Computers & Graphics 118 (2024) 184–193

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on RAIXR

Mixed reality human teleoperation with device-agnostic remote ultrasound: Communication and user interaction

David Black *, Mika Nogami, Septimiu Salcudean

Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada

ARTICLE INFO

Keywords:
Human computer interaction
Mixed reality
Augmented reality
Teleoperation
Tele-ultrasound

ABSTRACT

For many applications, remote guidance and telerobotics provide great advantages. For example, tele-ultrasound can bring much-needed expert healthcare to isolated communities. However, existing tele-guidance methods have serious limitations including either low precision for video conference-based systems, or high complexity and cost for telerobotics. A new concept called human teleoperation leverages mixed reality, haptics, and high-speed communication to provide tele-guidance that gives an expert nearly-direct remote control without requiring a robot. This paper provides an overview of the human teleoperation concept and its application to tele-ultrasound. The concept and its impact are discussed. A new approach to remote streaming and control of point-of-care ultrasound systems independent of their manufacturer is described, as is a high-speed communication system for the HoloLens 2 that is compatible with ResearchMode API sensor stream access. Details of these systems are shown in supplementary video demonstrations. Novel interaction methods enabled by HoloLens 2-based pose tracking are also introduced and tests of the communication and user interaction are presented. The results show continued improvement of the system compared to previous work in instrumentation, HCI, and communication. The system thus has good potential for tele-ultrasound, as well as possible other applications of human teleoperation including remote maintenance, inspection, and training. The remote ultrasound streaming and control application is made available open source.

1. Introduction

Remote guidance technologies can improve tele-medicine, inspection, maintenance, and teaching [1,2]. For example, tele-ultrasound (tele-US), consisting of remote control or guidance of US procedures, is useful for remote and under-resourced communities where sonographers are often unavailable [3,4]. Additionally, tele-US can be used in care homes for elderly patients with mobility issues [5], for COVID-19 to minimize exposure [6,7], for trauma assessment in ambulances [8], or for training of sonographers [9,10]. Existing methods for tele-US consist of robotic teleoperation and audiovisual, video-conferencing-based guidance on a smartphone or tablet application. However, both of these approaches face myriad practical challenges and shortcomings. Human teleoperation addresses these issues.

1.1. State of the art

Reviews of robotic US systems are found in [11] and more recently [12]. While one robotic tele-US system has been used in clinical trials [13], commercial success has been limited despite the robots' ability to provide precision, low latency, and haptic feedback [14–17]. This is likely due to practical limitations including cost, restricted workspaces, time consuming set-up, and complex maintenance and operation. The cost is especially relevant when compared to otherwise inexpensive US devices, and makes it difficult to deploy such systems in small communities. Despite this, a large body of literature has studied autonomous robotic US [18,19] using force-based positioning [20,21], depth camera-based planning [22], and reinforcement learning [23]. However, guaranteed robustly safe human–robot interaction is an issue, particularly for regulatory bodies, and robotic tele-US remains relatively impractical.

On the other hand, there are several commercially available video conferencing-based mobile systems. Butterfly Network, Clarius Mobile Health Corp., and Philips use a point of care US (POCUS) device with live imaging and video conferencing available via a cloud interface on a mobile phone or tablet. Some visual guidance can be given by overlaying arrows or pointers on the US image. Though accessible and inexpensive, these systems are designed rather for quick expert review of a capable sonographer's captured images instead of teleoperation of an inexperienced novice. The resulting interaction is thus very inefficient for the latter case, leading to low precision and high latency.

Neither of the existing solutions is both flexible and accessible while being accurate and efficient. However, recent advances in extended

* Corresponding author.
E-mail addresses: dgblack@ece.ubc.ca (D. Black), mikarei@student.ubc.ca (M. Nogami), tims@ece.ubc.ca (S. Salcudean).

https://doi.org/10.1016/j.cag.2024.01.003
Received 28 August 2023; Received in revised form 6 January 2024; Accepted 8 January 2024
Available online 10 January 2024

reality (XR) research may solve this issue. Within the umbrella of XR, virtual reality (VR) immerses a user in a virtual environment, while augmented reality (AR) takes the real environment and adds visual information in the form of video overlays [24]. There are many definitions and classifications pertaining to the approach of augmenting the user's view, including Milgram and Kishino's "reality-virtuality continuum" [25], and more recent additions [26]. For clarity, we refer to our system as mixed reality (MR) according to the convention used by Microsoft, the manufacturer of the headset we use. Unlike in AR where overlays are applied to videos, in MR visual guides can be located within the real environment itself through the use of optically transparent headsets and waveguides [27].

The ability of MR to project 3D visual information seamlessly into the real world is the key enabling technology in a new concept we call "human teleoperation", introduced in [28], which leverages MR, haptics, and high-speed communication to bridge the gap in remote guidance techniques. In this system, a human follower is controlled as if they were a flexible, cognitive robot through an MR interface. In this way, both the input and the actuation are carried out by people, but with tight coupling, leading to latency and precision similar to a tele-robotic system. This enables remote guidance that is more intuitive, accurate, and efficient than existing audiovisual systems, yet less expensive, more accessible, and more flexible than robotic teleoperation [28]. The architecture and general function of human teleoperation is described in Section 2.

While MR has also been used extensively in other fields to guide tasks, these uses differ fundamentally from the human teleoperation concept. The idea of using mixed reality to overlay ultrasound images into the radiologist's field of view was introduced in 1992 [29]. The same group later performed a randomized trial and found that MR guidance improved accuracy in reaching a target during needle biopsies, decreasing mean deviation from 2.48 mm to 1.62 mm [30]. Similar work has been used for ultrasound-guided needle biopsies on more modern MR headsets such as the Microsoft HoloLens [31]. In these systems, the US images are projected into the imaging plane of the US device, and the needle is extrapolated linearly to show its predicted trajectory. This effectively allows the radiologist to "see inside" the patient and better aim their needle. In [32], the overlay was displayed on a monitor, which is less immersive and intuitive than displaying it in an MR headset or 3D display as found in the surgeon console of the da Vinci surgical system (Intuitive Surgical, Sunnyvale, CA).

The concept of overlaying medical images and 3D volumetric models of anatomy in position on a patient has been applied to other branches of medicine including laparoscopic surgery [33,34], robot-assisted surgery [35,36], and even treatment of depression through transcranial magnetic stimulation [37]. A very attractive implementation for open resection of liver metastases is presented in [38]. In [39], the images and models are not overlaid directly, but available for the surgeon to drag into whichever position is convenient for them to look at. This is due in part to the difficulty of registering pre-operative images to intra-operative anatomy, which tends to move and deform. Some papers attempt to deal with this by registering and deforming the pre-operative images according to US images captured intra-operatively [40]. A different approach is to use fluorescent markers [41].

In addition to image overlays which mostly extend a physician's view into otherwise obscured anatomy, some systems for manufacturing include static labels or pointers [42]. The authors of [43] developed a framework to gather information from a scene and create an MR application offline which contains visual instructions that can be overlaid onto the scene.

What every one of these applications has in common is a relatively static overlay of images or pointers intended to extend a user's vision or indicate a target to reach. To our knowledge, no other system exists in which a virtual guiding tool is controlled in real time by a remote person to guide the user in a hand-over-hand manner with tight coupling. Even recent advances in "holoportation" focus on creating a natural social interaction rather than providing dynamic guidance [44]. By approaching the MR guidance from a controls and tele-robotics perspective, human teleoperation enables very flexible remote control of procedures such as US exams. In this way, unlike most other MR guidance research, human teleoperation does not necessarily enhance the user's sensory capabilities by overlaying images and guides, but rather transports an expert's knowledge and skill into a remote location where it is needed.

1.2. Objectives and contributions

In developing this system, we follow several design objectives and requirements, collected through review of the state-of-the-art and discussions with sonographers and radiologists, as well as community physicians. To outperform existing methods, the system must be intuitive to use, efficient for procedures, and flexible for use in different procedures and locations. To achieve this, the communication system must be high throughput, low latency, and easily used in different networks and signal conditions. The visual control system should be user friendly for the follower and lead to good accuracy and little lag. Similarly, the expert should have a sensation as close as possible to carrying out the procedure in person, called teleoperation transparency, which involves visual, positional, and force feedback.

