

Human Teleoperation - A Haptically Enabled Mixed Reality System for Teleultrasound

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

16-08-2021 / 19-08-2021

CITATION

Black, David; Yazdi, Yas Oloumi; Hadi Hosseinabadi, Amir Hossein; Salcudean, Septimiu (2021): Human Teleoperation - A Haptically Enabled Mixed Reality System for Teleultrasound. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.15175869.v1>

DOI

[10.36227/techrxiv.15175869.v1](https://doi.org/10.36227/techrxiv.15175869.v1)

Human Teleoperation - A Haptically Enabled Mixed Reality System for Teleultrasound

David Black^{*†}, Yas Oloumi Yazdi[†], Amir Hossein Hadi Hosseinabadi^{*}, Septimiu Salcudean^{*} *Fellow, IEEE*

^{*}Robotics and Control Lab, University of British Columbia, Vancouver, Canada [†]Engineering Physics, University of British Columbia, Vancouver, Canada

Abstract—Current teleguidance methods include verbal guidance and robotic teleoperation, which present tradeoffs between precision and latency versus flexibility and cost. We present a novel concept of "human teleoperation" which bridges the gap between these two methods. A prototype teleultrasound system was implemented which shows the concept's efficacy. An expert remotely "teleoperates" a person (the follower) wearing a mixed reality headset by controlling a virtual ultrasound probe projected into the person's scene. The follower matches the pose and force of the virtual device with a real probe. The pose, force, video, ultrasound images, and 3-dimensional mesh of the scene are fed back to the expert. In this control framework, the input and the actuation are carried out by people, but with near robot-like latency and precision. This allows teleguidance that is more precise and fast than verbal guidance, yet more flexible and inexpensive than robotic teleoperation. The system was subjected to tests that show its effectiveness, including mean teleoperation latencies of 0.27 seconds and errors of 7 mm and 6° in pose tracking. The system was also tested with an expert ultrasonographer and four patients and was found to improve the precision and speed of two teleultrasound procedures.

Index Terms—Teleoperation, Tele-medicine, Tele-ultrasound, Mixed Reality, Augmented Reality, Haptics

I. INTRODUCTION

The fourth industrial revolution, or Industry 4.0, is expected to bring higher industrial performance and efficiency through the adoption of emerging technologies in robotics, artificial intelligence, cloud computing, and mixed reality [1]. The same technologies are having an even more immediate impact on healthcare and medicine [2]. However, there is a certain disconnect between the technologies and their application. Many companies are unsure how to take advantage of Industry 4.0 to improve their business [3], while for many medical applications, the technology is not at a level where it can be used directly on patients, or it simply does not fit the application as well as desired.

One such problem is teleultrasound. In remote areas, access to expert care and diagnosis by sonographers is often severely lacking or infrequent [4]. By enabling expert sonographers to remotely guide or teleoperate ultrasound (US) procedures in these communities, teleultrasound has immense potential to improve the quality of care of patients, both in rural regions and in ambulances. It can also decrease costs associated with transporting the patients or medical workers, and increase safety in a pandemic such as COVID-19 [5].

Ultrasound teleguidance systems have been implemented by numerous groups. For trauma patients, verbal guidance via radio while viewing a stream of the ultrasound images was explored by Boniface et al. [6]. More modern systems sold by Clarius Mobile Health Corp. and Butterfly Network combine a mobile phone application with a wireless ultrasound transducer and remote access to the images and video conferencing via a cloud interface [7]. However, in all these solutions the instructions for probe positioning, orientation, and force are given verbally or with limited augmented reality overlays of arrows or pointers, which is very inefficient, leading to high latency and low precision.

Conversely, robotic teleultrasound systems have also been developed which provide low latencies and high precision, as well as haptic feedback [8][9][10][11]. These involve a robotic arm with ultrasound probe end effector which is teleoperated by a remote expert sonographer. Salcudean et al. presented a robot whose control was shared between the expert and a visual servoing system to maintain correct positioning on the carotid artery [12]. Another system, named OTELO [13][14], has demonstrated clinical utility in trials [15]. Recent work has even investigated the control of such systems over 5G and in the context of COVID-19 [16].

However, there are many drawbacks with robotic systems. While some are designed to be inherently backdriveable and lightweight [17], the issues of safe human-robot interaction and predictable and consistent autonomy remain unsolved [18]. As a result, a human follower is usually still needed on-site to monitor the robot [19], and potentially check and approve planned motion trajectories. This limits the efficiency of such systems. Furthermore, such robots have restricted workspaces, are time consuming to set up, too large to store on ambulances, and incongruously expensive compared to ultrasound systems. While ultrasound is usually an inexpensive procedure and is thus well suited to being a standard of care in remote communities, installing an expensive robot in every small town is infeasible.

