record the video streams to disk.





Fig. 2. The current 3DMC prototype. Left: a camera-lighting array with eight Firewire cameras and high-frequency area lights. Right: co-author Andrei State uses the transportable rack with five high-performance servers and a head-tracking interface to view live 3D imagery from the eight cameras.

of textureless regions and specular highlights—conditions that are common in medical scenes.

As described in [10] we use the graphics hardware on the 5<sup>th</sup> PC (see above) to perform the 3D reconstruction very quickly as the images arrive from the camera server PCs. The basic idea is to use the graphics hardware to rapidly render the camera images onto a series of virtual (computer graphics) planes swept through the scene, searching in parallel for the best color matches (least variance) at a dense set of points on the planes. The result is a relatively dense depth map that we can then render, again using the graphics hardware.

In some early experiments we constructed a one-meter-cubed rig with eight downward looking cameras. Figure 3 shows a sequence of images reconstructed from a recording of Dr. Bruce Cairns, M.D. performing a mock surgical procedure on a physical patient model. Those views (images) were reconstructed off line after the procedure. Figure 4 shows some results from our current prototype (Figure 2). Those views were reconstructed on line, in real time.

Note that in both cases the views were reconstructed and rendered from completely novel view points—none the same as any of the cameras, at different times during the sequence.

### B. Displays

When a medical advisor is on duty in a hospital, it is reasonable to expect that they might have access to facilities for stereoscopic, head-tracked viewing of dynamic 3D reconstructions of the remote patient and advisee (Figure 1, center). Our current prototype addresses this scenario with a high-resolution LCD panel and an Origin Instruments optoelectronic tracker with software for head-tracked visualization. As shown on the right in Figure 2, the user wears a head

band with three infrared LEDs that are tracked in real time by a small sensor unit. From this we compute the location of the user’s dominant eye and render the reconstructed imagery from that point of view. Thus the user can observe the reconstructed view with natural/intuitive monoscopic head-motion parallax. We plan to add time-division multiplexing (shuttered) stereoscopic visualization soon, and have even been investigating *autostereoscopic* displays.<sup>2</sup>

We also want to provide the best possible 3D experience when the medical advisor is away from the hospital (Figure 1, right). For a remote display we are looking at *personal digital assistants* (PDAs). Our goal is to develop or adapt tracking technology and user interface paradigms that will allow a remote medical advisor to use a PDA as a “magic lens” for looking at the remote scene [11]–[14].

We are currently investigating a two-handed patient “prop” paradigm [15]. The basic idea is that the advisor would have a physical prop that serves as a surrogate for the patient, and a PDA that is tracked *relative to the prop*. For example, the PDA cover could serve as the prop. The advisor would then hold the prop in one hand and the PDA in the other, moving them around with respect to each other as needed to obtain the desired view. Figure 5 shows two different prototypes.

For our specific application the prop paradigm provides the user with an instant visual target to aim their “magic lens” at. The approach also affords new ways of looking at the data. For example, an advisor can rotate the prop to quickly get a different view, rather than spending time and energy walking around to the other side. As a bonus, tracking a PDA relative to

<sup>2</sup>Autostereoscopic displays provide one more viewers with a fixed number of stereo views (for example eight) of a 3D scene, without the use of special user-worn glasses. See <http://www.opticalitycorp.com>.



