A First-Person Mentee Second-Person Mentor AR Interface for Surgical Telementoring

Chengyuan Lin*[1], Daniel Andersen[1], Voicu Popescu[1], Edgar Rojas-Muñoz[1], Maria Eugenia Cabrera[1], Brian Mullis[2], Ben Zarzaur[2], Kathryn Anderson[2], Sherri Marley[2], and Juan Wachs[1]

[1]Purdue University
[2]Indiana University School of Medicine

Abstract

This application paper presents the work of a multidisciplinary group of designing, implementing, and testing an Augmented Reality (AR) surgical telementoring system. The system acquires the surgical field with an overhead camera, the video feed is transmitted to the remote mentor, where it is displayed on a touch-based interaction table, the mentor annotates the video feed, the annotations are sent back to the mentee, where they are displayed into the mentee's field of view using an optical see-through AR head-mounted display (HMD). The annotations are projected from the mentor's second-person view of the surgical field into the mentee's first-person view. The mentee sees the annotations with depth perception, and the annotations remain anchored to the surgical field entities they describe as the mentee move their head. Average annotation display accuracy is 1.22cm. The system was tested in the context of a user study where participants ($n = 20$) were asked to perform a lower-leg fasciotomy on cadaver patient models. Participants who benefited from telementoring using our system received a higher Individual Performance Score, and they reported higher usability and self confidence levels.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality;

1 Introduction

As surgery continues to specialize more and more narrowly and deeply, it will be more and more challenging to always provide all needed surgical expertise at all points of care. Surgical telementoring is a promising approach for transmitting surgical expertise over large geographic distances promptly and efficiently. Consider a rural surgery center that is staffed with a general surgeon, but with no subspecialty surgery experts. An expert surgeon from a major hospital could "virtually scrub in" to assist with a procedure that the general surgeon is not entirely comfortable performing alone. Consider the scenario when a critical patient cannot be transported urgently to a facility where the required surgical expertise is available. This could be the case, for example, in a combat zone where a compartment syndrome relieving fasciotomy procedure has to be performed urgently at a forward operating base to save a patient's leg, and evacuating the patient is too slow or too dangerous. An orthopaedic trauma surgeon from a major military hospital could assist from thousands of miles away via telementoring. As a third example, a novel surgical procedure can be more rapidly disseminated through surgical telementoring, than by having the surgeon who invented the procedure travel around the world. Finally, telementoring could also benefit surgical training, with a single instructor working in parallel with multiple surgical residents, providing assistance on demand, to the trainees who need it.

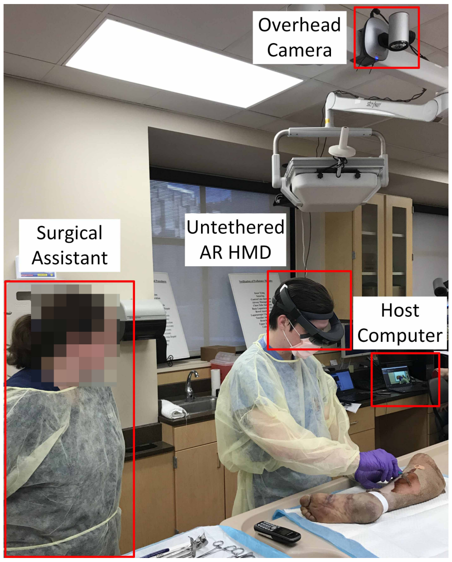
The conventional approach for surgical telementoring is based on a telestrator that allows a remote mentor to annotate graphically a video feed of the surgery, much the same way a sports broadcast commentator can annotate a play sequence. The annotated video feed of the surgery is shown to the mentee on a nearby display. This requires the mentee to shift focus away from the surgery, and to map mentally the instructions from the nearby display to the surgical field, which can lead to surgery delays and even errors. Augmented Reality (AR) is a promising alternative for surgical telementoring because it allows to integrate the mentor-authored annotations directly into the field of view of the mentee. The mentee sees the annotations as if the mentor actually drew them onto the surgical field, which avoids focus shifts and the high cognitive load of having to map annotations to the surgical field.

One possible AR interface for surgical telementoring is a transparent display placed between the mentee and the surgical field and that shows the mentor annotations. However, truly transparent displays are not yet available, with the exception of some low transmittance experimental displays. Video see-through transparent displays simulate transparency by showing the real world scene with the help of a video camera. Such a display supports only monoscopic viewing of the surgical field, which reduces depth perception and can decrease surgical performance. Both optical and video see-through displays pose the challenge of work-space encumbrance, as the surgeon has to reach around the display that is placed between them and the surgical field. An alternative AR interface is an optical see-through AR head-mounted display (HMD). The AR HMD avoids work space encumbrance and it allows the mentee to see the surgical field directly, with natural depth perception.

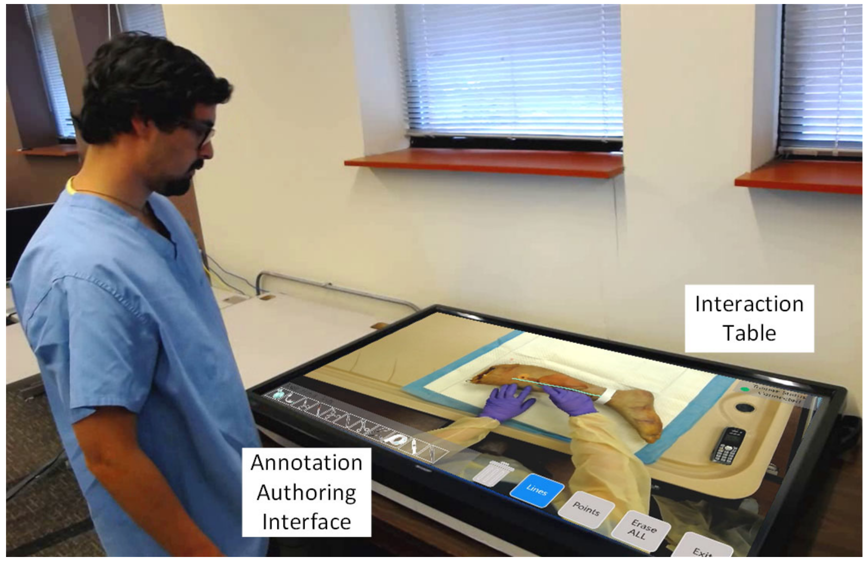
We are a group of computer science researchers, industrial engineering researchers, trauma surgeons, orthopaedic trauma surgeons, and surgery educators. In this application paper we describe a novel system for surgical telementoring based on an AR HMD, as well as an initial evaluation of the system in the context of a user study where surgery residents performed lower-leg fasciotomies on cadaver patient models.