In this paper, we introduce the concept of "human teleoperation" to bridge the gap between teleguidance and robotic systems. In human teleoperation the follower, or person carrying out the procedure on site, is guided by a remote expert through a real-time, mixed reality (MR) interface on a Microsoft HoloLens 2. A 3-dimensional (3D) virtual ultrasound transducer controlled by the expert is projected into the follower's environment for the follower to follow. In terms of classical teleoperation concepts [20], the "remote" robot acting on the

environment is replaced by a human "follower" (Fig. 1). The follower copies the desired position (P_o) and force (F_o) of the "master" or "expert" by aligning the tool to its MR projection on a HoloLens 2 worn by the follower. In turn, the expert is presented visually with the end-effector pose (P') via an MR capture of the follower's environment with the virtual tool in place, as well as the forces (F'), if sensed, returned through a haptic device.

The key enabling technology for this system is mixed reality. While augmented reality (AR) captures the real environment and renders it on a screen, for example on a smartphone or tablet, where virtual cues can be embedded into the scene, MR projects the 3D virtual objects into the real environment using a partially-transparent headset. This allows the follower wearing the MR headset to interact seamlessly with both the real environment and the virtual objects. The idea of using augmented and mixed reality to aid in medical procedures has been explored extensively, from providing guidance for tissue biopsies by overlaying medical images and guiding pointers [21][22], to training and simulation [23][24][25]. In teleultrasound, several patents for using augmented reality interfaces to guide ultrasound procedures have been filed by Butterfly Networks, Inc. [26] [27] and others [28] [29]. However, these current implementations are limited to instructional text and overlaid arrows and indicators that are placed by the expert.

The use of AR and MR to provide remote assistance has been used not only in telemedicine, but in countless industries including manufacturing and remote maintenance. Masoni et al. created an augmented reality system that places helpful labels and 2D text in the follower's scene to assist them in their task [30]. Conversely, Mourtzis et al. developed a framework to obtain information about a scene and create an AR application off-line which contains visual instructions that can be overlaid onto the scene [31].

All the AR/MR tele-assistance solutions mentioned above are static or pre-planned, are applied only to predefined, known environments, or include only simplistic labels and arrows for guidance. Thus, our human teleoperation concept provides several contributions, which we frame here in terms of a teleultrasound system, but which are trivially extended to other applications. Our system:

- Allows the expert to dynamically control a 3D virtual object such as a virtual ultrasound probe in the follower's scene in real time, so the follower can follow its pose with their real probe.
- Captures the 3-dimensional follower-side scene on demand and relays it to the expert so the expert can interact with it visually and haptically.
- Allows the expert to provide input by directly manipulating a dummy ultrasound probe.
- Includes haptic feedback so the expert has the sensation of touching the actual patient, and can guide the follower's input force.

These contributions form the basis of the human teleoperation system proposed in this paper. They allow teleguidance that is more precise, intuitive, and with lower latency than verbal guidance, yet more flexible, inexpensive, accessible, and more feasible than robotic teleultrasound. By providing

a control framework where both the input and the actuation are carried out by people, this system can be deployed in any new, unfamiliar environment, and faces none of the regulatory problems related to unpredictable and potentially unsafe behaviour of robotic systems.

In the following sections, the human teleoperation system will be introduced in the context of teleultrasound. First, the application-specific requirements and design objectives are discussed (Section II-A). In Section II-B and those following, the implementation of a prototype system is shown, before illustrating how it can be extended to other applications. Finally, tests were carried out to validate the effectiveness of the system. The results are found in Section III, and the system's limitations are discussed in Section IV.

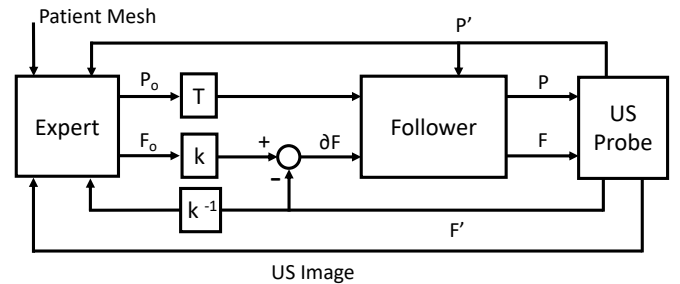


Fig. 1. Block Diagram of the Human Teleoperation System: The desired position, P_o , is transformed to the follower coordinate system using transform T . The forces are scaled down at the expert side by a factor of k^{-1} . Instead of a controller and actuators at the "slave" side, there is a human "follower".

