Avoiding Focus Shifts in Surgical Telementoring Using an Augmented Reality Transparent Display

Daniel ANDERSEN[a,] , Voicu POPESCU[a], Maria Eugenia CABRERA[a], Aditya SHANGHAVI[a], Gerardo GOMEZ[b], Sherri MARLEY[b], Brian MULLIS[b], and Juan WACHS[a,1]

[a] *Purdue University*

[b] *Indiana University School of Medicine*

Abstract. Conventional surgical telementoring systems require the trainee to shift focus away from the operating field to a nearby monitor to receive mentor guidance. This paper presents the next generation of telementoring systems. Our system, STAR (System for Telementoring with Augmented Reality) avoids focus shifts by placing mentor annotations directly into the trainee's field of view using augmented reality transparent display technology. This prototype of we tested with pre-medical and medical students. Experiments were conducted where participants were asked to identify precise operating field locations communicated to them using either STAR or a conventional telementoring system. STAR was shown to improve accuracy and to reduce focus shifts. The initial STAR prototype only provides an approximate transparent display effect, without visual continuity between the display and the surrounding area. The current version of our transparent display provides visual continuity by showing the geometry and color of the operating field from the trainee's viewpoint.

Keywords. Augmented reality, telementoring, telemedicine, transparent displays

1. Introduction

Surgical telementoring has the potential to improve surgery outcomes by delivering remote assistance when it is most needed. For example, a rural general surgeon can rely on telementoring to complete a specialized procedure under the guidance of a remote expert surgeon [1]. In another example, a battlefield general surgeon can treat orthopedic trauma injuries under remote specialist guidance in the austere environment of a forward operating base when evacuation is infeasible. Finally, a specialist who invented a novel surgical procedure can rapidly and effectively disseminate the new procedure by providing personalized, interactive guidance to trainee surgeons from a distance.

With a conventional surgical telementoring system, the remote mentor annotates a video feed of the surgery, and then the annotated video is sent back to the trainee to be displayed on a nearby monitor [2]. To view the mentor's annotations, the trainee must first look away from the operating field, view and memorize the annotations, look back

[1]Juan P. Wachs, Ph.D., School of Industrial Engineering, Purdue University, 315 N. Grant Street, West Lafayette, IN, USA 47907; E-mail: jpwachs@purdue.edu.

at the operating field, mentally remap the memorized annotations onto the actual operating field, and finally perform the indicated instruction. Focus shifts repeat many times during the course of a surgical procedure, placing additional cognitive load on the trainee surgeon, which can lead to dangerous delays or errors [3].

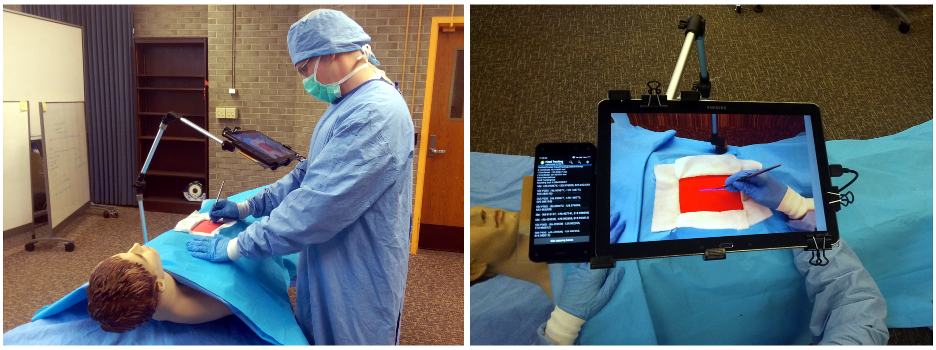

Figure 1. Our telementoring system integrates mentor annotations into the trainee's field of view using an augmented reality transparent display *(left)*. The trainee follow the mentor-suggested incision line *(right)*.