Fig. 1 gives an overview of our system. The surgical field is acquired at the mentee site with an overhead camera whose feed is sent to the remote mentor site. The overhead feed is displayed on a custom full-size patient touch-based interaction table that allows the mentor to annotate the surgical field with touch based gestures. The annotations are sent back to the mentee site where they are integrated into the mentee's view of the surgical field using an AR HMD worn by the mentee. The annotations are converted from 2D to 3D by projection from the overhead camera view, where they were authored, to the 3D geometry of the surgical field acquired by the AR HMD. This way, the remote mentor can annotate the surgical field in real time, and the annotations are shown to the mentee with

*lin553@purdue.edu



(a) Mentee subsystem



(b) Mentor subsystem

Figure 1: Our telementoring system based on an AR HMD at the mentee and on a full-size touch-based interaction table at the mentor.

correct depth perception, anchored to the surgical field entities that they describe. We call our AR interface first-person for the mentee, as the annotations are directly integrated into the mentee's field of view, and second-person for the mentor, since the mentor sees the surgical field and authors annotations in the fixed view of the overhead camera.

We have measured annotation display accuracy by drawing a dot A on a patient simulator (i.e. a mannequin), by asking the mentor to place a virtual dot annotation at A, and then by asking the mentee to draw a dot B on the mannequin where they see the virtual annotation. The annotation display error achieved by our system is the distance between dots A and B. The maximum/average annotation error achieved by our system over several dots that spanned the mannequin is 1.60cm and 1.22cm, respectively.

We have conducted a user study to test our system with fourteen surgery residents and six medical students, who were asked to perform a lower-leg fasciotomy on a cadaver patient model. The participants were assigned to two groups: a control group (CG), which performed the fasciotomy after studying the procedure from printed surgery course materials, and an experiment group (EG), which performed the fasciotomy under telementoring guidance provided with our system. Participant performance was rated by an expert surgeon who witnessed the procedure and quantified performance using an Individual Procedure Score (IPS) metric. The *EG* participants received an IPS score 16% higher than the *CG* participants. The two groups were also evaluated using a system usability questionnaire. The answers to all eight questions indicate a usability advantage for our system, and for four of the questions the advantage was statistically significant. Finally, the two groups were also evaluated based on a self-reported confidence in knowledge of fasciotomy before and after performing the procedure. The *EG* group showed a statistically significant growth for all four confidence metric questions, and they ended up with a higher confidence level than the *CG* group.

We also refer the reader to the accompanying video that illustrates the operation of our system and the user study we have conducted.

2 Prior Work

The number of surgeons with specialized expertise is limited and with an uneven geographic distribution. In rural areas, developing countries, and combat areas, specialized surgical expertise is typically not available. When a patient needs urgent care, transporting them to a health care facility where the required surgical expertise is available might be unfeasible [10]. Telementoring has been shown to be helpful in such scenarios by allowing remote surgeons with the needed expertise to provide guidance to local surgeons [14]. Combat injuries typically affect more than one organ system, and adequate care requires surgical expertise in a large number of subspecialty areas. Furthermore, it is not always possible to evacuate the patient to a medical facility, so surgical telementoring is particularly useful in these cases [11]. Beyond the emergent need for surgical expertise, telementoring was shown to be useful for learning new surgical skills, by avoiding the high time costs of having the expert instructor actually travel to all the hospitals where instruction is needed [27]. Telementoring can also be used for monitoring the quality with which a particular surgical procedure is performed [3, 28].

The conventional approach for surgical telementoring is based on a telestrator. The live video feed of the surgical field is transmitted to the remote mentor, who annotates it, the annotations are sent back to the mentee, and the annotated video is shown to the mentee on a nearby display [5]. Both 2D annotations, authored and shown in an image of the surgical field, and 3D annotations, authored and shown in the 3D space of the surgical field, have been investigated with telestrator systems. 2D annotations are widely used due to their simplicity [12, 25], but they are not naturally seen by the mentee due to the lack of depth perception, due to the lack of parallax, and due to occasional occlusions. 3D telestration was developed and tested using a da Vinci surgical visualization console [1]. By acquiring the surgical field with a binocular camera and with the help of stereo vision algorithms, 3D annotations are created and shown to the mentee, and the 3D annotations lead to lower error rates, compared to 2D annotations. Although annotations do not remain anchored to the surgical field as the user viewpoint changes, 3D telestration is a first step towards a true AR interface for surgical telementoring. However, telestrator techniques have inherent shortcomings. One shortcoming is the need for the trainee to shift focus repeatedly from the surgical field to the nearby display. Each time, the mentee has to remember the position and type of individual annotations, and then to map from memory the annotations to the actual surgical field. These focus shifts increase the cognitive load of the mentee, which can translate to surgery delays or even surgical errors [5].