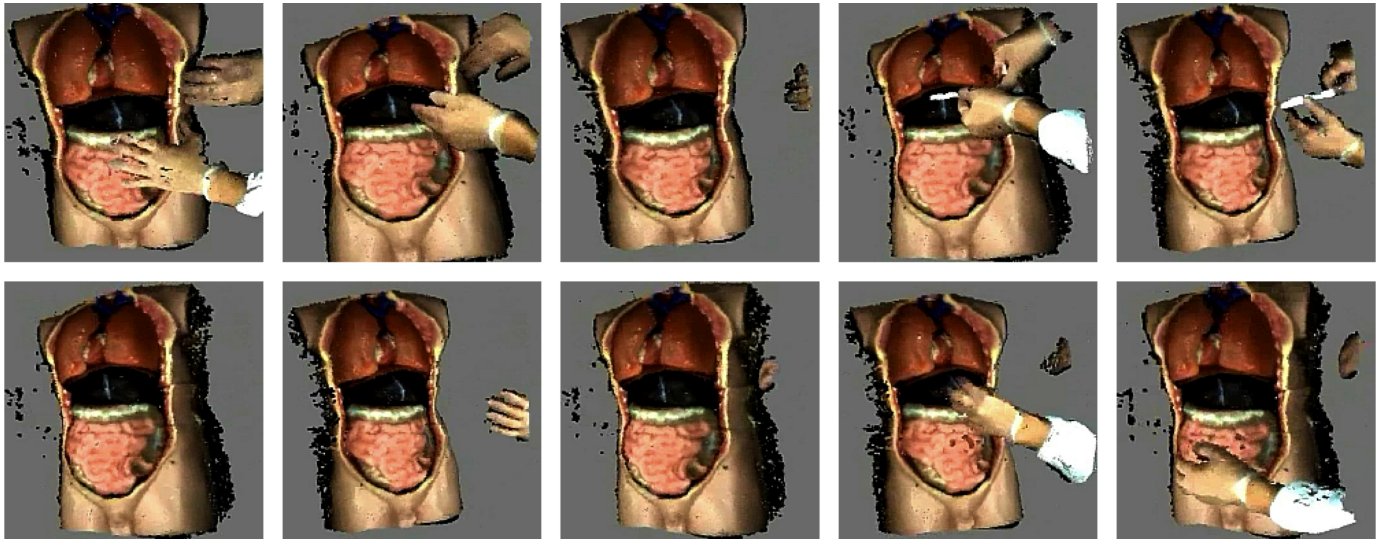


Fig. 3. Novel view images from an early mock surgical procedure on a training torso. Each image is from a different time during the session, and a completely novel viewpoint (none the same as the cameras).

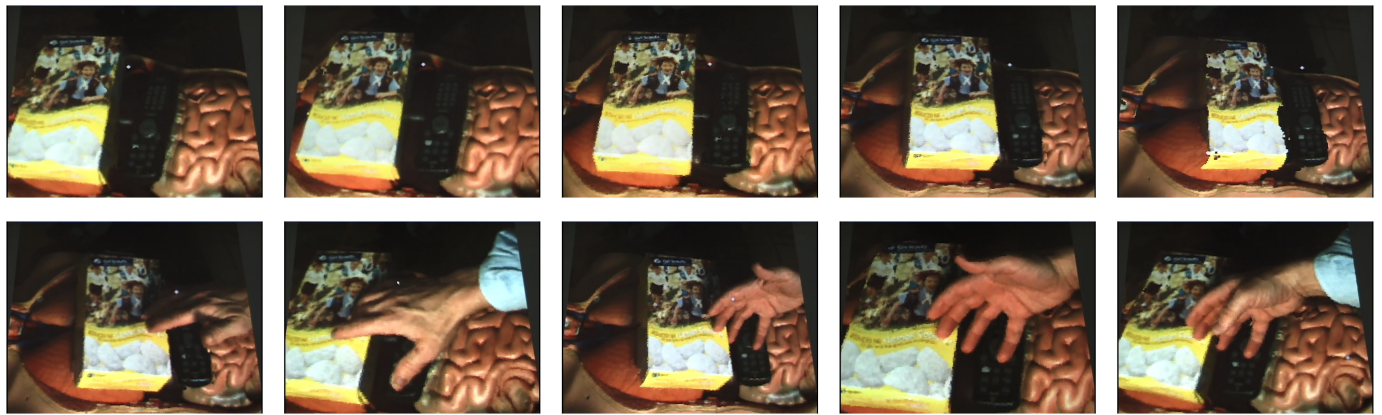


Fig. 4. A more recent sequence of novel view images reconstructed using the system shown in Figure 2. We set a box of Girl Scout cookies on top of the training torso to provide more obvious scene geometry. As in Figure 3, each image is from a different point in time and a completely novel viewpoint.

another object is a much more tractable problem than tracking a PDA relative to the world in general, opening up a number of potential tracking solutions that were otherwise not feasible.

### C. Networking

Managing adaptation is a particularly interesting and challenging aspect of remote 3D medical consultation. In many of our target 3DMC scenarios, the network path represents a significant bottleneck resource. Thus, we must carefully manage this resource in order to ensure that at all times we transmit the data (images or 3D) that is most useful to the overall application and the goals of the user. While adaptation has long been a focus of active research in multimedia systems, 3DMC raises a number of new and interesting problems.