II. METHODS

A. Design Objectives

Our research goal was to design and build a system that has the high precision and low latency of robotic teleultrasound without all the disadvantages listed in Section I. In particular, we aimed to achieve a small error between the desired and actual pose and force, and low latency between issuing a command and achieving the desired state. It has also been shown that haptic feedback for the expert improves teleoperation task performance [32] and is more intuitive for the expert, so transparency was also an objective in this system. The expert should have the sensation of touching the actual patient and should be able to guide the follower's force without distracting the follower from following the pose. While these objectives can be achieved in a robotic system, we additionally aimed to make the patient-side interface wireless and portable. The system should be fast to set up, accessible, inexpensive (compared to a robot), and intuitive to use for both the expert and the follower. Furthermore, through meetings with expert sonographers of the British Columbia Ultrasonographers' Society, it was established that high quality ultrasound image transmission and a video conferencing interface are essential.

B. System Overview

The teleultrasound system consists of two distinct halves, the follower side and the expert side, which communicate

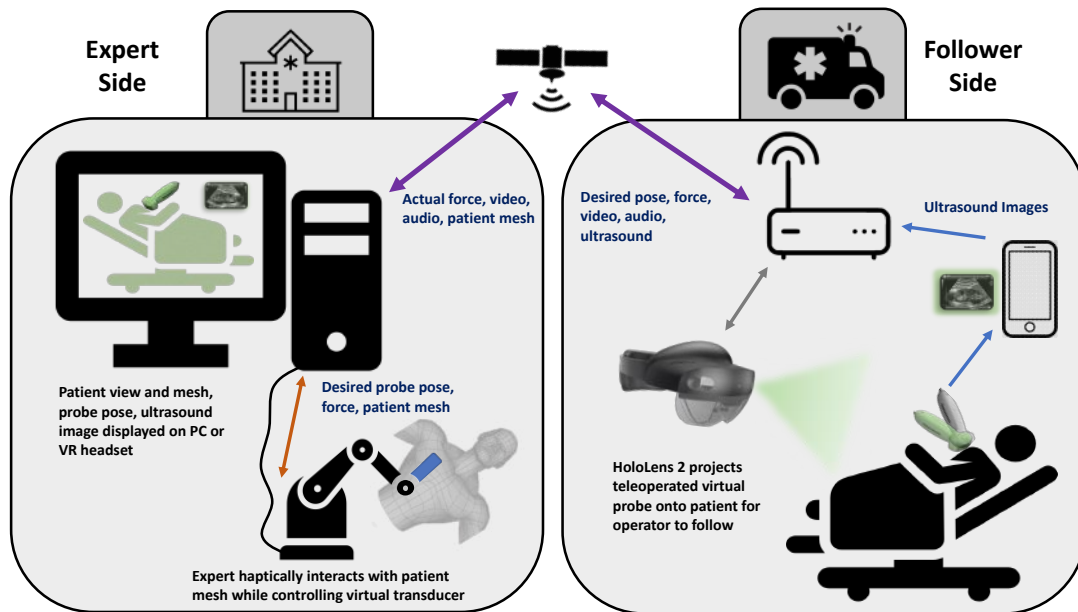


Fig. 2. System Level Diagram: The follower wears a Microsoft HoloLens 2 which projects a virtual transducer into the follower's scene. The expert controls this virtual probe using a haptic controller while observing the ultrasound images obtained by the follower. The expert and follower communicate via a (Mixed Reality WebRTC) video call interface and the probe pose, force, and patient mesh are sent by a WebSocket.

wirelessly as explained in Section II-C. A conceptual overview of the system is seen in Fig. 2.

The follower wears a Microsoft HoloLens 2 which projects a virtual ultrasound transducer into the follower's scene. The expert remotely controls this virtual probe using a haptic controller (Phantom Omni, 3D Systems, Inc) to input the desired pose (position and orientation) and force. The follower follows the virtual probe with the real probe, thus achieving the human teleoperation. The follower-side interface is seen in Fig. 3, with a few frames showing the teleoperation. At the same time, the live ultrasound images are transmitted wirelessly from a handheld ultrasound device (C3HD, Clarius Mobile Health, Vancouver, BC) to the follower's smartphone and the expert PC. The HoloLens 2 also captures an MR video of the scene with the MR overlays in position (known as an MR capture) and shares these live with the expert via a WebRTC interface for positional feedback. In this way, the expert receives the high quality ultrasound images in real time, can see the actual patient with the virtual and real probes, and is in verbal communication with the follower.

Additionally, the follower can send a spatial mesh of the patient, generated automatically by the HoloLens 2, to the expert on demand (Section II-F). This mesh is rendered haptically as a virtual fixture for the Phantom Omni, giving the expert the sensation that they are physically touching and interacting with the patient (Section II-D). Finally, the mesh is shown on the expert PC along with the virtual transducer in position for further pose feedback. This also allows the virtual transducer pose to be registered to the real patient, as explained in Section II-E.