The problem of focus shifts can be eliminated by directly inserting the graphical annotations into the mentee's field of view. Therefore, surgical telementoring can greatly benefit from AR interfaces, which can provide a natural approach for overlaying annotations onto the surgical field, as if the mentor actually drew them there. The potential of AR in surgery has been noted and investigated for a long time [26]. The recent leap forward of AR technology has intensified anew research efforts aimed at bringing AR in the operating room.

There are two major options for designing the AR interface: based on a display (e.g. computer tablet) interposed in between the mentee and the surgical field, and based on an AR HMD [8]. In previous work we have explored the tablet option [2]. A video-see through display, implemented by a computer tablet, was suspended above the surgical field. The camera built into the tablet acquires the surgical field, the video feed is sent to the mentor, and the mentor uses a touch-based interface to enhance the surgical field with annotations such as lines, text, and surgical instrument icons. The annotations are transmitted back to the trainee site, where they are shown on the tablet, superimposed onto the live view of the surgical field. The trainee can then follow the instructions from the mentor to complete the surgery, without having to switch focus away from the surgical field. Compared to a conventional telestrator system, a user study revealed that our system led to 57% smaller surgical port and instrument placement errors, and to 65% fewer focus shifts. One of the shortcomings of such a tablet-based AR interface is the lack of depth perception that ensues from the monoscopic visualization of the surgical field. This means that the mentee has to move their hand slowly in search of the correct depth where the tip of an instrument makes contact with the surgical field. This problem is encountered in all scenarios where the surgeon sees the operating field through a video feed, such as in laparoscopic or endoscopic procedures. A second important shortcoming of tablet-based AR interfaces is the workspace encumbrance brought by the tablet, which might require the mentee to deviate from their preferred arm and hand poses and motions during surgery.

In this paper we investigate the use of an optical see-through AR HMD interface, which has the potential to address these shortcomings. The mentee sees the surgical field directly, with natural depth perception, and the annotations are drawn in 3D, with correct parallax between the left and right eyes, which provides depth perception. Furthermore, the HMD does not interfere with the mentee's arm motions. Prior work investigations of the use of AR HMD interfaces in the operating room have found benefits in the context of overlaying a static image or model onto the patient [4, 29], or of overlaying a visualization of patient specific data acquired with an imaging system [19].

3 AR HMD STAR PLATFORM

The goal of surgical telementoring is to allow the mentee to see the mentor-authored annotations naturally, as if the mentor actually drew them on the patient. In order to achieve this, we need to show the mentor the surgical field, allow the mentor to annotate it, and then to integrate the annotations into the mentee's field of view of the surgical field. We first discuss the design of the AR interface at the mentor and mentee that enables telementoring (Sect. 3.1), and then we give an overview of the calibration (Sect. 3.2) and operation (Sect. 3.3) of our system that implements the designed AR interface.

3.1 AR Interface Design

We developed the AR interface of our surgical telementoring system based on the following design considerations. First, we wanted the mentee to see the mentor authored annotations directly overlaid onto the surgical field. This was satisfied by using an AR interface, that allows augmenting the view of the real world with anchored computer graphics annotations. The second consideration was to provide the mentee with depth perception, so the annotations should be displayed with accurate interpupillary parallax, in 3D. This was satisfied by resorting to an optical see-through AR HMD, with which the mentee can see the real world directly and which visualizes the annotations stereoscopically. A video see-through display might show annotations in stereo when the scene is acquired with two cameras, but field of view, vergence, and accommodation issues preclude seeing the real world as effectively as with an optical see-through AR HMD. The third consideration was to avoid encumbering the mentee workspace with hardware in between the mentee and the surgical field. This consideration reinforced our choice for an HMD AR interface, as opposed to, for example, interposing a computer tablet in between the mentee and the patient.

The fourth consideration is to provide the mentor with appropriate situational awareness, as they need to provide prompt and detailed instructions to the mentee. It is clear that the mentor needs to benefit at least from a video stream of the surgical field. Our first attempt was to acquire this video feed with the on-board camera already built into the AR HMD. This way, the mentor would see the mentee's first-person view. However, in our preliminary tests that used this configuration, it became apparent that conveying the surgical field to the mentor through the mentee's first person view has two important challenges. First, such a visualization can be ineffective, as the visualization changes for the mentor drastically as the mentee changes view direction. For example, every time the mentee looks at their surgical assistant or instrument tray, the mentor loses sight of the surgical field, which is particularly disconcerting when the mentor is in the process of trying to decide where to place an annotation, or, even worse, of actually drawing one. The mentor had to repeatedly ask the mentee to stand still, which is an unacceptable interference with the mentee's performance of the surgery. In other words, the mentor should not use the mentee as a servo mechanism for orienting the remote camera to their preference. Second, the rapid and surprising change of view can disorient the mentor, and even induce nausea [13]. To avoid these problems, we decided to deploy an external overhead camera that captures the surgical field from a stationary position above the surgical field.

In conclusion, our design settled on a first-person AR interface for the mentee and a second-person AR interface for the mentor, which we implemented in a surgical telementoring system, and which we tested in a lower-leg fasciotomy user study.

3.2 System Calibration

Fig. 2 shows an overview of the architecture of our surgical telementoring system (Fig. 1). We describe our system using the $\xi_{A,B}$ notation for the $SE(3)$ transformation [6] from a coordinate system A to a coordinate system B.