In particular, 3DMC represents an instance of a much larger and more general class of multimedia applications in which many media streams are generated, with complex interstream semantic relationships. In other words, unlike traditional video conferencing applications in which the content of each par-

ticipant's video stream is largely independent of the other participants, the content of media streams produced in the context of 3DMC are likely to be very much related. The most obvious example of this can be seen in the case of video data in which many cameras are capturing the scene from a number of different angles. Thus the visual characteristics of one video stream (i.e., amount of motion, lighting, etc.) can be expected to be shared to some extent by the other streams as well. Furthermore, the utility of the visual information in one stream may depend on the quality and utility of visual information in a peer stream. For example, given two video streams that share a significant overlap of field of view, it may be preferable to allocate available bandwidth to capture and transmit a high-quality image for only one of the two streams while allowing the quality of the other stream to degrade. Alternatively, it may be better to allocate bandwidth equally in order to achieve similar quality for both streams (useful for good stereo correlation and high-quality 3D reconstruction from these two streams). Which of these two adaptation

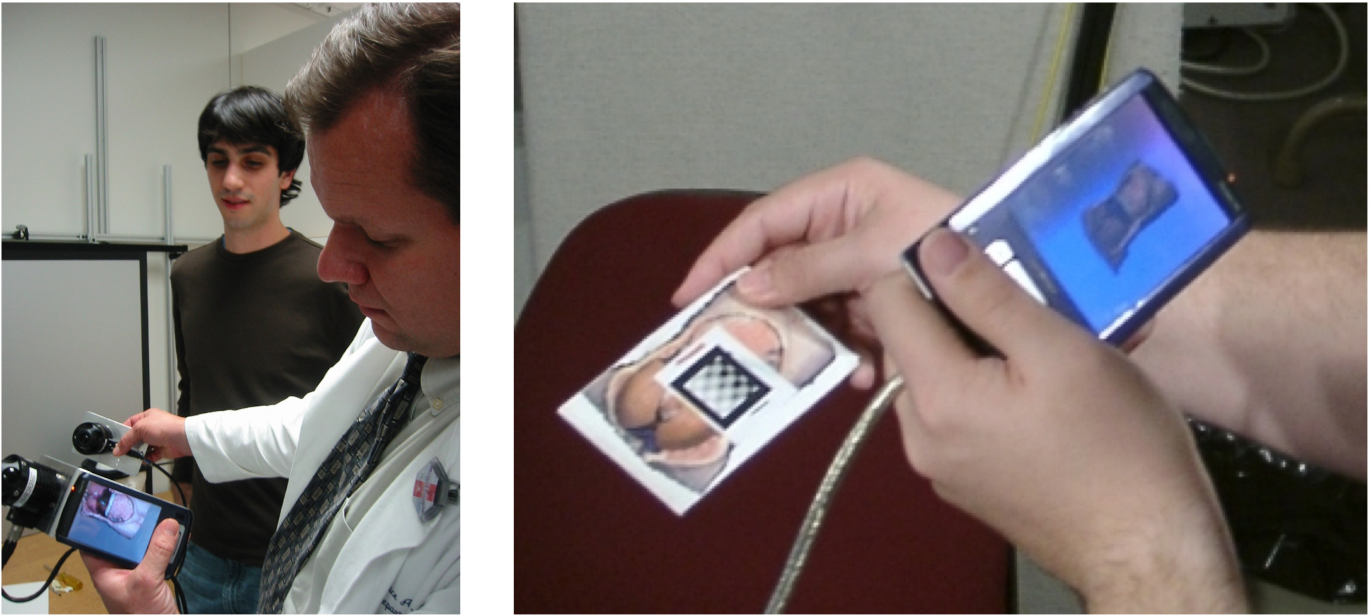


Fig. 5. Left: Our first tracked PDA prototype used a HiBall-3000<sup>TM</sup> tracking system [16], with sensors mounted on the PDA (Toshiba e800, left hand) and the surrogate (right hand). Right: Our current prototype uses a PointGrey DragonFly camera [17] mounted on the PDA (left hand). The prop (right hand) has a printed image of our training torso on it, along with a grayscale pattern. We use the ARToolkit [18] to track the surrogate with respect to the PDA (camera).

solutions is preferable depends on any number of factors including what the application is trying to achieve with the video data and how the user is interacting with the system.