While the haptic device is used to control fine pose, the rough positioning can be changed on the expert side using the PC's arrow keys, and on the follower side by pinching

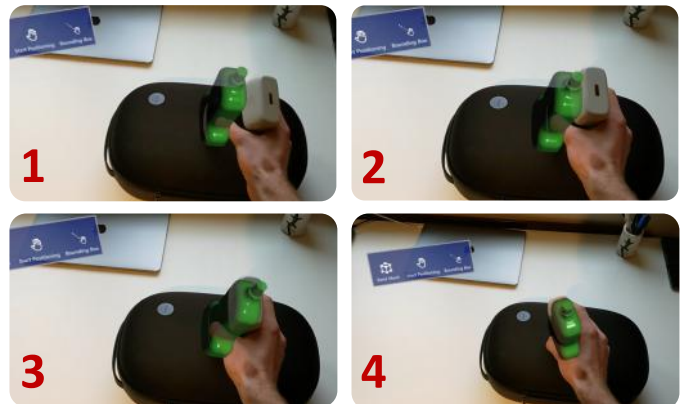


Fig. 3. Follower-Side Interface and Teleoperation: The follower sees the virtual transducer and a control menu. In frames 1-3 the follower starts in a random position and matches the virtual probe pose precisely within < 0.5 sec, before the probe pose changes and is immediately matched by the follower again in frame 4. The transducer is green to indicate the expert is happy with the applied force.

and dragging the virtual probe. When the follower changes the probe position, the input from the haptic device is ignored to avoid conflicting pose commands. The haptic controller is also used to input the desired force, which is displayed on the follower side by changing the color of the virtual transducer. In this way, the follower receives feedback on the applied force without being distracted from the pose control. The force applied to the ultrasound probe by the follower is an important part of obtaining a quality ultrasound image.

Finally, the expert views the ultrasound images, MR capture, and patient mesh with virtual transducer in position on the monitor of the expert PC, as shown in Fig. 4. The expert PC application can be viewed immersively on a virtual reality



Fig. 4. Expert Workstation: the Phantom Omni haptic controller (front) is used to input pose and force and provide haptic feedback. The virtual probe relative to the patient mesh is visualized with the live ultrasound images and MR capture video call shown above it.

headset, if desired. This further increases the immersive and realistic nature of the expert side teleoperation interface, and allows more intuitive visualization of the virtual probe on the patient mesh in 3D.

With this overview in mind, the following subsections explain the system design in more detail.

C. System Architecture and Communication

This section explores in detail the implementation of each component and how they all communicate. Fig. 5 shows the different communication layers and what data is sent through which interface. This mirrors Fig. 1, but shows how each connection is implemented. The required bandwidths are listed in Table I.

Chan et al. showed that data speeds of at least 1Mbps are needed for high quality transmission of ultrasound images [33]. However, with more modern imaging systems and higher expectations for quality and frame rate, this may be substantially higher. In addition, the sonographers stressed the importance of an audio/video conferencing system, which adds another several Mbps. The transducer pose and force have to be transmitted at a high rate for haptic feedback, and finally a spatial mesh of the patient measured by the HoloLens 2 is sent as well (See Section II-D). The bandwidth accounting is shown in Table I. In total, the data being communicated may amount to up to 10Mbps peak. Given these large bandwidths, a 5G system would be ideal for the remote operation. However, this proof-of-concept prototype was developed to run on local networks only, and extension to 5G is a future improvement.

Starting on the right side of Fig. 5, the HoloLens 2 provides the main interface for the follower through a Unity application built with the Microsoft Mixed Reality Toolkit (MRTK). It must receive the desired pose and force from the expert and the actual force from the follower side, and send the patient mesh as well as MR captures of the scene. All communication between the expert PC and the HoloLens 2 is achieved via the rosbridge suite [34] except the MR capture and audio com-

TABLE I
BANDWIDTH ACCOUNTING: APPROXIMATION OF REQUIRED BANDWIDTH FOR TELEULTRASOUND SYSTEM. THE TOTAL IS ABOUT 7-10MBPS. THE VIDEO BANDWIDTH IS FOR A 720P, 30FPS VIDEO, WHICH WAS DEEMED SUFFICIENT. POSE AND FORCE ARE 16-BIT FLOATING-POINT NUMBERS, AND ORIENTATION IS SENT AS A QUATERNION.

Data	Size	Rate	Bandwidth
Force/Torque	48 bits	100Hz	4.8kbps
Position	48 bits	100Hz	4.8kbps
Orientation	64 bits	100Hz	6.4kbps
Video	-	30Hz	4Mbps
Audio	-	-	96kbps
Ultrasound	-	-	3Mbps

munication which are sent using Microsoft's Mixed Reality WebRTC API.