There is a one-time calibration process after which the system becomes operational. We use an untethered, self-tracking AR HMD, which, for every frame, provides the position and orientation of the HMD with respect to the world. The goal of the calibration stage is to determine the pose $\xi_{oc,w}$ of the overhead camera (OC) in the world coordinate system (W) of the AR HMD. Our AR HMD has a built-in video camera which we leverage for this calibration process. We use a standard calibration procedure [32] that first calibrates the intrinsics of the overhead camera and of the AR HMD built-in camera using a checkerboard pattern. Then the overhead and built-in camera extrinsics are found by showing the checkerboard to both cameras simultaneously (Fig. 3). The overhead camera sends its image to the host computer (c1 in Fig. 2), where the checker corners are detected and the pose $\xi_{oc,cp}$ relative to the checkerboard pattern (CP) is computed by solving a perspective-n-point problem [9]. The pose of the AR HMD relative to the checkerboard pattern $\xi_{hmd,cp}$ is computed similarly. $\xi_{oc,cp}$ is sent to the AR HMD (c2), where the pose of the overhead camera $\xi_{oc,w}$ is finally computed with the following concatenation of transformations (Equation 1), where

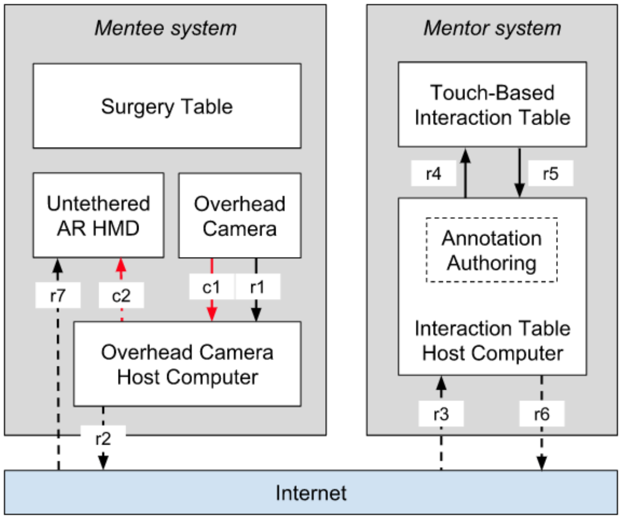


Figure 2: System diagram. Solid and dotted arrows correspond to wired and wireless communication, respectively. Red illustrates system calibration, and black illustrates system operation.

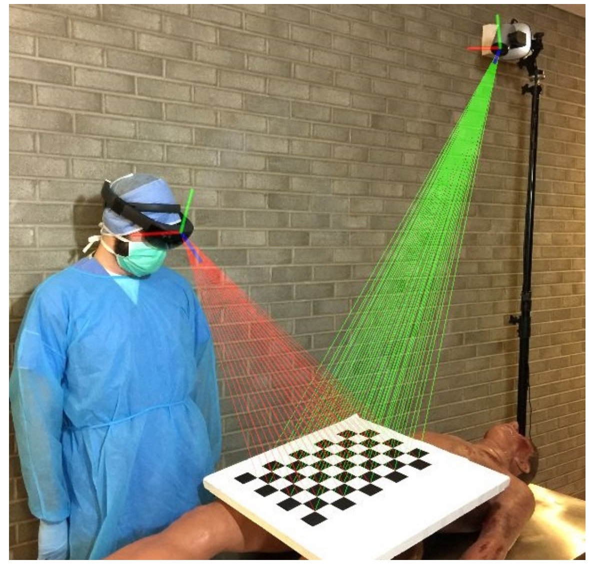


Figure 3: Calibration process. The overhead camera (green ray visualization) is registered with respect to the camera built into the AR HMD (red rays) using a standard calibration checkerboard.

$\xi_{hmd,w}$ is the HMD pose tracked for the frame that captures the checkerboard pattern. $\xi_{oc,w}$ is stored on the AR HMD and used during operation to visualize the mentor annotations.

$$\xi_{oc,w} = \xi_{oc,cp} \cdot \xi_{hmd,cp}^{-1} \cdot \xi_{hmd,w} \tag{1}$$

3.3 System Operation

The overhead camera captures a live video feed of the surgical field (r1), which is sent to the remote mentor via the Internet (r2). The feed is received at the mentor subsystem (r3), where it is displayed on the touch-based interaction table (r4). The mentor examines the surgical field, zooms in (digitally) and pans the view, and authors annotations as needed using touch-based gestures. The annotation authoring commands are collected (r5) and sent to the mentee subsystem via the Internet (r6). The AR HMD is connected to the Internet and directly receives the annotation commands (r7), which it uses to show the annotations to the mentee in 3D.

Since the annotations were authored by the mentor in 2D, in the overhead camera video feed, the 2D annotations have to be converted to 3D annotations suitable for stereo AR HMD visualization. Annotations have to be anchored to the surgical field entities that they describe. For example, an incision line has to be drawn at the correct depth such that it appears to actually touch the patient surface. In another example, a text label annotation should be displayed in a flag that is above the surface of the patient, with the flag planted exactly at the surface point the label describes. In a third example, a scalpel could be shown at the proper angle with the incision surface, with the blade touching the surface. All annotations have one or more points of contact with the surgical field geometry.

Given a 2D annotation point p in the overhead camera image plane, its 3D position P in the world coordinate system is computed by unprojection to the overhead camera ray r_{oc}, by transforming the ray to world coordinates $r_W = \xi_{oc,w} r_{oc}$, and by intersecting the ray with the surgical field geometry G, i.e. $P = r_w \cap G$. We approximate G with the coarse geometric model of the scene acquired by our AR HMD. Fig. 4 illustrates the process of mapping 2D authored annotations to 3D by projection onto surgical field geometry along overhead camera rays.

Figure 4: Annotation projection. The incision line, the scalpel tip, and the textual label stem tip are projected from the overhead camera perspective onto the geometry of the surgical field. The incision line lies on the patient, whereas the scalpel and the label annotations float above the patient.