The challenge we face when tackling the adaptation problem in 3DMC is twofold. First, how can we compactly and intuitively specify an adaptation policy to support specific user-level goals? Second, given a particular adaptation policy and set of user-level goals, how can we efficiently evaluate that policy relative to available resources and the way the data are represented and organized?

The 3DMC we envision involves many different media sources (video, audio, patient vital signs, etc.) used in complex ways. As such we need a framework for addressing the problems of adaptation that is more formal and mechanical than previous approaches, which often rely on statically defined priorities (e.g., prioritize audio over video) or simple rule-based decisions (e.g., when available bandwidth is X, do Y).

In our current work, we are initially limiting ourselves to the problem of adaptation among the different video sources in a 3DMC system capturing the scene. Our current model for each camera assumes that each camera is able to produce low-, medium-, or high-resolution imagery and that each camera's view defines a region of interest within the scene. Furthermore, our camera model captures representational dependencies that are the by-product of the underlying compression algorithms (e.g., interframe motion compensation, rate-distortion trade-offs, etc.). Now the problem becomes, given feedback about the user's preference for a region of interest within the scene, how do we allocate bandwidth to the different cameras in order to maximize the utility of the images transmitted?

To address this problem, we are leveraging an abstract

general adaptation framework that we have developed to evaluate available tradeoffs using a utility-cost metric [19]. In this framework, all possible tradeoffs available to the application are mapped as nodes in an N-dimensional space. Each dimension represents a particular axis for adaptation. In 3DMC, we have identified five dimensions for adaptation: one each for time, resolution, relative change of visual content, and two that capture the notion of region of interest. The frames produced by the cameras are represented within this space as nodes. Any encoding dependencies between frames are encoded as directed edges between nodes. The user is represented as a node within the appropriate subspace (i.e., time, and desired region of interest). The user is also associated with a vector that indicates how this position is changing. Adaptation is driven by a utility function that relates the user's position and change vector to the nodes eligible for transmission. Thus, adaptation policy is now very compactly represented as a distance function within the utility space and the evaluation of the adaptation policy becomes mechanical.

At this stage we are experimenting with simple greedy solutions where the node yielding the maximum utility for the minimum cost is selected until resources are exhausted. The process is iterative and the graph evolves as time passes and new frames are produced. We have finished constructing the adaptation framework software and built a simulator in order to experiment with different utility and cost functions. Preliminary experiments show the system is able to make complex, non-trivial adaptation decisions in an emulated eight-camera setup such as in Figure 2. Much of the remaining challenge is to develop and evaluate specific utility functions that correspond to the actual perceived quality of real users.





Fig. 6. *Human patient simulator* staged as an accident victim for the experimental evaluation. Note the 2D video cameras and display mounted on the overhanging structure.

### III. EVALUATION

Previous research in computer supported cooperative work (e.g., [20], [21]) and theory of language [22] suggests that working remotely using 2D video-conferencing lacks the richness of collocation and face-to-face interaction, e.g., multiple and redundant communication channels, implicit cues, and spatial co-references, all that are difficult to support via computer-mediated communications. This lack of richness is thought to impair performance because it is more difficult to establish the common ground that enables individuals to understand the meaning of one another's utterances. Other research [23]–[25] suggests that working remotely may not be compatible with existing work practices, and thus is unlikely to find many adopters. Because we do not know if 3D medical consultation will overcome the limitations found in 2D video-conferencing, or be compatible with existing work practices, there is a need to evaluate the technology and its potential before investing millions of dollars in development.

To evaluate the potential of the technology, we are conducting a controlled experiment that compares the effectiveness and relative advantages of diagnosing and treating an accident victim (mannequin) under three conditions: a paramedic

working alone, a paramedic working in consultation with a remote physician via 2D video, and a paramedic working in consultation with a “remote” physician via a 3D proxy—a physician actually standing nearby. The use of a 3D proxy allows us to investigate the potential benefits and real life issues related to the 3D technology early on, before the technology is fully developed.