Finally, the system should be usable with any POCUS transducer. While most POCUS devices give remote access to images through a proprietary application, only some, such as Clarius, give low-level access to image streams and controls through open source APIs. For those without accessible APIs, it is not possible to integrate the device into the human teleoperation system, while for those with APIs, a different software interface or hardware abstraction layer has to be written for every make of device, which is impractical. Instead, we describe a system that streams screen-grabs of the Android smartphone connected to the device in real time to the remote site, and which enables remote control of the phone from the remote site. This gives the sonographer real-time access to the ultrasound images and controls.

This paper, which significantly extends a conference workshop paper [45], presents the human teleoperation system, building on previous work on the development and validation of several modules [28,46,47], and making multiple contributions to the state of the art. The overall system (Section 2), haptics (Section 5), human MR tracking performance (Section 6), and initial patient tests (Section 7) are first described. The following contributions are then presented:

- A new communication system to enable the use of the HoloLens2 Research Mode APIs [48] (Section 3),
- Design and integration of device-agnostic remote streaming and control of POCUS ultrasound probes (Section 3),
- A new initial start-up, registration, and calibration procedure that is easier, more accurate, and >85% faster than previously reported [28] (Section 4),
- New active visual aides to eliminate steady-state error in follower positioning using new probe pose sensing [49] (Section 8),
- Tests and comparison with a different MR headset (Section 8),
- New insights into potential future work (Section 9)

Most of the new developments are demonstrated as videos in Supplementary Material, and the code for device-agnostic remote ultrasound streaming and control is open-sourced.

2. Human teleoperation for ultrasound

In this section we describe the specific prototype system we built for tele-US. Other applications for human teleoperation are discussed in Section 9. A diagram of the entire system is found in Fig. 1.

In general teleoperation systems, there is a local robotic agent which interacts with its environment, and a remote operator who receives

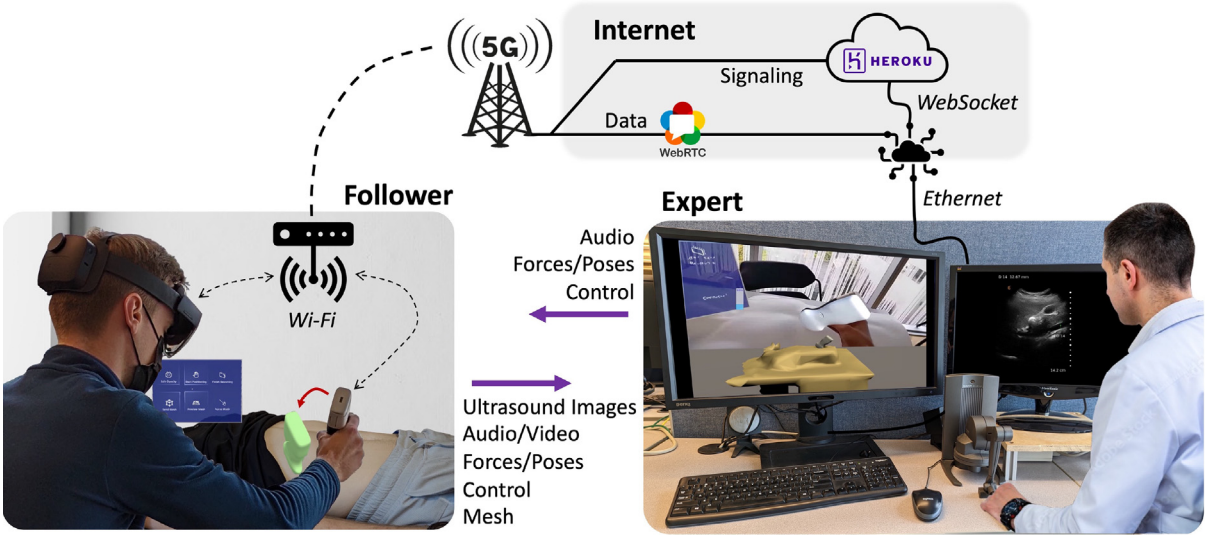

Fig. 1. Overview of human teleoperation system.

Table 1
Worst-case uplink and downlink data throughput requirements on the follower side.

Type	Uplink	Contents	Downlink	Contents
Timing	1.28 kbps	64 bit long int × 20 Hz	3.84 kbps	3 64 bit long ints × 20 Hz
Force	16 kbps	3 32 bit floats + 1 64 bit timestamp × 100 Hz	16 kbps	Same as uplink
Pose	28.8 kbps	7 32 bit floats + 1 64 bit timestamp × 100 Hz	28.8 kbps	Same as uplink
Video	≈1–2 Mbps	960 × 540 px H.264 encoding, 25 Hz, Variable quality	0	No video
Audio	128 kbps	Typical MP3 bitrate (Part of MPEG-4 stream)	128 kbps	Same as uplink
US	4.64 Mbps	58 kB JPEG image (worst-case) × 10 Hz	0	No US
Mesh	2.3 Mbps	≈12k mesh triangles × 3 points and 3 indices × 32 bit floats	0	No downlink mesh
Total	6.81 Mbps	Mesh sent rarely on demand. Peak throughput 9.11 Mbps	180 kbps	Sum

feedback from the local agent and provides instructions on what actions to carry out. Traditionally, these are called the "master" and "slave" respectively, though we avoid this terminology. In human teleoperation, the local agent is a human, the "follower", while the remote operator is experienced in the task being performed, and is referred to as the "expert". For tele-US, the expert is a sonographer or radiologist with ultrasound experience, and the follower is an inexperienced novice, whose interactions with his/her environment constitute moving an US probe on a patient, as instructed by the expert.

The follower wears an MR headset, the HoloLens 2 (Microsoft, Redmond, WA), which projects a virtual US probe into their field of view. To perform the desired procedure, they align their real US probe with the virtual one and track it as it moves around on the patient. In this way, the desired position and orientation (pose) of the probe are achieved. The desired force is similarly reached through a visual control system, as explained in Section 5.

The desired pose and force of the probe are set in real time by the expert, who manipulates a haptic device (Touch X, 3D Systems, Rock Hill, SC). This is a small, 6-degree-of-freedom (DOF) serial manipulator that measures the 6-DOF pose of its pen-like handle and applies 3-DOF forces to the handle's tip for haptic feedback [50]. We replaced the Touch X conventional handle with a 3D-printed US probe-shaped end effector. As the expert grasps the end-effector and carries out his/her desired motion, the probe pose is transmitted to the virtual handle guiding the follower. The haptic device applies forces back to the expert. By pushing against the haptic device, the expert also inputs his/her desired force. This is described in detail in Section 5.

One of the primary objectives of the system is teleoperation transparency, or making the expert feel as if they are performing the procedure in person by matching the expert and follower positions and forces. To this end, a three or four-channel architecture is required, in which force and velocity are sent from expert to follower and vice versa at a high rate [51,52]. Hence, the desired forces and poses are sent from expert to follower, along with an audio stream and some control packets. Conversely, the measured force and pose of the follower are returned to the expert, along with an audio stream, a video stream captured by the HoloLens which includes the virtual objects in place (called an MR capture), the live ultrasound images, control packets, and occasional mesh data. The communication system for transmitting this data is described in Section 3, and the meshes are discussed in Section 4.

In this way, the expert sees the patient, ultrasound probe, and ultrasound images live, and feels the applied forces. She/he consequently decides how to move and performs the motion on the haptic device, which updates the input signal to the follower through the visual control system, which the follower tracks. At the same time, the expert and follower are in verbal communication. Preliminary experiments with carrying out a remote ultrasound examination by a novice follower when teleoperated by an expert have shown promising results (Section 7) [28].

3. Communication

As described in Section 2, a large amount of data must be transmitted at high speed. In particular, the throughput requirements on the follower side are outlined in Table 1. As the expert is stationary and attached via a wired connection, their throughput is of less concern.