4 Results and Discussion

We implemented our system using a Microsoft Hololens AR HMD [20] which has the advantages of being untethered, allowing the mentee to move freely, of having a built-in video camera, allowing for overhead camera calibration, of self-tracking, allowing annotation anchoring as the mentee moves, and of acquiring a geometric proxy of the scene, allowing for annotation projection. The active part of the HoloLens display has a 1268 by 720 resolution and it is refreshed at 60fps. An important shortcoming of the HoloLens AR HMD is the small field of view of the active part of the display (i.e. about 30 by 17.5 degrees), which limits the region of the mentee's field of view that can be annotated. The overhead camera is a Logitech PTZ Pro 2 [15], acquiring 1920×1080 pixel frames at 30fps. The camera has a variable focal length, which is set before calibration to show all and only the relevant part of the surgical field. Audio communication between the mentor and the mentee was provided with a conventional phone in speaker mode. The interaction table at the mentor was built from a touch-screen Sharp LCD (1920×1080 resolution, 60 fps, physical size of 52.3×29.4 inches) [24] positioned horizontally, that provides multi-touch interaction.

Figure 5: Annotation accuracy measurement on patient simulator (i.e. mannequin). Black dots are drawn with a dry erase pen on the patient simulator, the mentor creates annotations at these black dots, and the mentee draws green dots on the patient simulator with a pen where they see the annotations. The annotation display error is given by the distance between a pair of black and green dots.

Figure 6: *EG* participant in the fasciotomy user study. The virtual incision line and instruments are only seen by the participant, and they were added here for illustration purposes.

We first discuss system performance based on technical metrics (Sect. 4.1), then we describe a user study where we tested our system in the context of fasciotomy telementoring (Sect. 4.2), and we end the section with a discussion of the limitations of our system (Sect. 4.3). We also refer the reader to the accompanying video that illustrates the operation of our system and the user study we have conducted.

4.1 System Performance

One important aspect of our real-time visual communication system is latency. One latency is the delay with which the overhead camera video feed is transmitted from the mentee site to the mentor site. This latency depends on the distance between the two sites, the network state (i.e. ping times), and video encoding and decoding efficiency. We have measured ping times from 50ms within our Purdue servers, to over a second from Purdue to universities in South-East Asia and Australia. We use the WebRTC platform [30] to encode the overhead camera feed on the mentor's site host computer, and to decode the feed on the mentee's site host computer, which is done with negligible delay. In our experiments network bandwidth was not a concern as it was sufficient to transmit the overhead camera feed at full resolution with levels of compression that did not affect video quality. Another latency is the delay between the mentee head movement and the required repositioning of annotations, which for our AR HMD is an almost unnoticeable 16ms. In other words, when the mentee moves their head, the annotations appear stationary in the 3D world, and do not "follow" the mentee's view.

There is no inaccuracy in the visual communication from mentee to mentor. For example, if the mentee points at a surgical field location with their finger, the mentor will see the correct location in the overhead camera video feed. However, there is inaccuracy in the mentor to mentee communication.

We have analyzed the annotation display error, as witnessed by the mentee. This error accumulates from the combination of camera calibration error, surgical field geometry approximation error, HMD fitting error, and head tracking error. Intrinsic and extrinsic calibration of the overhead and built-in cameras was done with a checkerboard corner average reprojection error of 0.4pixels. The surgical field geometry approximation error and the user head tracking errors are not under our control, as they are provided by the AR HMD. We measured the error of the Hololens in terms of tracking using a checkerboard pattern. The average error is 2° and 2cm for rotation and translation, respectively. The HMD fitting error stems from the variable user head geometry, which makes the HMD sit in different positions and orientations with respect to the user's eyes. The AR HMD provides a transformation from the on-board camera to the user's left and right eye, which we apply during rendering, however this generic transformation is only an approximation of the true transformation needed for every user.

We have measured the annotation display error witnessed by the mentee empirically, by placing a physical marker A in the surgical field, asking the mentor to annotate the position A of the marker in the overhead camera feed, and then by asking the mentee to place a second physical marker B at the location where they see the annotation A drawn. The annotation error is the distance between markers A and B which we measure with a measuring tape. In Fig. 5 A and B markers are the black and green dots, and the maximum and average annotation display error is 1.60cm and 1.22cm, respectively.

As Fig. 5 shows, the direction and magnitude of the offset between green and black markers is rather consistent, so much of the annotation display error is systematic. We have devised an optional additional calibration procedure that improves annotation display accuracy under the assumption that most of the systematic annotation display error is due to an consistent overestimation of scene geometry by the AR HMD. Indeed, using the built-in Kinect-like depth camera, the AR HMD builds an approximate geometric model of the scene that consistently overestimates scene geometry, by wrapping a coarse geometric mesh over the actual detailed geometry. The additional calibration procedure is based on interaction between mentor and mentee. The mentor places an annotation and then asks the mentee to place and hold their index where they see the virtual annotation. The annotation display error is apparent to the mentor in their overhead camera view as a distance between the mentee's finger tip and where the mentor drew the annotation. Using this visualization, the mentor shifts the approximate geometric model of the surgical field to reduce the annotation display error.

4.2 User Study

We have conducted a user study at the Indiana University School of Medicine with $n = 20$ participants: 14 surgery residents and 6 medical students. The *task* was a four-compartment release by dissecting lower-leg fascia on cadaver models. Such a fasciotomy intervention is an emergency procedure for treating compartment syndrome, which is a lack of blood circulation to the limb due to excessive swelling as the result of blunt trauma. If left untreated, compartment syndrome leads to the loss of the affected limb. Fasciotomies remain challenging surgical procedures. In a recent systematic review on the surgical management of chronic exertional compartment syndrome, the overall success rate was reported at 66%, the satisfaction rate was 84%, and the rate of return to previous or full activity was 75% [7]. Furthermore, symptom recurrence was up to 44.7%, reoperation rate up to 19%, and overall complication rate was 13%.