In this paper, we describe our project that aims to eliminate focus shifts in surgical telementoring by displaying the annotations directly into the trainee's field of view. The annotations are shown on a transparent display positioned between the trainee surgeon and the operating field (Figure 1). When viewing the display, the trainee sees the operating field and their hands as if the display were not there. The transparent display is implemented using a tablet enhanced with a depth sensor and a head tracker. Live video of the operating field, acquired by the tablet's camera and enhanced with depth, is reprojected to the trainee's point of view, making the display appear transparent. This approach allows the trainee to see mentor-provided graphical annotations directly overlaid onto the operating field, avoiding focus shifts.

To achieve this aim, we have developed a surgical telementoring system prototype, called STAR (System for Telementoring with Augmented Reality). STAR allows a mentor to create and edit graphical annotations on live video of the operating field, and displays the annotations directly onto the trainee's field of view, anchored to relevant regions of the operating field [4]. Next, we validated our system with a user study which found that our system led to greater accuracy and fewer focus shifts than a conventional system [5]. The first STAR prototype only provides an approximate transparent display effect, without visual continuity between the area of the operating field seen through the display and the surrounding area seen directly. We have developed a transparent display that achieves visual continuity. The transparent display is implemented with a tablet augmented with an on-board depth sensor and head tracker.

2. Methods and Materials

In this section, we describe the architecture of the first STAR prototype (Section 2.1), the first user studies that compare STAR to a conventional telementoring system (Section 2.2), and the tablet-based transparent display (Section 2.3).

2.1. STAR: an augmented reality surgical telementoring system

Figure 2. Our surgical telementoring prototype system, showing the trainee module *(left)*, the mentor module *(center)*, and the trainee's first-person view of the operating field through the trainee module *(right)*.

Figure 2 shows the first STAR prototype. The trainee module (Figure 2, left) is implemented with a tablet held in place between the trainee and the operating field. Live video is captured through the trainee tablet and is sent to the remote mentor site. The mentor uses the mentor module, also implemented with a tablet, to annotate a specific video frame via a touch-based user interface (Figure 2, center). The mentor can create points, lines, text, and icons of surgical instruments. Annotations are transmitted to the trainee module and overlaid onto the video feed of the operating field (Figure 2, right). The annotations are anchored to the surgical points of interest during operating and remain fixed as the tablet is repositioned, or as the operating field deforms or becomes occluded (e.g. when the trainee's hands are between the tablet and the operating field).

Annotations are anchored by first detecting features and extracting descriptors from the reference frame where the annotations were defined. Then, for each subsequent frame in the video feed, features are detected and descriptors are extracted. The current features are matched to the reference features and a homography is computed between the reference and the current frame.

This first STAR prototype only appears approximately transparent. Video frames of the operating field are acquired from the tablet camera's point of view, not the trainee's. As a result, a visible mismatch exists between the part of the operating field shown on the tablet screen, and the surrounding area which the trainee sees directly.

2.2. Precise communication of operating field locations using STAR

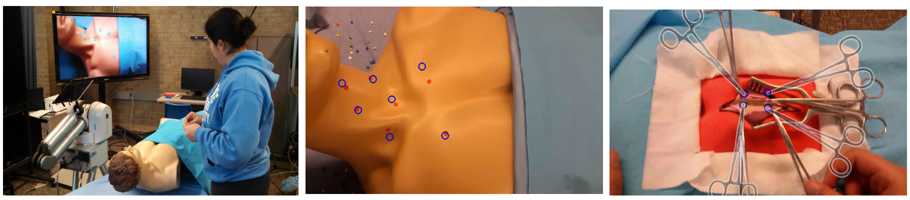

Figure 3. Marker locations indicated by mentor on nearby monitor using a conventional telementoring system *(left)*; marker locations (circles, *center*) and instrument placement (circles, outlined clamps, *right*) indicated by the mentor directly into the trainee's view of the operating field using STAR.