To achieve the large throughputs required by such tight coupling between expert and follower, a high-performance communication system was developed and is described in [47,53]. 5G provides ultra-reliable low latency communication and enhanced mobile broadband that are well suited to this application due to the large throughput, high reliability, and low latency requirements. The University of British Columbia was the first campus in North America equipped with a 5G network, so

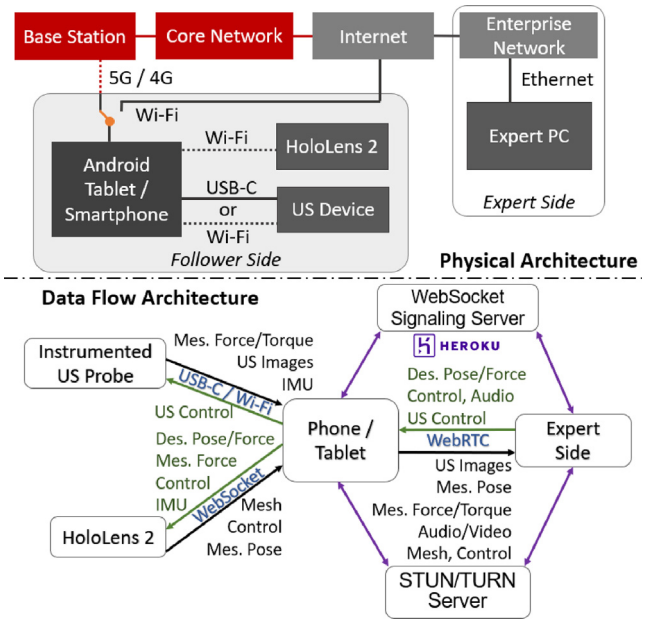

Fig. 2. Top: Physical, data link, and network layers of the communication architecture. Bottom: Session and application layers of the communication architecture, showing data flow (expert to follower in green and follower to expert in black). The control channels include various user settings, button-presses, gestures, and other functions. STUN: session traversal using NAT (network address translation). This is a server that sends back a device's public address and connectivity information, which is then shared with the peer over a signaling server to establish the peer-to-peer WebRTC connection.

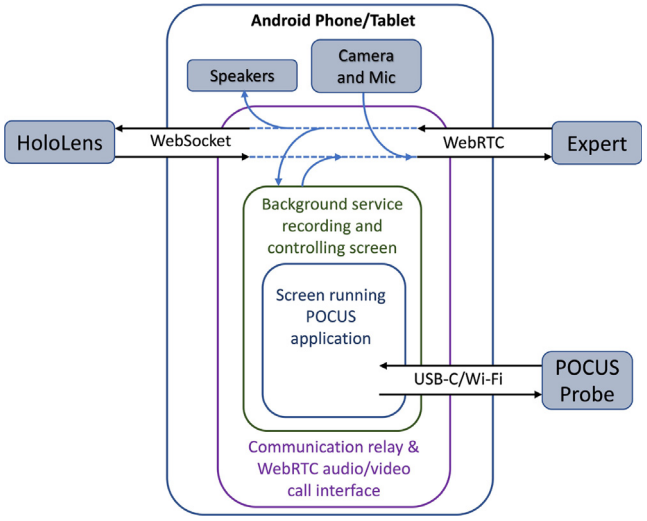

Fig. 3. Android application for ultrasound image streaming and remote control of any POCUS device, as well as relaying of WebRTC data from the expert to WebSocket data to the HoloLens and vice versa.

we partnered with Rogers Communications Inc., a Canadian telecommunications company, to utilize their non-standalone sub-6 GHz 5G network, shown in red in Fig. 2.

Several open systems interconnection (OSI) model layers of the communication architecture are shown in Fig. 2. Since the follower side is mobile and could be installed anywhere, including a moving ambulance or a remote location, it is designed to work over mobile networks. A Sercomm 5G modem with a wired connection to a Wi-Fi router provides the interface between the mobile network and the user devices. This is equivalent to 5G CPE devices sold by Yeastar (Xiamen, China), Oppo (Dongguan, China), and others, or to "MiFi" devices from Inseego (San Diego, CA). Alternatively, as explained below, all data can be routed through an Android smartphone or tablet which is connected to the mobile network. The HoloLens 2, US device, and force and pose sensors connect to a local Wi-Fi network, which in turn connects to the Internet via mobile networks or directly. Conversely, the static expert side connects to the Internet via a wired connection.

Built on this infrastructure, the communication system was designed to minimize latency using the WebRTC (Web Real Time Communication) standard. As shown in Fig. 2, WebRTC is a peer-to-peer architecture which eliminates server-related delays. It is built on stream control transmission protocol [54] and real time transport protocol [55], both of which are related to the user datagram protocol (UDP) which prioritizes speed over reliability. Instead of retransmission upon packet loss, which adds latency, packets are sent at a sufficient rate that lost ones are quickly replaced and local consistency checks are in place. To establish connectivity initially, WebRTC uses Session Description Protocol (SDP) and Interactive Connectivity Establishment (ICE) to find an optimal connection between two peers over any network and through most router NAT (Network Address Translation) schemes or firewalls. Both peers determine their own information from a STUN server, and then exchange this SDP over a signaling server which we implemented using Python WebSockets and host on a cloud platform called Heroku. Using this information, the peer-to-peer connection is established. The signaling is encrypted locally, and the signaling server requires authentication. Similarly, WebRTC is built upon Datagram Transport Layer Security (DTLS), so all communication is encrypted and secure, which is important for medical applications.

Unfortunately, WebRTC for a Universal Windows Platform (UWP) device such as the HoloLens 2 only runs on a 32-bit ARM CPU architecture. Conversely, the HoloLens 2 Research Mode APIs [48], used for inside-out pose tracking of the ultrasound probe [49], only work in an ARM64 architecture. Hence, to enable pose tracking and other computer vision, WebRTC cannot be used from the HoloLens. Furthermore, every type of ultrasound probe has different APIs for accessing images and controlling the device; in fact most permit no such access. As a result, the communication system from [47,53] had to be updated.

3.1. Ultrasound control

The follower side now includes an Android smartphone or tablet which acts as communication hub, audio/video source, and ultrasound probe interface. This is shown in Fig. 3. The smartphone or tablet connected to the POCUS device runs a background service which connects to the expert over WebRTC, and to the HoloLens over a WebSocket. Using either WebRTC's APIs directly, or a library called Unity Render Streaming, the screen of the device is captured and streamed to the expert side. This shows the POCUS application, which includes the ultrasound images and controls in real time, in high quality, and with low lag. When the expert sees the images, they may decide to change certain imaging parameters. When they click and/or drag with their mouse on the stream of the tablet screen, the start and end locations and times of the gesture are captured and sent over WebRTC back to the device, where they are built into a gesture and executed using the Android Accessibility Services. In this way, the expert can fully control the device's screen and thus the imaging parameters and controls. To emulate pinching gestures, the expert's mouse wheel is captured and used, for example, to zoom in or out of the ultrasound image. Since the follower can view the image on the Android screen as well, the expert can point out image features to the follower by pressing a key which shows their cursor as an icon on the Android screen.

In addition, all packets from the expert are forwarded to the HoloLens and vice versa, and the camera and microphone of the Android device are streamed to the expert, while the expert's audio is played on the device's speakers, allowing for verbal and visual contact between expert and follower. Having the video stream stem from the Android rather than the HoloLens 2 as described previously [28,45] has the advantage of being able to move the camera to a desired

Table 2
Round trip time (RTT; mean ± std; ms) versus data throughput in good signal conditions for various networks. The last row shows slightly increased RTT (by 1.64 ms on average) due to the packet forwarding over a local WebSocket explained in Section 3.

Throughput	Ethernet	WiFi	4G LTE	5G NR
1.28 kbps	1.07 ± 0.57	5.80 ± 3.30	38.41 ± 6.63	26.95 ± 7.72
2.17 Mbps	0.93 ± 0.59	5.87 ± 1.75	43.49 ± 30.18	39.61 ± 6.14
4.97 Mbps	–	–	52.30 ± 59.30	47.64 ± 22.81
6.81 Mbps	1.07 ± 0.88	7.82 ± 9.90	66.58 ± 123.00	70.44 ± 78.29
2.17 Mbps	–	7.41 ± 1.43	45.63 ± 2.21	40.85 ± 1.64

location to show something specific, or to mount it so it is more stable than the HoloLens. Additionally, the WebSocket connection and packet forwarding from the HoloLens is far less optimized for video streaming than WebRTC, so sending the video from the Android rather than the HoloLens greatly reduces delay.