The experiment is a post-test design. The task set before each paramedic is the management of a difficult airway for an accident victim. The victim is actually a METI *human patient simulator* that is staged as if at an accident scene (see Figure 6). The paramedic must diagnose and manage a difficult airway to successfully treat the victim. This approach to evaluation has been successfully used in previous research [26]. It allows evaluation of new technology to begin before the technology is fully developed and ready for the field. Early feedback provides valuable insights regarding the future potential and desired features for technology at a lower cost. In addition, with this approach evaluation occurs with no impact on patient care, an important consideration in medicine. The use of realistic tasks and rigorously developed measures helps ensure the validity and reliability of the evaluation results.



In our experiment, effectiveness is operationalized as task performance which is assessed with two measures. One measure is a task/subtask performance protocol based on task/subtask order and quality. The protocol was developed in consultation with physicians specializing in emergency medicine and will be used to assign a numeric task performance score. The second measure, based on the mannequin's blood Oxygen saturation level (O<sub>2</sub>), is the integral of 90%–O<sub>2</sub> over the period when O<sub>2</sub> < 90%. This is based on medical best practice, which recognizes that the longer O<sub>2</sub> is under 90%, the worse off a victim is.

Relative advantage is operationalized as self-efficacy, usefulness and trustfulness of information provided by the remote physician, and compatibility with current ways of working. These data are obtained through a post-session questionnaire.

We are currently running experiment sessions, with the goal of having 20 sessions per condition. It is not possible to perform a power analysis to determine the sample size needed to produce statistically significant results because we do not have an estimate for the effect size. Thus, we reviewed the literature for similar studies and found that a sample size of 10 to 20 was common (e.g., [27], [28]). We anticipate finishing data collection in fall 2006. Early, preliminary results indicate that the usefulness of information provided by the remote physician is significantly higher in the 3D proxy condition.

#### ACKNOWLEDGEMENTS

At UNC-Chapel Hill we want to acknowledge Jim Mahaney and John Thomas for their technical support; and Graduate Research Assistants Adrian Ilie, Vincent Noel, Hua Yang, Michael Noland, and Max Smolens for their contributions to the effort. At the University College of Borås in Sweden we want to acknowledge graduate student Hanna Maurin for her tireless work on the formal evaluation described in Section III.

This effort is primarily supported by National Library of Medicine contract N01-LM-3-3514: "3D Telepresence for Medical Consultation: Extending Medical Expertise Throughout, Between and Beyond Hospitals," and in part by NSF grant EIA-0303590: "Ubiquitous Pixels: Transforming Collaboration & Teaching with Pervasive Wall Displays."

#### REFERENCES

- [1] M. Michael G. Kienzle, "Rural-academic integration: Iowas national laboratory for the study of rural telemedicine," National Laboratory for the Study of Rural Telemedicine, Tech. Rep., 2000.
- [2] M. David G. Ellis, P. James Mayrose, M. Dietrich V. Jehle, M. Ronald M. Moscati, and M. Guillermo J. Pierluisi, "A telemedicine model for emergency care in a short-term correctional facility," *Telemedicine Journal and e-Health*, vol. 7, no. 2, pp. 87–92, 2001.
- [3] S. Tachakra, "Depth perception in telemedical consultations," *Telemed J E Health*, vol. 7, no. 2, pp. 77–85., 2001.
- [4] C. Hendrix and W. Barfield, "Presence within virtual environments as a function of visual display parameters," *Presence: Teleoperators and virtual environments*, vol. 5, no. 3, pp. 274–289, 1996.
- [5] M. P. Snow, "Charting presence in virtual environments and its effects on performance," Ph.D. dissertation, Virginia Polytechnic Institute and State University, December 1996.
- [6] R. Pausch, M. A. Shackelford, and D. Proffitt, "A user study comparing head-mounted and stationary displays," in *Proc. of IEEE Symposium on Research Frontiers in Virtual Reality*. IEEE Press, 1993, pp. 41–45.