Participants were randomly assigned to one of two groups: a control group (CG), which received instruction on how to perform the

Figure 7: *CG* participant in the fasciotomy user study.

fasciotomy from an illustrated brochure, i.e. the Advanced Surgical Skills for Exposure in Trauma [22] course material on fasciotomies, and an experiment group (EG), which received real-time guidance with our telementoring system. The *EG* group did not receive any fasciotomy instruction prior to actually performing the procedure. Fig. 6 and Fig. 7 show a participant in the experiment group and control group, respectively. The additional interactive calibration procedure was performed by the mentor with each mentee, as the procedure depends on the actual surgical field geometry, and the cadaver lower leg models had great shape and size variability.

The two groups were compared based (1) on expert rating, (2) on self-reported usability, (3) on self-reported confidence in procedure knowledge, and (4) on procedure completion time. To analyze the data, we first check the data normality assumption using the Shapiro-Wilks test [23] and in our case no data was normal. For the unpaired (between subject) data (1, 2 and 4), we use the Mann-Whitney *U* test [17] to test for statistical significance. For the paired (i.e. within subject) data (3), statistical significance is tested with the Wilcoxon signed-rank test [31].

(1) An expert surgeon evaluated the performance of each participant during and after the experiment using the Individual Procedure Score metric [16], which we adapted to fasciotomy. IPS is a test that assesses whether a training course is being effective on improving the overall surgical expertise of a participant. The test includes an objective analysis of the participants execution of the required procedural steps, as well as a subjective analysis to identify any errors that occur during procedure execution. *EG* participants received a median IPS of 81.15 with an interquartile range of ± 23.25, which was 16% higher than for *CG* participants (69.55 ± 33.40). The interquartile range is defined by the score received by the 25th percentile participant and the 75th percentile participant, and was used here as the data pointed to non-normality. However, the greater *EG* IPS scores were not statistically significant ($p = 0.26$).

(2) The two groups were compared based on self-reported usability through a five-level Likert scale questionnaire (Table 1). *EG* participants reported a higher preference for their condition than *CG* participants. For four out of the eight questions, the difference was statistically significant.

(3) The two groups were also compared in terms of self-reported confidence in performing a fasciotomy procedure. Table 2 and Table 3 report the increase in participant confidence level from before to after the experiment, for *EG* and *CG* participants, respectively. The confidence scores are assigned on a scale from 1 to 5. *EG* participants reported a statistically significant improvement in all four confidence categories, whereas *CG* participants reported statistically significant improvements in only half of the categories. Table 4 and Table 5 in Appendix A provide the initial and final confidence levels, for the two participant groups. The *CG* participants were more confident than the *EG* participants in their knowledge of the procedure before the task, but *EG* participants were more confident after the task.

(4) *EG* participants completed the procedure marginally faster (i.e. 4% faster, 1,379s median completion time with a ± 380s interquartile range) than *CG* participants (1,444s ± 685s).

This first study indicates that our AR surgical telementoring has the potential to provide surgical expertise remotely in an effective way. Not all advantages detected are statistically significant. One reason is the great variability and low number of participants. Another reason is that the remote mentor was a faculty member overseeing the surgery residency program, who was known to the participants,

Table 1: Self-reported support method usability. P-values with an asterisk (*) represent a statistically significant difference between the two groups. For questions 6 and 8, a lower score is indicates a higher preference.

Question	EG	CG	p-value
[1] Sufficient information provided	5.0 ±1.00	4.0 ± 0.50	0.024*
[2] Instructions easy to follow	5.0 ± 1.00	4.0 ± 1.25	0.018*
[3] Instructions conveyed effectively	4.0 ± 1.25	4.0 ± 1.00	0.415
[4] Cleared procedure doubts	4.0 ± 1.25	3.0 ± 1.50	0.063
[5] Expedited procedure completion	5.0 ± 2.25	3.5 ± 2.25	0.111
[6] Generated frustration	2.0 ± 1.25	3.0 ± 2.00	0.037*
[7] Better than side-by-side mentoring	2.0 ± 2.00	2.0 ± 1.00	0.139
[8] Worse than side-by-side mentoring	2.5 ± 2.25	4.0 ± 2.00	0.028*

Table 2: *EG* participant self-reported confidence scores. All p-values report a significant improvement.

Confidence Assessment Aspect	Self-Reported Confidence Difference	p-value
Identify anatomical landmarks	1.0 ± 1.25	0.014*
Knowledge of procedural steps	1.0 ± 1.00	0.006*
Instrument handling technique	1.0 ± 1.25	0.014*
Perform procedure alone	1.5 ± 1.00	0.006*

Table 3: *CG* participant self-reported confidence scores. p-values with an asterisk (*) represent a statistically significant improvement.

Confidence Assessment Aspect	Self-Reported Confidence Difference	p-value
Identify anatomical landmarks	1 ± 1.00	0.022*
Knowledge of procedural steps	1 ± 2.00	0.036*
Instrument handling technique	0 ± 1.00	0.225
Perform procedure alone	1 ± 0.25	0.11

which added significant performance pressure on *EG* participants, whereas *CG* participants worked without the pressure of being evaluated by one of their professors. Furthermore, the telementoring sessions turned into practical lessons of surgery, which included revisiting of fundamental concepts in anatomy and in surgical procedures. This was of course not the case for *CG* participants. Not counting the tangential teaching mixed in with fasciotomy telementoring is difficult to do objectively, but it is likely to reduce the overall procedure completion times considerably for *EG* participants.

4.3 Limitations

Both the mentee and the mentor complained occasionally that the annotation showing the incision line would obstruct the view of the actual incision, as the incision progressed as it was executed. The mentor has the ability to remove an annotation completely, but was is needed is to erase the annotation gradually as the incision progresses. Another possible solution for this problem that we will explore in a future study is to ask the mentee to transfer the annotation on the actual skin of the patient with a surgical marker before actually performing the incision.