Rosbridge is an API which allows Robot Operating System (ROS) communication networks to be extended from a single device to a distributed set of devices on a local wireless network. These remote devices each run one of the rosbridge client libraries (ROS# for C#, roslibpy for Python, roslibjs for Javascript) through which they can publish and subscribe to ROS topics, actions, and services. The ROS messages are first serialized into JSON (JavaScript Object Notation) before being sent to the rosbridge server on the expert PC via a WebSocket interface, which facilitates the high-speed, persistent connection needed for this application.

In the teleultrasound system, the rosbridge server is set up on a Windows Subsystem for Linux (WSL) running Ubuntu 18.04 on the expert PC. This allows for seamless integration with the expert's Unity application and Phantom Omni drivers, which require Windows. Both the expert and follower user interfaces are 3D graphics applications built in Unity (Unity Technologies, Inc) using C#. The expert and follower interfaces therefore communicate with rosbridge via ROS#, an open source rosbridge client library from Siemens. The HoloLens runs a different build of the library called ROS#-UWP, which is compatible with the Universal Windows Platform (UWP) architecture of the device. In order to minimize latency, the orientation of the probe is encoded as a quaternion. The mesh is also preprocessed to decrease the required data transfer. This is discussed in the following section.

The expert Unity application uses OpenHaptics SDK to drive the Phantom Omni and the haptic interactions, as well as OpenVR SDK to provide an optional immersive view on an Oculus Rift DK2 VR headset. As shown in Fig. 4, the live ultrasound images and MR capture are shown in the Unity application along with the patient mesh and virtual transducer. This gives the expert multiple channels of information to work with and make clinical and diagnostic decisions. Clarius Cast API by Clarius Mobile Health Corp. allows real time streaming of the ultrasound images from the wireless transducer to devices on the local network. The audio/video call uses the HoloLens 2's microphones and front-facing cameras to stream an MR capture, as described before.

D. Haptics

The control of pose and force, as well as force feedback to the expert are achieved using a Phantom Omni haptic

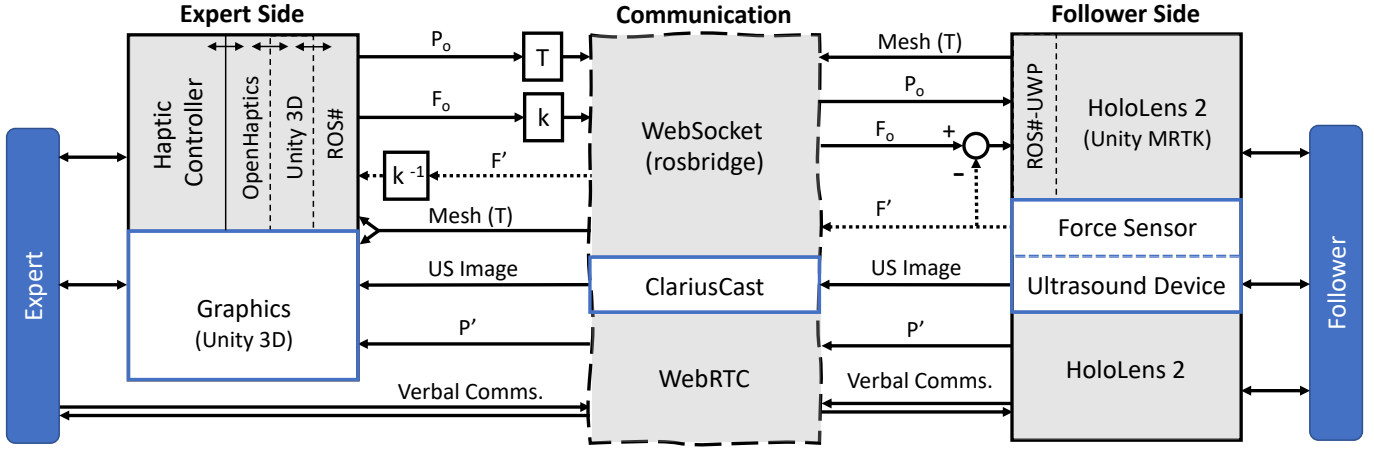


Fig. 5. Communication Architecture: The wireless communication (centre) utilizes rosbridge, Mixed Reality WebRTC, and ClariusCast API. The rosbridge core runs on the expert side PC, on a Windows Subsystem for Linux (WSL) Ubuntu 18.04 machine. Everything else on the expert PC requires Windows. The expert and follower user interfaces are both implemented in Unity, using OpenHaptics SDK and Mixed Reality Toolkit (MRTK) respectively to interface with the Phantom Omni and the HoloLens 2. Both sides access rosbridge via ROS#. The patient mesh captured by the HoloLens 2 provides the coordinate transform, T , as explained in Section II-E.

device. The Phantom Omni is a 6 degree of freedom serial arm with three actuated arm joints that can provide haptic feedback, three passive spherical wrist joints, and a stylus-like end effector with two buttons.