In [47] we carried out tests of the communication performance over various networks, in different signal conditions, and with different throughput values to quantify the response to different conditions. We also experimented with configurations including retransmission, packet ordering guarantees, channel splitting, and queuing. Some of the key results are shown in Table 2. We found that the communication performed sufficiently well over 4G or 5G with good signal conditions up to about 5 Mbps continuous throughput, and well beyond this for WiFi or Ethernet. Typical round trip times (approximately two times latency) for 5G were 25 to 50 ms which is appropriate for teleoperation with direct force feedback in the case of a relatively soft contact environment like a patient [56–58].

The added latency of the WebSocket and packet forwarding has now been tested using the same method [47], and is shown in the last row of Table 2. On average, the re-routing adds only 1.64 ms to the round-trip time. In the case of 5G, this is approximately a 4.14% increase and is not noticeable for the users. Importantly, the lag-sensitive video and audio feedback stem directly from the Android device and are not subject to the same delays. On the other hand, pose and force data for the teleoperation has the additional delay. The new communication system is shown working in Supplementary Material.

4. Spatial registration

The expert is presented with a live video stream from the follower's perspective, and the ultrasound images of the patient. Based on these, he/she decides how to move. If the expert sees the follower's US probe positioned too far left in the follower's view, it is natural for them to move their probe to the right. However, the follower coordinate frame is positioned by the HoloLens and varies every time the application is started. Therefore, a registration between the expert and follower spaces is required to ensure that directions in both frames correspond intuitively.

Furthermore, the HoloLens 2 continuously captures a 3D spatial mesh of its environment for SLAM purposes. This mesh can be accessed and sent to the expert side to provide 3D visual feedback as well as a haptic interaction, described in Section 5. This process was also discussed in detail in [28].

To extract the mesh of only the patient, not the rest of the room, and to enable the spatial registration, the position and orientation of the patient must be determined. To this end we previously developed two methods [28]. ArUco markers were placed [59] in four corners of the test surface simulating the patient bed. The HoloLens automatically determined the poses of the markers, which were previewed to the user and individually confirmed by them. It thus found the patient bounding box, orientation, and the transform of the electromagnetic tracker. These steps are shown in Fig. 4. This was, however, time consuming and too complex for novice users of the HoloLens 2.

With improved pose tracking [49], however, a simplified method can be used and no registration to an external pose tracker is required.

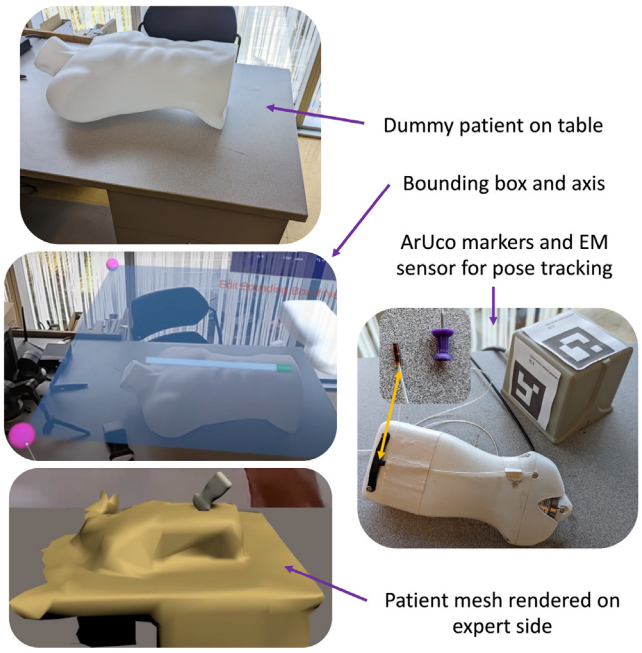

Fig. 4. Aspects of expert-follower spatial registration. The US probe is a 3D printed dummy used for testing of the instrumentation. The electromagnetic (EM) transmitter is the box with two ArUco markers. This is replaced with vision-based pose tracking from the HoloLens 2.

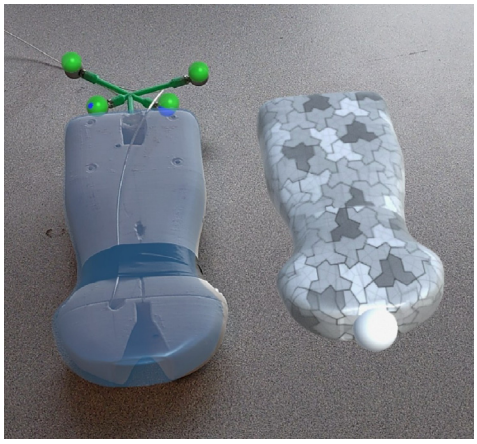

Fig. 5. MR Capture from HoloLens 2 of (Left) the dummy ultrasound probe with IR markers showing successful pose tracking. The green spheres are the measured marker positions while the blue ones are the output of the unscented Kalman filter, and the blue virtual probe shows the resultant calculated pose. This preview can be disabled during actual operation. And (Right) the virtual ultrasound probe controlled by the expert, with patterns for better contrast and depth perception. These can also be seen in videos in Supplementary Material.

At the start of the exam, the follower can simply place the ultrasound probe along the craniocaudal axis, with its tip in the approximate area of the scan, and press a virtual button. The axis of the probe, projected onto the horizontal, is taken to be the left–right axis for the expert, and the virtual probe's initial position is set to the position of the real probe. The probe tracking is shown in Fig. 5. Thus, the procedure is ready to start without any need for external fiducials, alignment of markers, or manual setting of the initial probe position by dragging the hologram into place. To determine the bounding box, a box of fixed size and offset around the probe can be generated and then adjusted if necessary by the follower.

In a simple test in which the startup process was repeated 10 times each using the old and new methods, the new approach took on average

Fig. 6. Color and error-bar force rendering schemes, showing the desired follower action for each rendering. Note, the color and bar length vary continuously and linearly between the shown states.

4.2 ± 3.1 s compared to 29.3 ± 3.8 s for the ArUco markers, thus decreasing startup time by 85.7%. These tests were carried out by an experienced user accustomed to manipulating the virtual objects, so inexperienced users would likely derive an even greater advantage. Additionally, the initial probe position was found to be more practical and accurate than when it was manually dragged into position. The new and old methods are demonstrated in Supplementary Material.

5. Force control

We have so far discussed the visual pose control method on the follower side. However, force is almost equally important in ultrasound, determining what anatomical structures are visible and ensuring they are not too deformed. Moreover, sonographers usually look almost exclusively at the US images and rely on their sense of touch to move the probe on the patient. Hence, both controlling the applied force of the follower and feeding back that force to the expert are essential in human tele-operated ultrasound.

Necessary for both of these topics is force sensing at the ultrasound probe. This can be achieved by fabricating a shell around the US probe, connected to the probe using an off-the-shelf force sensor [14,60–63]. However, this is bulky and heavy. We are developing an alternative method [64,65], but for preliminary testing of our algorithms for force feedback we have 3D printed a dummy US probe with an ATI Nano25 force/torque sensor embedded in the tip.

To generate an error signal, the measured force is subtracted from the expert's desired force applied to the haptic device. This error signal is then displayed visually for the follower so they can move to minimize it and thus match the desired force. The visualization could involve an arrow that grows, shrinks and changes direction, or a second virtual probe, offset by an amount proportional to the force error and the patient tissue stiffness, in the direction of the force error. As the follower must track the virtual probe's pose, however, it is beneficial to have them focus exclusively on the probe. Additionally, the desired force is usually normal to the patient tissue, as the slippery ultrasound gel does not allow for large transverse forces. To avoid excessive cognitive load, therefore, we tested two different force rendering schemes that change only the virtual probe itself and do not contain direction information [46].

The two rendering schemes are color and error-bar and are shown in Fig. 6. In the former, the ultrasound probe color varies smoothly between blue (follower should apply more force), green (force error is small), and red (follower should apply less force). Similarly in the latter, the error bar grows toward the patient or away and changes color depending on the force error. In tests described in Section 6, we found that the error-bar rendering was superior, likely because it was easier to resolve errors close to zero using the error-bar than by differentiating slightly blueish green from slightly reddish green.