We conducted a user study comparing STAR to a conventional telementoring system that displays the mentor annotation on a monitor near the operating field. The study involved twenty pre-medical and medical students.

Each participant completed two telementored tasks under simulated mentor guidance. The first task was to mark mentor-indicated locations on a patient simulator using adhesive markers (representing the locations of surgical ports). The second task was to complete a simulated abdominal incision using multiple surgical instruments.

Participants were randomly assigned to a Conventional condition, using a conventional telementoring system, or to an AR condition, using STAR for guidance. In the Conventional condition, participants received mentor graphical annotations on a separate monitor near the operating field (Figure 3, *left*). In the AR condition, participants received telementored guidance in graphical form as they were looking at the operating field through the transparent display (Figure 3, *center* and *right*).

We measured the time the participants took to complete the task, the number of times the participants shifted focus away from the operating field, and the placement error defined as the distance between the indicated and actual location of the marker or of the surgical instrument.

2.3. An augmented reality transparent display

In the first version of STAR, the trainee tablet only approximates a transparent display effect. The video acquired by the tablet is displayed as is, without reprojecting it to the trainee's viewpoint. In this section, we describe the process to obtain a true transparent display effect, as shown in Figure 1.

A simulated transparent display effect requires reprojecting the 3D geometry of the real-world scene to the trainee's viewpoint. Therefore, the color and geometry of the operating field, and the trainee's head position, must be captured in real time and transformed to a common coordinate system. This information is acquired using a depth camera and a head tracker mounted on-board with the transparent display.

During rendering, a color image from the tablet's onboard camera is initially acquired. Then, depth data is acquired as either a texture map or a point cloud, depending on the depth sensor used. The head tracker acquires the trainee's current head position.

From the depth map, a mesh is created, that matches the real-world scene's geometry. In the case of holes in a depth map, which can appear in regions beyond the depth sensor's supported range or from dark or specular surfaces, a pull-push process is applied to fill holes with an average of nearby valid depth samples [6]. When depth is acquired as a point cloud, the points are triangulated into a mesh in the display plane using Delaunay triangulation.

The mesh vertices are projected to the video frame to be assigned texture coordinates. The textured mesh is finally rendered from the trainee's viewpoint using the display as the image frame to create the transparency effect.

3. Results

In the following sections, we discuss the results of each of the three main steps described above.

3.1. STAR: an augmented reality surgical telementoring system

The trainee and mentor modules are each implemented using a 12.2-inch Samsung tablet, wirelessly connected to each other. The ORB algorithm in OpenCV was used to anchor annotations to the operating field [7]. The system runs at 12 frames per second.

We tested annotation anchoring robustness during trainee display repositioning (i.e. translation, rotation, scaling) and during operating field occlusion and deformation. We overlaid annotations onto the trainee module's video frames of both an anatomical poster and a surgical dummy. Annotation placement error was measured against ground truth locations. Annotation anchoring was more robust to translation (89-98% success) than to rotation (55%-90%) or zoom (78%-80%). Anchoring was also robust to minor occlusion (96%-100%) but often failed during major occlusion (60-74%) or deformation (15-63%). Our initial homography-based anchoring assumes that the operating field is a planar surface.

3.2. Precise communication of operating field locations using STAR

When using STAR, participants completed telementored tasks with significantly ($P<0.001$) reduced placement error, a 45-68% improvement over the conventional system. Likewise, participants using STAR shifted focus away from the operating field significantly less often ($P<0.003$).

However, participants using STAR completed telementored tasks 19% more slowly on average ($P<0.076$). Several participants performed a slow "searching" motion when placing surgical tools until they had aligned the tool with the mentor annotation. In other words, STAR allowed participants to be more precise, requiring more time. In the conventional condition, the participants were less able to improve over an initial approximate position over which they could not improve.