The expert determines whether more/less force is needed based on the quality of the ultrasound image, the video feed of the patient, and verbal communication with the follower. They then indicate the desired force through the haptic controller. Though the Phantom Omni used in this prototype can apply forces precisely, it is limited to 3.3N. In the 2-10 N force range, the human hand's just noticeable difference (JND) in force is about 10% [35], so for ultrasonographers accustomed to working in the 5-20 N range [36], a 10% JND is comparable in magnitude to the entire force range of the haptic device. Thus, in practice it was found to be very difficult to precisely modulate the applied force without saturating the device, making it impractical for the expert to directly input a force by pressing harder. Instead, for the proof-of-concept, the two buttons on the stylus end-effector are used to indicate "more force", "less force", or "good force". On the follower side this is shown by changing the color of the transducer. "More force" makes the probe red, "less force" turns it blue, and "good force" is green. In this way, the follower can remain completely focused on following the desired pose, and does not have to look away to determine the desired force.

In future work, methods for force sensing at the follower's ultrasound device will be investigated, as discussed in Section IV. For testing purposes in this work, a Raspberry Pi was set up to simulate force data and connect to rosbridge using Python's roslibpy library. This could in future be used to obtain the readings from a force sensor.

E. Pose Registration

The current haptic feedback system relies on the spatial mesh of the patient being used as a virtual fixture for the haptic controller to interact with. In addition, the mesh provides visual feedback for the expert regarding transducer positioning,

and facilitates the pose registration between the expert side virtual probe, the follower-side virtual probe, and the real patient as mentioned in Section II-B.

Let \mathcal{C}_o be the follower's world coordinate frame and \mathcal{C}_1 be the world frame in the expert application. The patient mesh is measured by the HoloLens 2 as a set of points in space, $\{x_i\}$, represented in the follower's head coordinate frame (where the HoloLens is worn) as $\mathcal{C}_h x_i$. The HoloLens provides accurate Simultaneous Localization and Mapping (SLAM) through its spatial awareness interface, so the transform $\mathcal{C}_h = \mathcal{C}_o \circ \mathcal{C}^h$ is known, which gives the real patient's vertices in space: $x_i = \mathcal{C}_o \circ \mathcal{C}^h x_i$. Now the virtual transducer is roughly positioned by the follower relative to the patient, as explained before. This sets the pose of the probe, \mathcal{C}_p , relative to the mesh on the follower side,

$$\mathcal{C}_p = \mathcal{C}_o \circ \mathcal{C}^p = \mathcal{C}_h (\mathcal{C}^h)^{-1} \circ \mathcal{C}^p \quad (1)$$

When the mesh is sent, it is placed in the expert's scene in the centre of the screen, at a comfortable distance from the camera. This determines the location of \mathcal{C}_h in the expert's world ($\mathcal{C}_h = \mathcal{C}_1 \circ \mathcal{C}^h$) since the mesh was defined relative to that coordinate system. Hence, using Eqn. 2 we can ascertain the probe's pose in the expert world:

$$\mathcal{C}_p = \mathcal{C}_1 \circ \mathcal{C}^h (\mathcal{C}^h)^{-1} \circ \mathcal{C}^p \quad (2)$$

Thus the registration is achieved. This gives the transform T in Figs. 1 and 5. The coordinate transforms are visualized in Fig. 6.

F. Mesh Management

The mesh is sent via the WebSocket and rosbridge, as explained in Section II-C, after some data preprocessing. The HoloLens constantly captures a spatial mesh of as much of the environment as it sees. However, for the teleultrasound system, only the patient's mesh is desired. Thus, a bounding box is defined which delineates from which region of space the

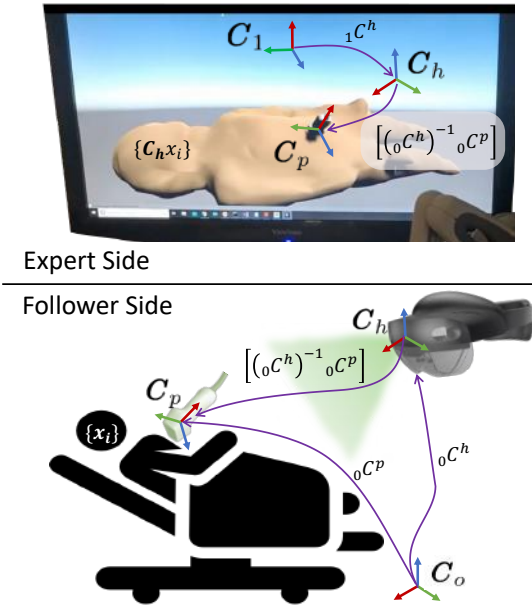


Fig. 6. Coordinate transforms on expert and follower sides, giving the registration of the virtual probe to the real patient and patient mesh on the follower and expert sides respectively.

mesh vertices for the patient should be extracted. This is shown in Fig. 7. This is done as follows: the follower is presented with three spherical markers when starting up the application. The follower pinches and drags the markers into position at any three corners of the patient's bed. The fourth corner is calculated automatically by finding a rectangle that minimizes the sum of the squared displacements required to make the other three markers coincident with its corners, and placing the final marker at the fourth corner. A semi-transparent plane spanning the rectangle is then shown and can be dragged to set the height of the bounding box to eliminate mesh points from the ceiling. The markers and plane are hidden a few seconds after hitting the "Finished" button, and can be recalled by pressing a button on the control menu to edit the bounding box.