The force sensor can also be used for force feedback to the expert in one of several teleoperation configurations. These include a 2-channel force-position [66,67] or dual-hybrid force/position [68] controller, a 3-channel system with local feedback [69], or a 4-channel bilateral parallel force/position [52] or matched impedance [70] controller. The problem with many of these approaches is that performance and stability are limited by time delays. In human teleoperation, delays stem not only from the communication system, which is relatively fast, but also from the reaction time of the follower person, which can be substantial. Wave variable-based schemes have been proposed to maintain stability in time-delayed force-reflecting teleoperation [71], but position tracking is sometimes sluggish and wave reflections can be disorienting [72].

Many papers explore robust time-delayed teleoperation, but our system presents an opportunity for a potentially simpler approach. As explained in [28], a 3D mesh of the patient, captured by the HoloLens 2 and transmitted to the expert, can be rendered as a virtual fixture on the haptic device. When the expert displaces the haptic device into the region delineated by the mesh, it is met with an opposing force proportional to configurable stiffness and damping coefficients. Using the force and pose sensing of the US probe, it is possible to estimate the impedance parameters of the patient tissue [73]. The estimated parameters can be fed back along with the patient mesh at a much lower rate than forces would have to be, and as they are relatively constant, a temporary packet delay or loss in communication would have no great effect. This approach, with constant approximate parameters, was used in early tests of the system described in Section 7 [28].

6. Human performance

To evaluate the feasibility of human teleoperation, we carried out experiments of human pose and force tracking ability, and quantified the responses of the human and visual control system [46].

Methods: The tests involved 11 participants who ranged in age and background (20–64 years, mean 32, 36% female). As with other control systems, we performed step and frequency response tests and arbitrary series of motions to determine the reaction time, overshoot, settling time, steady-state error, frequency dependence of relative phase and magnitude, RMS tracking error, and tracking lag for position, orientation, and force. Each test was performed with both force rendering methods and two different pose renderings.

To carry out these tests, the desired motion sequence was rendered using the HoloLens 2, and the subject's response was measured using a force sensor (ATI Nano25) and electromagnetic pose sensor (NDI driveBAY) embedded on a 3D printed ultrasound probe dummy held by the subject. For the pose and force step responses, the virtual probe jumped back and forth between positions and forces respectively, holding each position or force for 5 s. For position, the direction was changed every time so the follower could not predict the motion and bias the results. For the frequency response, the subject tracked a frequency-swept sinusoid until they were unable to follow. In each of these tests, the position and force tracking accuracy, speed, and lag were analyzed.

Results: The step and frequency response tests are shown in Fig. 7, and example force, position, and orientation tracking tests are found in Fig. 8. The key results of these tests are outlined in Table 3. While extensive discussion is found in [46], these results show that tracking both pose and force separately or simultaneously is feasible, with relatively small tracking error and lag of approximately 0.35 s. The worst case reaction time is to a step change in pose or force and varies based on the step amplitude, but remains well under 1 s. Similarly, the frequency response depends on the input signal's amplitude but for good performance the human force and pose tracking are limited to about 0.25 Hz and 1 Hz respectively. This is slower than the human

Table 3
Human performance test key results from [46]. This shows that tracking pose or force is easier than both at once, but that when tracking both simultaneously the speed and accuracy are sufficient for relatively high-performing teleoperation.

	Force	Pose
Single parameter (pose or force):		
Continuous tracking lag	255 ± 118.88 ms	346.23 ± 118.15 ms
Continuous tracking rms error	0.99 ± 0.29 N	6.24 ± 1.93 mm, $5.93 \pm 1.85°$
Dual parameter (pose and force simultaneously):		
Continuous tracking lag	345.5 ± 87.60 ms	345.5 ± 87.60 ms
Continuous tracking rms error	1.25 ± 0.33 N	8.5 ± 1.4 mm, $7.27 \pm 2.25°$
Step responses:	*(10 N step)*	*(10 cm step)*
Reaction time to step changes	171.5 ± 85.9 ms	628.3 ± 102.3 ms
Steady state error	0.26 ± 0.16 N	2.8 ± 2.1 mm
Frequency responses:	*(10 N magnitude)*	*(10 cm magnitude)*
Max. frequency for good tracking	0.25 Hz	1 Hz
For smaller input magnitudes, max. frequency	Decreases	Increases

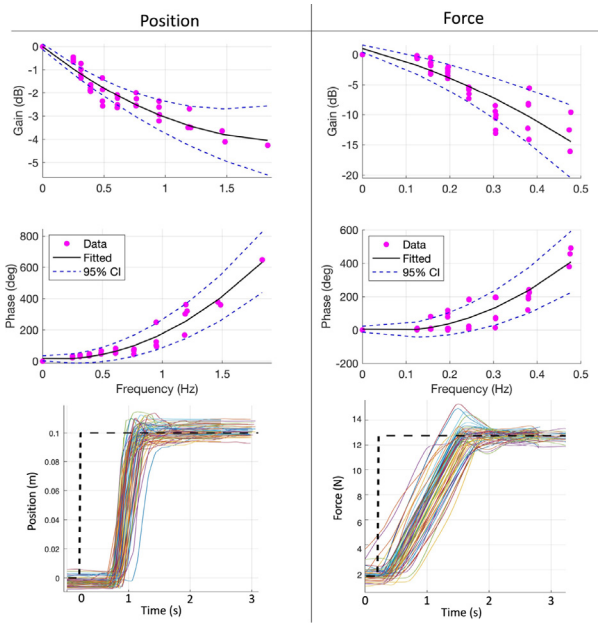

Fig. 7. MR human tracking performance step and frequency response test results [46]. The dotted step response is the desired signal. The other lines are the subjects' measured responses. The subject frequency responses are plotted, as is a fitted second degree polynomial (black solid line), and the 95% confidence interval of the fit (dashed lines). Numerical results from these tests are found in Table 3.

hand's force bandwidth presented in [74], though in the same order of magnitude, showing that the actuations are limited by cognitive rather than motor factors.

Discussion: All of these results point to excellent performance potential for human teleoperation. The motions in ultrasound, for example, are much less demanding than in the tests. Additionally, they indicate that careful consideration of time delays is essential for force feedback, as discussed in Section 5, because the time delays imposed by the human response time vary substantially and can be up to a large fraction of a second.

7. Patient tests

Though the human performance tests in Section 6 and the communication experiments in Section 3 show good potential for human teleoperation, neither evaluates the performance of the system in a practical clinical setting. The human teleoperation system should improve precision, efficiency, and completion time of US procedures compared to existing audiovisual guidance methods, and improve set-up time, ease-of-use, and accessibility compared to robotic systems.

Table 4
Results from preliminary volunteer study with expert sonographer. Each procedure and method has a completion time and measurement percent error. Kidney values are displayed as transverse × craniocaudal dimension error. This shows that human teleoperation is faster and more precise than existing audiovisual guidance ($p = 0.052$).

	Kidney	Vena cava
Direct	1:28 ± 0:21 –	0:42 ± 0:04 –
Audiovisual	4:13 ± 3:58 3.5% × 14%	3:55 ± 0:25 12.8%
Human Teleop.	1:36 ± 0:23 3.5% × 4.6%	0:49 ± 0:02 8.3%

While larger-scale trials with patients as well as comparative studies between human and robot teleoperation are planned future work, we have completed a preliminary, small-scale study [28].

Methods: This involved four healthy volunteers and a physician with extensive US experience on two example procedures: kidney measurement and inferior vena cava (IVC) examination. The volunteer followers had no ultrasound experience. Both procedures involved quantitative endpoints (measurement of IVC diameter and kidney transverse and craniodaudal dimensions) and were carried out three times: once directly by the physician, once using the Clarius video conferencing interface, and once using human teleoperation. Each volunteer follower carried out one procedure using human teleoperation and the other procedure using audiovisual guidance, to avoid bias from learning effects if they performed the same procedure twice. The direct US measurements were taken as a reference, and the measurement error and completion time of the different methods were compared.

Results: The average completion times and measurement errors compared to the direct US are outlined in Table 4 for each procedure and method. Despite the small sample size, it was found that human teleoperation was faster and more accurate than audiovisual guidance ($p = 0.052$) and not much slower than direct ultrasound.

8. Follower MR experience

It is important that the human teleoperation system is as intuitive and easy to use as possible since its purpose is to guide novices. To evaluate this, user feedback and a questionnaire were used.