3.3. An augmented reality transparent display

Figure 4. Our fixed-viewpoint *(left)* and our multiple-viewpoint *(right)* transparent display prototypes.

Two prototype implementations of our transparent display were developed. The first implementation, which supports a transparent effect from a fixed viewpoint, uses the Google Project Tango tablet [8], which has an integrated depth sensor (Figure 4, left). The second implementation, which allows the user to move freely with respect to the display, uses a Samsung tablet with a Structure sensor [9] for depth acquisition and an Amazon Fire Phone [10] for head tracking (Figure 4, right). Our system requires no eyewear or tracking gear to be worn by the user. The transparent display rendering pipeline runs at 30fps. There is some latency (about 100 ms) between motion in the scene or the user's head position, and the view of the scene updating on the screen.

To quantify the transparent display's visual alignment with the surrounding scene, we compared images taken from a user perspective, both with and without the tablet. The relative change in position, rotation, and scale of objects between these images was measured. On average, the translation error was 2.88%, the rotation error was 3.06 degrees, and the scaling error was 4.95%.

4. Conclusion

This paper discussed the work done toward reducing focus shifts during surgical telementoring using an augmented reality transparent display. We have developed a system that allows mentors to create annotations and transmit them directly into the trainee's field of view, anchored to relevant areas of the operating field. The system's ability to reduce focus shifts and improve trainee accuracy was validated through objective metrics. Finally, the sense of a truly transparent display was achieved by acquiring color and depth data and projecting it to the user's viewpoint.

In the future, the transparent display will be integrated into the next version of STAR, to improve the display of graphical annotations into the trainee's field of view. We will also improve the mentor module by incorporating a full-size interaction table gesture interactions. Finally, we will use acquired operating field geometry to overlay multi-stage predictive 3D surgical simulations onto the view of the operating field.

Acknowledgment

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-14-1-0042. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

References

[1] H. Sebajang, P. Trudeau, A. Dougall, S. Hegge, C. McKinley, M. Anvari, The role of telementoring and telerobotic assistance in the provision of laparoscopic colorectal surgery in rural areas. *Surgical Endoscopy and Other Interventional Techniques* **20**(9) (2006), 1389-1393.

[2] Y. Guo, O. Henao, T. Jackson, F. Quereshy, A. Okrainec, Commercial videoconferencing for use in telementoring laparoscopic surgery. *Med. Meets Virtual Real* **21** (2014), 147.

[3] E.M. Bogen, K.M Augestad, H.R. Patel, R.O. Lindsetmo, Telementoring in education of laparoscopic surgeons: An emerging technology, *World Journal of Gastrointestinal Endoscopy* **6**(5), (2014), 148-155.

[4] D. Andersen, V. Popescu, M.E. Cabrera, A. Shanghavi, G. Gomez, S. Marley, B. Mullis, J. Wachs, Virtual annotations of the surgical field through an augmented reality transparent display, *The Visual Computer* (2015), 1-18.

[5] D. Andersen, V. Popescu, M.E. Cabrera, A. Shanghavi, G. Gomez, S. Marley, B. Mullis, J. Wachs, Medical telementoring using an augmented reality transparent display. Unpublished manuscript. Retrieved from https://engineering.purdue.edu/starproj/wp-content/uploads/2015UserStudy.pdf

[6] S.J. Gortler, R. Grzeszczuk, R. Szeliski, M.F. Cohen, The lumigraph, *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*. ACM, (1996), 43-54.

[7] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, ORB: an efficient alternative to SIFT or SURF. *IEEE International Conference on Computer Vision*, (2011), 2564-2571.

[8] Google, Project Tango (2015). Retrieved from https://www.google.com/atap/project-tango/.

[9] Occipital, The Structure sensor is the first 3D sensor for mobile devices (2014). Retrieved from http://structure.io/.

[10] Amazon.com, Amazon Fire Phone (2014). Retrieved from http://www.amazon.com/firephone.