When the follower presses the "Send Mesh" button on their menu, for example because the patient's position has changed, the follower is first encouraged to scan the patient with the HoloLens for 5 seconds to capture the required details. During this process, the mesh edges are projected onto the real world to give an idea of its quality and which areas should be improved by scanning over them. Each vertex of the mesh is then iterated through to check if it is within the bounding box. To do so, the point is first projected down into the plane of the defined rectangle. Each edge of the rectangle represents a half-space partition $\mathbf{a}_i^\top \mathbf{x} \leq b_i$, so in total the rectangle is a convex set of points defined by the intersection of the four half-spaces. By placing the four \mathbf{a}_i^\top vectors as the rows of a matrix, A , a mesh point's inclusion in the rectangle can easily be determined by checking if $A\mathbf{x} \leq \mathbf{b}$ (component-wise) and the vertical component is less than the bounding box height. This calculation is very low cost.

Any mesh triangles with only one vertex left are ignored,

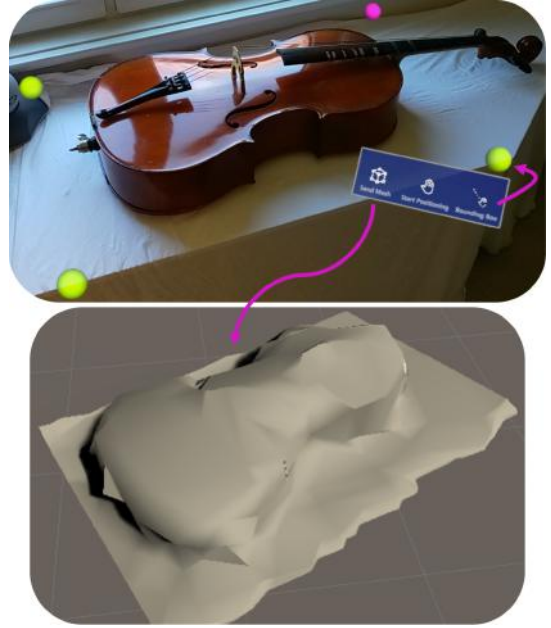


Fig. 7. Example Bounding Box Definition, Scanning, and Mesh Transfer of Patient-Like Shape: The top image is a mixed reality capture from the HoloLens 2. A bounding box is delineated using the green virtual markers after pressing the "Bounding Box" button on the menu. The fourth, pink, corner is automatically calculated, and the height is set. When the "Send Mesh" button is pressed, only the mesh from within the bounding box is sent. This is seen on the expert console in the bottom image.

while mesh triangles with two vertices in the bounding box are completed by replacing the third vertex with an average of the included two. This smooths the edges of the cropped patient mesh, which is then expressed as a list of vertex points (3-vectors) and a list of indices defining which points form triangles together. These are sent via ROS as a simple message containing float and int arrays, and are converted back to a Unity mesh on the expert side.

III. TESTING AND VALIDATION

In the design objectives, a number of goals were outlined involving latency, precision in position, orientation, and force, and fast and easy setup. In addition, the system aimed to be intuitive and easy to use for both the follower and expert. To verify that the human teleoperation concept implemented in the described prototype achieves these objectives and is effective in improving teleultrasound procedures, a number of tests were carried out by an ultrasound expert.

A. Data Latency:

To determine the latency of the rosbridge system for sending forces and poses, the time taken to receive 100 messages from the HoloLens and Raspberry Pi was measured for both types of data. The times were then divided by 100, to find the mean latencies for the rosbridge communication channels. The test was repeated 100 times, and the resultant latency histogram is shown in Fig. 8. Since the force sensor is not yet implemented, this only considers the latency of the communication, not of how long it takes to actually measure the forces. On the other

hand, the pose test first measures the current pose of the Phantom Omni. The latency for the video conferencing was determined by making an obvious, sharp sound which was picked up by the HoloLens 2, transmitted to the expert PC, and replayed loudly. A microphone recorded both sounds, and the delay time was determined in MATLAB. This test was repeated 20 times, and the same value was found on every trial. These results are summarized in Table II. The force and pose latencies are similar even though the pose involves more data, showing that they are limited by Unity rather than the communication system. Both meet the design objectives, and the WebRTC video conferencing system is sufficiently fast for efficient communication.