In the patient tests as well as numerous ($n = 86$) demos given to technical, non-technical, medical, and non-medical people of all ages, feedback has uniformly been that tracking the virtual tool is very easy and intuitive. All participants were able to track motions within seconds of putting on the HoloLens 2, sometimes with no specific instruction. The sole difficulty was an occasional initial incorrect depth perception, leading to the follower person holding the real probe closer to the headset than the virtual probe was. After a single prompting telling the

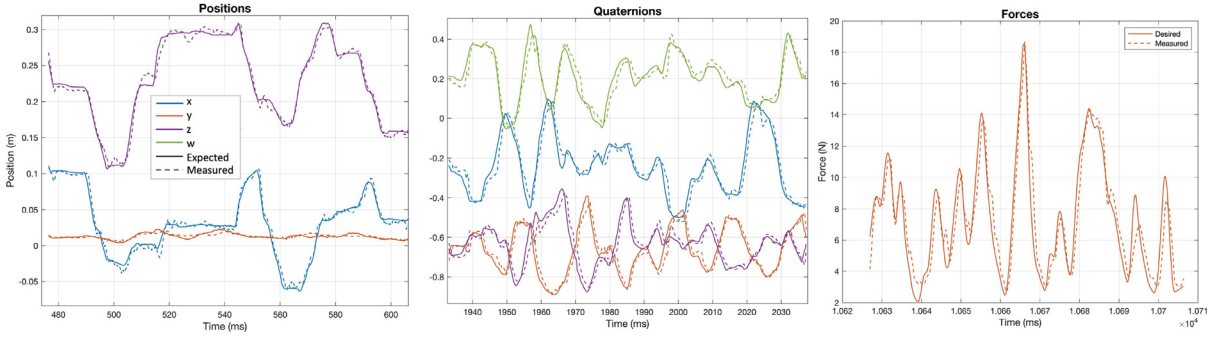

Fig. 8. Results from one example human performance test [46]. We see good tracking in all axes of position and orientation. The force with error-bar rendering also exhibits low tracking error. Solid lines are desired values while dotted lines are the measured response.

Table 5
User questionnaire scores out of 5 for the MR follower interface [46] (5 = a lot; 1 = a little).

Participants who became dizzy	0
Preference for error-bar over color force	4.55 ± 0.52
Mental demand of force tracking	3.36 ± 0.5
Mental demand of pose tracking	2.18 ± 0.6
Dual param tracking more difficult than single?	4.36 ± 0.81
Intuitiveness of tracking interface	4.82 ± 0.4
Physical demand	2.36 ± 0.8

user the approximate location of the virtual probe, this was rectified in all cases and there were no recurrences.

To address the issue of depth perception, the virtual ultrasound probe was given a distinct and unique pattern that contrasts with its environment, as shown in Fig. 5. Stereo pattern matching is important for binocular depth perception [75], and previously the virtual probe was monochromatic and lacking a distinct outline in well-lit environments. Additionally, contrast between the texture of an object and the background, particularly when the luminance of the object is similar to that of the background, significantly affects depth perception [76]. Thus, the pattern can help with more accurate pose tracking, though this has not yet been tested.

In the human performance tests [46] described in Section 6, users filled out a questionnaire regarding their experience. The questions and results are shown in Table 5. It was found that again tracking both force and pose is very intuitive, though requires some focus to perform well. In dimmer lighting conditions, the holographic US probe can occlude the real one, thus blocking visual positioning feedback for the follower. In this case, using a virtual probe with transparent sections or decreasing its overall opacity was sometimes useful, depending on user preference.

Several users commented that it was sometimes difficult to know how well they were matching the virtual probe, likely due to similar factors. However, the new pose measurements can be used to improve the steady-state error of the follower by giving a visual indication of how well the follower is matching the desired pose. For example, if the user position or orientation is off by more than a threshold, a second virtual probe can be animated temporarily, moving smoothly from the current pose to the desired one a few times to indicate to the follower how they should move. This is demonstrated in Supplementary Material. Alternatively, colors can be used to indicate the quality of matching, though this should not be confused with force information. Adding further visual cues may also lead to distraction or sensory overload, as users already reported relatively high cognitive demand.

9. Future work

In addition to the further patient tests and comparison to robotics mentioned in Section 7, this project presents many avenues for further research. For example, the human teleoperation concept can be applied to other fields beyond ultrasound, including any application where tightly coupled hand-over-hand guidance is needed. This could include maintenance, inspection, and training. Furthermore, because the actuations are ultimately carried out by a human, not a computer or robot, this presents an interesting application for artificial intelligence. The human expert could be replaced by an AI agent, trained through learning from demonstration or reinforcement learning, for example, to guide procedures or sub-tasks autonomously on demand. This is not limited by the concerns of robustly safe human–robot interaction that affect autonomous robotics. Since the transducer pose, force, ultrasound image, and video are all captured and can be recorded, rich datasets can be built for training such an AI system.

In terms of communications, it is possible to leverage further aspects of 5G, including exporting costly computations used for computer vision or AI to proximal edge servers with minimal latency. In addition, we plan to test the communication over mm-wave band 5G, which promises the greatest increase in throughput.

Furthermore, the spatial registration process could be made faster and easier using computer vision to segment the patient automatically. For improved haptics, local convex meshes could be approximated and streamed in real time to remove artefacts such as splinters and to take into account a moving patient. Different visual control architectures have the potential to reduce cognitive load and thus remove the decreased performance in Table 3 when tracking pose and force simultaneously compared to one at a time. Moreover, delay-robust force feedback should be implemented and tested, potentially in a shared architecture with mesh and impedance feedback.

Finally, future work will evaluate further alternative MR headsets such as the Magic Leap 2, Apple Vision Pro, HoloBoard, or Creal LightField display. For example, an alternative MR system, the Nreal Light, has been tested for comparison. The holograms were found to be very sharp and vibrant, and relatively positionally stable, though not as stable as the HoloLens 2. While the SDK is easy to use and well integrated with Unity, the hand tracking is currently poor, and a physical controller is needed to interact with the virtual environment. It was found that the most efficient mode of interaction with virtual buttons was by using the gaze pointer (which tracks the head, not the pupils) to point at buttons, and pressing the button on the controller to select them. External control of the virtual probe was effective, including changing the opacity of the hologram. Thus, the core functionality of human teleoperation system works on the Nreal Light. More advanced features such as probe pose tracking, mesh-based haptics, and automated startup, which rely on IR cameras and effective mesh generation, are not possible without additional sensing equipment. However, the small form factor and low cost of the device along with the effectiveness of the fundamental aspects of human teleoperation make this potentially usable for a lower-cost human teleoperation system in lower-resource environments. Other devices such as the Magic Leap 2 or Apple Vision Pro may be more suitable but face cost trade-offs. This will be explored further in future work.

10. Conclusion

This paper has described multiple novel and broadly useful aspects of a new system for teleoperating a human with precision and latency similar to that of a robot through a mixed reality interface. The work of several papers is summarized and key results are shown which demonstrate the performance and feasibility of human teleoperation. Furthermore, new developments and tests in HCI and communications are described. Not only does this system bring up interesting basic science questions about mixed reality, teleoperation, and HCI, but it has the potential to be of use in myriad industrial applications, and to impact many remote communities that otherwise have difficult access to healthcare.

CRediT authorship contribution statement

David Black: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mika Nogami:** Visualization, Validation, Software, Methodology, Investigation. **Septimiu Salcudean:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

We gratefully acknowledge scholarship support from the Vanier Canada Graduate Scholarships program, Canada, infrastructure support from CFI and funding support from NSERC, Canada and the Charles Laszlo Chair in Biomedical Engineering, Canada, as well as infrastructure, technical, and funding support from Rogers Communications, Canada and MITACS, Canada.

Appendix A. Supplementary material

Note: these videos are demonstrations of the concepts and do not represent the achievable performance. This is due to significant lag introduced when the mixed reality capture is running on top of our application, and because the mixed reality capture creates an apparent offset between real and virtual objects that is not present for the user.