- [7] R. Schroeder, A. Steed, A.-S. Axelsson, I. Heldal, A. Abelin, J. Wideström, A. Nilsson, and M. Slater, "Collaborating in networked immersive spaces: As good as being there together?" *Computers & Graphics, Special Issue on Mixed Realities - Beyond Conventions*, vol. 25, no. 5, pp. 781–788, October 2001. [Online]. Available: <http://www.elsevier.com/geom/10/13/20/57/34/32/article.pdf>
- [8] D. W. Mizell, S. P. Jones, M. Slater, B. Spanlang, and D. Swapp, "Immersive virtual reality vs. "flat-screen" visualization: A measurable advantage," (*Submitted for publication.*), 2003.
- [9] R. Yang, "View-dependent pixel coloring—a physically-based approach for 2d view synthesis," Ph.D. Thesis, University of North Carolina at Chapel Hill, 2003.
- [10] R. Yang, M. Pollefeys, H. Yang, and G. Welch, "A unified approach to real-time, multi-resolution, multi-baseline 2d view synthesis and 3d depth estimation using commodity graphics hardware," *International Journal of Image and Graphics (IJIG)*, vol. 4, no. 4, pp. 1–25, 2004.
- [11] W. Fitzmaurice, George, "Situating information spaces ad spatially aware palmtop computers," *Communications of the ACM*, vol. 36, no. 7, July 1993.
- [12] W. Fitzmaurice, George and W. Buxton, "The chameleon: Spatially aware palmtop computers," in *ACM CHI94*, Boston, MA USA, 1994.
- [13] M. Mohring, C. Lessig, and O. Bimber, "Video see-through AR on consumer cell-phones," in *Proc. of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 252–253.
- [14] W. Pasman, A. van der Schaaf, L. Lagendijk, R., and W. Jansen, F., "Accurate overlaying for mobile augmented reality," *Computers and Graphics*, vol. 23, pp. 875–881, 1999.
- [15] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell, "Passive real-world interface props for neurosurgical visualization," in *CHI '94: Proc. of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM Press, 1994, pp. 452–458.
- [16] 3rdTech, "3rdTech," September 2005. [Online]. Available: <http://www.3rdtech.com/>
- [17] "Point Grey Research," September 2005. [Online]. Available: <http://www.ptgrey.com/>
- [18] U. o. W. HIT Lab, "Artoolkit," September 2005. [Online]. Available: <http://www.hitl.washington.edu/artoolkit/>
- [19] D. Gotz and K. Mayer-Patel, "A general framework for multidimensional adaptation," in *MULTIMEDIA '04: Proc. of the 12th annual ACM international conference on Multimedia*. New York, NY, USA: ACM Press, 2004, pp. 612–619.
- [20] P. Dourish, A. Adler, V. Bellotti, and A. Henderson, "Your place or mine? learning from long-term use of audio-video communication," *Computer Supported Cooperative Work*, vol. 5, no. 1, pp. 33–62, 1996. [Online]. Available: [citeseer.nj.nec.com/dourish96your.html](http://citeseer.nj.nec.com/dourish96your.html)
- [21] G. Olson and J. Olson, "Distance matters," *Human-Computer Interaction*, vol. 15, no. 2-3, pp. 139–178, 2000.
- [22] H. Clark, *Using Language*. Cambridge, UK: Cambridge University Press, 1996.
- [23] J. S. Olson and S. Teasley, "Groupware in the wild: lessons learned from a year of virtual collocation," in *Proc. of the 1996 ACM conference on Computer supported cooperative work*. ACM Press, 1996, pp. 419–427.
- [24] W. Orlikowski, "Learning from notes: Organizational issues in groupware implementation," *The Information Society*, vol. 9, no. 3, pp. 237–252, 1993.
- [25] S. L. Star and K. Ruhleder, "Steps toward an ecology of infrastructure: Problems of design and access in large information systems," *Information Systems Research*, vol. 7, pp. 111–134, 1996.
- [26] D. H. Sonnenwald, M. C. Whitton, and K. L. Maglaughlin, "Evaluating a scientific collaboratory: Results of a controlled experiment," *ACM Trans. Comput.-Hum. Interact.*, vol. 10, no. 2, pp. 150–176, 2003.
- [27] C. Gutwin, M. Roseman, and S. Greenberg, "A usability study of awareness widgets in a shared workspace groupware system," in *CSCW '96: Proc. of the 1996 ACM conference on Computer supported cooperative work*. New York, NY, USA: ACM Press, 1996, pp. 258–267.
- [28] A. Lane, "Sectional anatomy: strategy for mastery," in *Proc. of the Visible Human Project Conference*, 1998. [Online]. Available: <http://www.nlm.nih.gov/research/visible/vhpconf98/AUTHORS/LANE/LANE.HTM>