Another limitation of our system is that annotations do not stick to the surgical field as it deforms. Our AR HMD does acquire surgical field geometry continually, but the geometry is updated relatively infrequently, i.e. only once the geometry has reached another stable configuration. For example, when an object moves from one location to another, it takes a few seconds for the geometric mesh corresponding to the old position of the object to be erased and for the mesh for the new position to be created. An additional real depth camera, e.g. a Microsoft Kinect [21], could capture surgical field geometry in real time with great fidelity, which opens the door to accurate dynamic annotation anchoring.

Another limitation of our system is that the AR HMD is not very bright, and annotations appear faint when the background is brightly lit, as it is the case of surgical fields illuminated by powerful lights. For this user study we turned off the surgical lights, which also interfere with the quality of the overhead camera video feed, which has limited dynamic range. Low annotation brightness is a fundamental limitation of optical see-through AR HMDs, which draw semi-transparent annotations on top of the user's direct view of the real world. A video see-through AR HMD allows for good control of the user's view of the real world, enabling opaque annotation pixels that completely erase the real world pixels, no matter how bright these are. Of course, the video see-through has the disadvantage of a smaller field of view and of an unnatural visualization of the real world scene, which can potentially hinder hand eye coordination. Since surgeons are used to operate by seeing images of and not the actual real world (e.g. laparoscopic, microscope, and endoscopic procedures), comparing an optical to a video see-through AR interface is important future work.

Our system inherits additional limitations of the AR HMD, such as a small field of view of the active part of the display, which confines annotation display to the center of the mentee's field of view. Another limitation is the poor ergonomics of operating with a heavy and sometimes poorly fitting contraption attached to one's head. Several participants reported back and neck strain, especially the ones with little surgical experience who would tilt their head forward, moving the weight of their head and of the display away from their body.

5 Conclusions and Future Work

In this application paper we have presented the design and implementation of a surgical telementoring AR interface, and we have validated our system in a user study where participants performed a cadaver-leg fasciotomy under telementoring. Our system promises surgical telementoring benefits, although not all benefits measured were statistically significant in this initial study.

In addition to the future work possibilities mentioned above, another direction of future work is to improve the mentor's sense of presence in the operating room. One option is to do away with the overhead camera and to rely on the video feed already acquired by the AR HMD from the mentee's viewpoint. As discussed in Sect. 3.1, one challenge is to stabilize this other-first-person view, borrowing from prior work on video stabilization [18]. This not only simplifies the system, but also potentially increases the accuracy of the annotations, by authoring annotations in a view similar to the one from where they will be seen. Indeed, if the mentee assumes the same viewpoint as the viewpoint of the frame that was annotated by the mentor, the annotation can be shown to the mentee in the correct location without knowledge of the surgical field geometry. Another option for improving the mentor's situational awareness is to upload to the mentor not only a video feed of the surgery, but actually an RGBZ stream of frames with per pixel depth, which allows the mentor to choose their viewpoint interactively, to draw annotations more accurately in 3D (e.g. a non-planar incision curve), and even to visualize the surgical field immersively, e.g. with a Virtual Reality headset.

Telementoring could also benefit from extending the types of annotations supported with the ability to send a visual depiction of the mentor's hands, as surgical instruction includes mid-air gestures that sketch, for example, the use of an instrument. We foresee that the quickest path to achieving this is to capture the mentor hands with a video stream, to segment them, and to display them at the mentee. Placing the hands in the correct position with respect to the surgical field, and showing them from the mentee's viewpoint will require capturing the mentor hands with an RGBZ stream.

Our current surgical telementoring system relies on a high-quality network connection between the mentee and mentor sites, which is not always available in the case of austere environments. For this, the system should be enhanced with AI mentoring capabilities that can provide basic assistance to the mentee when the network connection is failing, or is not available at all. One of the major challenges is to recognize automatically the current state of the surgery, a difficult case for computer vision algorithms as surfaces are fragmented, with view-dependent reflective properties, with complex occlusions, and deforming rapidly.

Beyond system refinements, additional user studies are needed to specialize the interface and to optimize the surgical telementoring benefits of our system in the context of many other types of surgical procedures.

Acknowledgments

We thank our Augmented Reality Tea group for insightful comments and suggestions. This work was supported by both the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-14-1-0042, and the National Science Foundation under Grant DGE-1333468.

References

[1] M. R. Ali, J. P. Loggins, W. D. Fuller, B. E. Miller, C. J. Hasser, P. Yellowlees, T. J. Vidovszky, J. J. Rasmussen, and J. Pierce. 3-d telestration: a teaching tool for robotic surgery. *Journal of laparoendoscopic & advanced surgical techniques*, 18(1):107–112, 2008.

[2] D. Andersen, V. Popescu, M. E. Cabrera, A. Shanghavi, G. Gomez, S. Marley, B. Mullis, and J. P. Wachs. Medical telementoring using an augmented reality transparent display. *Surgery*, 159(6):1646–1653, 2016.

[3] K. M. Augestad, J. G. Bellika, A. Budrionis, T. Chomutare, R.-O. Lindsetmo, H. Patel, C. Delaney, and M. M. M. M. Project. Surgical telementoring in knowledge translationclinical outcomes and educational benefits: a comprehensive review. *Surgical innovation*, 20(3):273–281, 2013.