Video of application startup and initial registration/setup:
https://youtu.be/_ZOTvdS4taY

Video of pose tracking and visual correction:
https://youtu.be/KyAZMkdFzo

Video of patterned virtual probe for better depth perception:
https://youtu.be/i5_LtReP04k

Video of remote ultrasound streaming and control:
https://youtu.be/apzVsq1L8PI

Code for remote ultrasound streaming and control:
github.com/dgblack/AndroidStreamControl

References

[1] Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021;17(2):218–21.
[2] Yamamoto T, Otsuki M, Kuzuoka H, Suzuki Y. Tele-guidance system to support anticipation during communication. Multimodal Technol Interact 2018;2(3):55.
[3] Kaneko T, Kagiyama N, Nakamura Y, Hirasawa T, Murata A, Morimoto R, Miyazaki S, Minamino T. Effectiveness of real-time tele-ultrasound for echocardiography in resource-limited medical teams. J Echocardiogr 2022;20(1):16–23.
[4] Jemal K, Ayana D, Tadesse F, Adefris M, Awol M, Tesema M, Dagne B, Abeje S, Bantie A, Butler M, et al. Implementation and evaluation of a pilot antenatal ultrasound imaging programme using tele-ultrasound in Ethiopia. J Telemed Telecare 2022;1357633X221115746.
[5] Mariani PJ, Setla JA. Palliative ultrasound for home care hospice patients. Acad Emerg Med 2010;17(3):293–6.
[6] Smallwood N, Walden A, Parulekar P, Dachsel M. Should point-of-care ultrasound become part of healthcare worker testing for COVID? Clin Med 2020;20(5):486.
[7] Uschnig C, Recker F, Blaivas M, Dong Y, Dietrich CF. Tele-ultrasound in the era of COVID-19: A practical guide. Ultrasound Med Biol 2022.
[8] Montoya J, Stawicki S, Evans DC, Bahner D, Sparks S, Sharpe R, Cipolla J. From FAST to E-FAST: an overview of the evolution of ultrasound-based traumatic injury assessment. Eur J Trauma Emerg Surg 2016;42(2):119–26.
[9] Drake AE, Hy J, MacDougall GA, Holmes B, Icken L, Schrock JW, Jones RA. Innovations with tele-ultrasound in education sonography: the use of tele-ultrasound to train novice scanners. Ultrasound J 2021;13(1):1–8.
[10] Soni NJ, Boyd JS, Mints G, Proud KC, Jensen TP, Liu G, Mathews BK, Schott CK, Kurian L, LoPresti CM, et al. Comparison of in-person versus tele-ultrasound point-of-care ultrasound training during the COVID-19 pandemic. Ultrasound J 2021;13(1):1–7.
[11] Priester AM, Natarajan S, Culjat MO. Robotic ultrasound systems in medicine. IEEE Trans Ultrason Ferroelectr Freq Control 2013;60(3):507–23.
[12] Salcudean SE, Moradi H, Black DG, Navab N. Robot-assisted medical imaging: A review. Proc IEEE 2022.
[13] Vieyres P, Poisson G, et al. A tele-operated robotic system for mobile tele-echography: The OTELO project. In: M-Health. Springer; 2006, p. 461–73.
[14] Salcudean SE, Bell G, Bachmann S, Zhu W-H, Abolmaesumi P, Lawrence PD. Robot-assisted diagnostic ultrasound–design and feasibility experiments. In: International conference on medical image computing and computer-assisted intervention. Springer; 1999, p. 1062–71.
[15] Akbari M, Carriere J, Meyer T, Sloboda R, Husain S, Usmani N, Tavakoli M. Robotic ultrasound scanning with real-time image-based force adjustment: Quick response for enabling physical distancing during the COVID-19 pandemic. Front Robot AI 2021;8. http://dx.doi.org/10.3389/frobt.2021.645424.
[16] Delgorge C, Courrèges F, Bassit LA, Novales C, Rosenberger C, Smith-Guerin N, Brù C, Gilabert R, Vannoni M, Poisson G, et al. A tele-operated mobile ultrasound scanner using a light-weight robot. IEEE Trans Inf Technol Biomed 2005;9(1):50–8.
[17] Mathiassen K, Fjellin JrE, Glette K, Hol PK, Elle OJ. An ultrasound robotic system using the commercial robot UR5. Front Robot AI 2016;3:1.
[18] Li K, Xu Y, Meng MQ-H. An overview of systems and techniques for autonomous robotic ultrasound acquisitions. IEEE Trans Med Robot Bionics 2021;3(2):510–24.
[19] Wang S, Housden J, Noh Y, Singh D, Singh A. Robotic-assisted ultrasound for fetal imaging: Evolution from single-arm to dual-arm system. In: Annual conference towards autonomous robotic systems. Vol. 11650, July 2019, p. 27–38.
[20] Jiang Z, Grimm M, Zhou M, Hu Y, Esteban J, Navab N. Automatic force-based probe positioning for precise robotic ultrasound acquisition. IEEE Trans Ind Electron 2020;68(11):11200–11.
[21] Virga S, Zettinig O, Esposito M, Pfister K, Frisch B, Neff T, Navab N, Hennersperger C. Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms. In: 2016 IEEE/RSJ international conference on intelligent robots and systems. IROS, IEEE; 2016, p. 508–13.
[22] Huang Q, Lan J, Li X. Robotic arm based automatic ultrasound scanning for three-dimensional imaging. IEEE Trans Ind Inf 2018;15(2):1173–82.
[23] Ning G, Zhang X, Liao H. Autonomic robotic ultrasound imaging system based on reinforcement learning. IEEE Trans Biomed Eng 2021;68(9):2787–97.
[24] Carmigniani J, Furht B. Augmented reality: an overview. In: Handbook of augmented reality. Springer; 2011, p. 3–46.
[25] Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst 1994;77(12):1321–9.
[26] Skarbez R, Smith M, Whitton M. Revisiting milgram and kishino's reality-virtuality continuum. Front. In: Presence and beyond: evaluating user experience in AR/MR/VR. Frontiers Media SA; 2022, p. 8.
[27] Rokhsaritalemi S, Sadeghi-Niaraki A, Choi S-M. A review on mixed reality: Current trends, challenges and prospects. Appl Sci 2020;10(2):636.
[28] Black D, Oloumi Yazdi Y, Hadi Hosseinabadi AH, Salcudean S. Human teleoperation-a haptically enabled mixed reality system for teleultrasound. Hum–Comput Interact 2023;1–24.