[4] E. Azimi, J. Doswell, and P. Kazanzides. Augmented reality goggles with an integrated tracking system for navigation in neurosurgery. In

Virtual Reality Short Papers and Posters (VRW), 2012 IEEE, pp. 123–124. IEEE, 2012.
[5] E. M. Bogen, K. M. Augestad, H. R. Patel, and R.-O. Lindsetmo. Telementoring in education of laparoscopic surgeons: an emerging technology. *World journal of gastrointestinal endoscopy*, 6(5):148, 2014.
[6] N. Bourbaki. *Elements of Mathematics: Lie Groups and Lie Algeras*. Springer-Verlag, 1989.
[7] D. Campano, J. A. Robaina, N. Kusnezov, J. C. Dunn, and B. R. Waterman. Surgical management for chronic exertional compartment syndrome of the leg: a systematic review of the literature. *Arthroscopy*, 32(7):1478–1486, 2016.
[8] L. Chen, T. Day, W. Tang, and N. W. John. Recent developments and future challenges in medical mixed reality. *arXiv preprint arXiv:1708.01225*, 2017.
[9] X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng. Complete solution classification for the perspective-three-point problem. *IEEE transactions on pattern analysis and machine intelligence*, 25(8):930–943, 2003.
[10] P. Garcia. Telemedicine for the battlefield: present and future technologies. In *Surgical Robotics*, pp. 33–68. Springer, 2011.
[11] R. Gerhardt, J. Berry, R. Mabry, L. Flournoy, R. Arnold, C. Hults, J. Robinson, R. Thaxton, R. Cestero, J. Heiner, et al. Evaluation of contingency telemedical support to improve casualty care at a simulated military intermediate resuscitation facility: The em-angel study. *Journal of special operations medicine: a peer reviewed journal for SOF medical professionals*, 14(1):50–57, 2014.
[12] N. Hinata, H. Miyake, T. Kurahashi, M. Ando, J. Furukawa, T. Ishimura, K. Tanaka, and M. Fujisawa. Novel telementoring system for robot-assisted radical prostatectomy: impact on the learning curve. *Urology*, 83(5):1088–1092, 2014.
[13] J. J. LaViola Jr. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32(1):47–56, 2000.
[14] B. R. Lee, J. T. Bishoff, G. Janetschek, P. Bunyaratevej, W. Kamolpronwijit, J. A. Cadeddu, S. Ratchanon, S. O'kelley, and L. R. Kavoussi. A novel method of surgical instruction: international telementoring. *World journal of urology*, 16(6):367–370, 1998.
[15] Logitech ptz pro 2 video conference camera & remote. `https://www.logitech.com/en-us/product/conferencecam-ptz-pro2`. Accessed: 2018-03-12.
[16] C. F. Mackenzie, E. Garofalo, S. Shackelford, V. Shalin, K. Pugh, H. Chen, A. Puche, J. Pasley, B. Sarani, S. Henry, et al. Using an individual procedure score before and after the advanced surgical skills exposure for trauma course training to benchmark a hemorrhage-control performance metric. *Journal of surgical education*, 72(6):1278–1289, 2015.
[17] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pp. 50–60, 1947.
[18] Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H.-Y. Shum. Full-frame video stabilization with motion inpainting. *IEEE Transactions on pattern analysis and Machine Intelligence*, 28(7):1150–1163, 2006.
[19] G. C. Meng, A. Shahzad, N. Saad, A. S. Malik, and F. Meriaudeau. Prototype design for wearable veins localization system using near infrared imaging technique. In *Signal Processing & Its Applications (CSPA), 2015 IEEE 11th International Colloquium on*, pp. 112–115. IEEE, 2015.
[20] Microsoft hololens. `https://www.microsoft.com/en-us/hololens`. Accessed: 2018-03-12.
[21] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*, pp. 127–136. IEEE, 2011.
[22] A. C. of Surgeons Committee on Trauma, J. Fildes, J. Meredith, D. Hoyt, F. Luchette, M. Bowyer, P. Byers, E. Cornwell, J. Cuschieri, R. Gross, et al. *ASSET (Advanced Surgical Skills for Exposure in Trauma): Exposure Techniques When Time Matters*. American College of Surgeons, 2010.
[23] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965.
[24] Sharp model details. `http://siica.sharpusa.com/AQUOS-BOARD/Models/Details/PN-L603B`. Accessed: 2018-03-12.
[25] D. H. Shin, L. Dalag, R. A. Azhar, M. Santomauro, R. Satkunasivam, C. Metcalfe, M. Dunn, A. Berger, H. Djaladat, M. Nguyen, et al. A novel interface for the telementoring of robotic surgery. *BJU international*, 116(2):302–308, 2015.
[26] J. H. Shuhaiber. Augmented reality in surgery. *Archives of surgery*, 139(2):170–174, 2004.
[27] J. St Julien and N. D. Perrier. Video telementoring to accelerate learning of new surgical techniques. *JAMA surgery*, 151(7):671–672, 2016.
[28] S. Treter, N. Perrier, J. A. Sosa, and S. Roman. Telementoring: a multi-institutional experience with the introduction of a novel surgical approach for adrenalectomy. *Annals of surgical oncology*, 20(8):2754–2758, 2013.
[29] H. Wang, F. Wang, A. P. Y. Leong, L. Xu, X. Chen, and Q. Wang. Precision insertion of percutaneous sacroiliac screws using a novel augmented reality-based navigation system: a pilot study. *International orthopaedics*, 40(9):1941–1947, 2016.
[30] Webrtc. `https://webrtc.org/`. Accessed: 2018-03-12.
[31] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945.
[32] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.

Appendix A Additional tables

Table 4: Participants self-reported confidence before the experiment.

Confidence Assessment Aspect	EG	CG
Identify anatomical landmarks	3.00 ± 1.25	3.50 ± 1.00
Knowledge of procedural steps	3.00 ± 0.50	2.50 ± 2.00
Instrument handling technique	3.00 ± 2.00	4.00 ± 1.50
Perform procedure alone	2.00 ± 1.25	3.00 ± 1.25

Table 5: Participants self-reported confidence after the experiment.

Confidence Assessment Aspect	EG	CG
Identify anatomical landmarks	4.00 ± 1.25	4.00 ± 1.00
Knowledge of procedural steps	4.00 ± 0.00	3.50 ± 1.25
Instrument handling technique	4.00 ± 2.00	4.00 ± 2.00
Perform procedure alone	3.50 ± 1.00	3.50 ± 1.50