[29] Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. ACM SIGGRAPH Comput Graph 1992;26(2):203–10.
[30] Rosenthal M, State A, Lee J, Hirota G, Ackerman J, Keller K, Pisano ED, Jiroutek M, Muller K, Fuchs H. Augmented reality guidance for needle biopsies: an initial randomized, controlled trial in phantoms. Med Image Anal 2002;6(3):313–20.
[31] Groves L, Li N, Peters TM, Chen E. Towards a first-person perspective mixed reality guidance system for needle interventions. J Imaging 2022;8(1):7.
[32] Ameri G, Baxter JS, Bainbridge D, Peters TM, Chen E. Mixed reality ultrasound guidance system: a case study in system development and a cautionary tale. Int J Comput Assist Radiol Surg 2018;13(4):495–505.
[33] Navab N, Traub J, Sielhorst T, Feuerstein M, Bichlmeier C. Action-and workflow-driven augmented reality for computer-aided medical procedures. IEEE Comput Graph Appl 2007;27(5):10–4.
[34] Kolagunda A, Sorensen S, Mehralivand S, Saponaro P, Treible W, Turkbey B, Pinto P, Choyke P, Kambhamettu C. A mixed reality guidance system for robot assisted laparoscopic radical prostatectomy. In: OR 2.0 context-aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures, and skin image analysis. Springer; 2018, p. 164–74.
[35] Schneider CM, Dachs GW, Hasser CJ, Choti MA, DiMaio SP, Taylor RH. Robot-assisted laparoscopic ultrasound. In: Information processing in computer-assisted interventions: first international conference, IPCAI 2010, Geneva, Switzerland, June 23, 2010. Proceedings 1. Springer; 2010, p. 67–80.
[36] Kalia M, Avinash A, Navab N, Salcudean S. Preclinical evaluation of a markerless, real-time, augmented reality guidance system for robot-assisted radical prostatectomy. Int J Comput Assist Radiol Surg 2021;16(7):1181–8.
[37] Leuze C, Yang G, Hargreaves B, Daniel B, McNab JA. Mixed-reality guidance for brain stimulation treatment of depression. In: 2018 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct). IEEE; 2018, p. 377–80.
[38] Ntourakis D, Memeo R, Soler L, Marescaux J, Mutter D, Pessaux P. Augmented reality guidance for the resection of missing colorectal liver metastases: an initial experience. World J Surg 2016;40(2):419–26.
[39] Cartucho J, Shapira D, Ashrafian H, Giannarou S. Multimodal mixed reality visualisation for intraoperative surgical guidance. Int J Comput Assist Radiol Surg 2020;15(5):819–26.
[40] Song H, Moradi H, Jiang B, Xu K, Wu Y, Taylor RH, Deguet A, Kang JU, Salcudean SE, Boctor EM. Real-time intraoperative surgical guidance system in the da vinci surgical robot based on transrectal ultrasound/photoacoustic imaging with photoacoustic markers: an ex vivo demonstration. IEEE Robot Autom Lett 2022.
[41] Wild E, Teber D, Schmid D, Simpfendörfer T, Müller M, Baranski A-C, Kenngott H, Kopka K, Maier-Hein L. Robust augmented reality guidance with fluorescent markers in laparoscopic surgery. Int J Comput Assist Radiol Surg 2016;11(6):899–907.
[42] Nee AY, Ong S, Chryssolouris G, Mourtzis D. Augmented reality applications in design and manufacturing. CIRP Ann 2012;61(2):657–79.
[43] Mourtzis D, Zogopoulos V, Vlachou E. Augmented reality application to support remote maintenance as a service in the robotics industry. Procedia Cirp 2017;63:46–51.
[44] Orts-Escolano S, Rhemann C, Fanello S, Chang W, Kowdle A, Degtyarev Y, Kim D, Davidson PL, Khamis S, Dou M, et al. Holoportation: Virtual 3d teleportation in real-time. In: Proceedings of the 29th annual symposium on user interface software and technology. 2016, p. 741–54.
[45] Black D, Salcudean S. Mixed reality human teleoperation. In: 2023 IEEE conference on virtual reality and 3d user interfaces abstracts and workshops. VRW, IEEE; 2023, p. 375–83.
[46] Black D, Salcudean S. Human-as-a-robot performance in mixed reality teleultrasound. Int J Comput Assist Radiol Surg 2023;1–8.
[47] Black D, Andjelic D, Salcudean S. Evaluation of communication and human response latency for (human) teleoperation. 2022, http://dx.doi.org/10.36227/techrxiv.21432009.v1, TechRxiv.
[48] Ungureanu D, Bogo F, Galliani S, Sama P, Duan X, Meekhof C, Stühmer J, Cashman TJ, Tekin B, Schönberger JL, et al. Hololens 2 research mode as a tool for computer vision research. 2020, arXiv preprint arXiv:2008.11239.
[49] Black D, Salcudean S. Robust, infrared marker-based object tracking for augmented reality guidance and teleoperation. 2023, TechRxiv.
[50] Massie TH, Salisbury JK, et al. The phantom haptic interface: A device for probing virtual objects. In: Proceedings of the ASME winter annual meeting, symposium on haptic interfaces for virtual environment and teleoperator systems. Vol. 55, Chicago, IL; 1994, p. 295–300.
[51] Lawrence DA. Stability and transparency in bilateral teleoperation. IEEE Trans Robot Autom 1993;9(5):624–37.
[52] Hashtrudi-Zaad K, Salcudean SE. Transparency in time-delayed systems and the effect of local force feedback for transparent teleoperation. IEEE Trans Robot Autom 2002;18(1):108–14.
[53] Black D, Salcudean S. A mixed reality system for human teleoperation in tele-ultrasound. In: Hamlyn symposium for medical robotics. 2022, p. 91–2.
[54] RFC 4960: Stream control transmission protocol. 2007, URL https://www.rfc-editor.org/rfc/rfc4960.
[55] RFC 1889: RTP: A transport protocol for real-time applications. 1996, URL https://www.rfc-editor.org/rfc/rfc1889.
[56] Jay C, Hubbold R. Delayed visual and haptic feedback in a reciprocal tapping task. In: First joint eurohaptics conference and symposium on haptic interfaces for virtual environment and teleoperator systems. World haptics conference. IEEE; 2005, p. 655–6.
[57] Jay C, Glencross M, Hubbold R. Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. ACM Trans Comput-Hum Interact 2007;14(2):8–es.
[58] Kaber DB, Zhang T. Human factors in virtual reality system design for mobility and haptic task performance. Rev Hum Fact Ergonom 2011;7(1):323–66.
[59] Garrido-Jurado S, Muñoz-Salinas R, Madrid-Cuevas FJ, Marín-Jiménez MJ. Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognit 2014;47(6):2280–92.
[60] Schimmoeller T, Colbrunn R, Nagle T, Lobosky M, Neumann EE, Owings TM, Landis B, Jelovsek JE, Erdemir A. Instrumentation of off-the-shelf ultrasound system for measurement of probe forces during freehand imaging. J Biomech 2019;83:117–24.
[61] Gilbertson MW, Anthony BW. An ergonomic, instrumented ultrasound probe for 6-axis force/torque measurement. In: 2013 35th annual international conference of the IEEE engineering in medicine and biology society. EMBC, IEEE; 2013, p. 140–3.
[62] Harris-Love MO, Ismail C, Monfaredi R, Hernandez HJ, Pennington D, Woletz P, McIntosh V, Adams B, Blackman MR. Interrater reliability of quantitative ultrasound using force feedback among examiners with varied levels of experience. PeerJ 2016;4:e2146.
[63] Mylonas GP, Giataganas P, Chaudery M, Vitiello V, Darzi A, Yang G-Z. Autonomous eFAST ultrasound scanning by a robotic manipulator using learning from demonstrations. In: 2013 IEEE/RSJ international conference on intelligent robots and systems. IEEE; 2013, p. 3251–6.
[64] Black D, Hosseinabadi AHH, Pradnyawira NR, Pol M, Nogami M, Salcudean S. Towards differential magnetic force sensing for ultrasound teleoperation. In: 2023 IEEE world haptics conference. WHC, IEEE; 2023, p. 333–9.
[65] Black D, Pradnyawira NR, Nogami M, Hosseinabadi AHH, Salcudean S. Low-profile, 6-axis differential magnetic force/torque sensing. 2023, TechRxiv.
[66] Hannaford B, Anderson R. Experimental and simulation studies of hard contact in force reflecting teleoperation. In: Proceedings. 1988 IEEE international conference on robotics and automation. IEEE; 1988, p. 584–9.
[67] Daniel R, McAree PR. Fundamental limits of performance for force reflecting teleoperation. Int J Robot Res 1998;17(8):811–30.
[68] Reboulet C, Plihon Y, Briere Y. Interest of the dual hybrid control scheme for teleoperation with time delays for proceeding of ISER'95. In: Experimental robotics IV. Springer; 1997, p. 498–506.
[69] Hastrudi-Zaad K, Salcudean SE. On the use of local force feedback for transparent teleoperation. In: Proceedings 1999 IEEE international conference on robotics and automation (cat. no. 99CH36288C). Vol. 3, IEEE; 1999, p. 1863–9.
[70] Salcudean SE, Hashtrudi-Zaad K, Tafazoli S, DiMaio SP, Reboulet C. Bilateral matched impedance teleoperation with application to excavator control. IEEE Control Syst Mag 1999;19(6):29–37.
[71] Niemeyer G, Slotine J-J. Stable adaptive teleoperation. IEEE J Ocean Eng 1991;16(1):152–62.
[72] Ye Y, Liu PX. Improving trajectory tracking in wave-variable-based teleoperation. IEEE/ASME Trans Mechatronics 2009;15(2):321–6.
[73] Misra S, Okamura AM. Environment parameter estimation during bilateral telemanipulation. In: 2006 14th symposium on haptic interfaces for virtual environment and teleoperator systems. IEEE; 2006, p. 301–7.
[74] Stocco L, Salcudean SE. A coarse-fine approach to force-reflecting hand controller design. In: Proceedings of IEEE international conference on robotics and automation. Vol. 1, IEEE; 1996, p. 404–10.
[75] Julesz B. Binocular depth perception of computer-generated patterns. Bell Syst Tech J 1960;39(5):1125–62.
[76] Ichihara S, Kitagawa N, Akutsu H. Contrast and depth perception: Effects of texture contrast and area contrast. Perception 2007;36(5):686–95.