TABLE II

LATENCY OF THE COMMUNICATION SYSTEM: MEASURED LATENCIES OF THE 3 MAIN COMMUNICATION CHANNELS, TRANSMITTING FORCE FROM RASPBERRY PI TO EXPERT PC, POSE FROM EXPERT PC TO HOLOLENS 2, AND IMAGES TO EXPERT PC.

Datatype	Mean Latency
Force	11.2ms
Pose	11.6ms
MR Capture	160ms

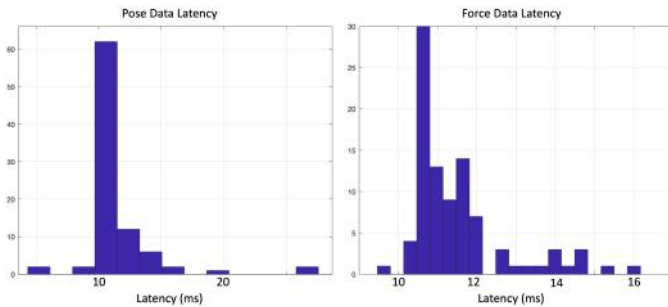


Fig. 8. Histograms of Force and Pose Communication Latencies: The time taken to send 100 messages of each datatype from the HoloLens and Raspberry Pi respectively to the expert PC was measured, divided by 100 to find the average, and repeated 100 times.

B. Teleoperation Latency and Precision:

The actual teleoperation is unlikely to be limited by latency in the communication system, but rather by the reaction times of the follower in following the virtual probe pose. To test the resulting latency of the system as a whole, as well as the precision of the teleoperation, two series of random motions were recorded using the haptic controller. Trial 1 consisted of smooth, continuous motions while trial 2 consisted of sharp motions followed by holding the pose for a few seconds (See Fig. 9). The latter series is much like a sequence of step response tests. Both series lasted about 150 seconds. An end-effector similar to the shell of the ultrasound device was mounted on the haptic controller, and each series was played back on the HoloLens while the follower followed the virtual probe pose with the real "probe" mounted on the haptic controller. In this way, the follower probe pose was also recorded by the controller so the expert and follower signals could be compared precisely.

C. Precision:

The precision was characterized separately for the position and orientation of the probe. For position, each axis was compared individually and an error signal was obtained by subtracting the leader and follower position elements. The signals for the series of sharp motions are plotted in Fig. 9. The RMS positional error of each axis and the resulting Euclidean displacement for both trials are found in Table III. Both trials show very similar positional results despite the different character of the motion. Both average values are slightly inflated because they include the initial large position error. The sharper motions in trial 2 are likely the reason why the mean offset in that trial is larger. The mean error is 36% of the width of the transducer head, which was 2 cm in these tests.

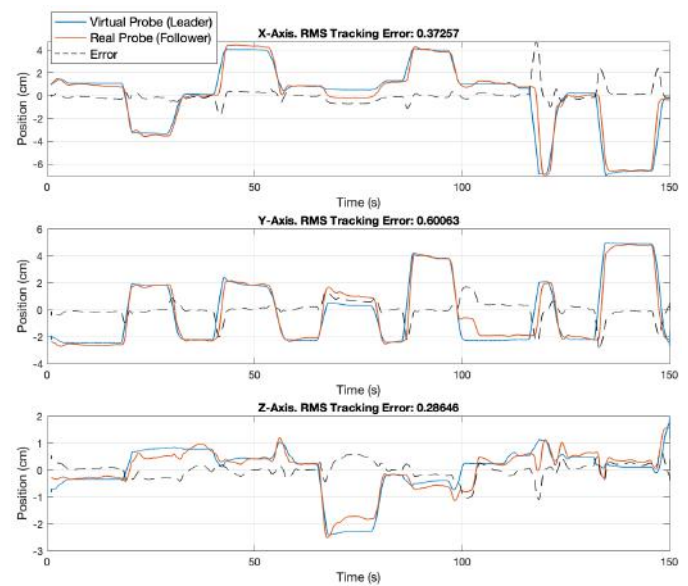


Fig. 9. Positional tracking of follower with error signal. The RMS position error was 3.7 mm, 6.0 mm, and 2.9 mm in the x, y, and z axes respectively.

TABLE III

RMS TRACKING ERROR OF EACH AXIS AND THE RESULTING EUCLIDEAN DISPLACEMENT IN TELEOPERATION. WE SEE SIMILAR ERRORS IN BOTH TRIALS DESPITE THE DIFFERENT CHARACTER OF THE MOTIONS.

Axis	X	Y	Z	Eucl.
Trial 1 Error (mm)	4.6	3.1	3.7	6.7
Trial 2 Error (mm)	3.7	6.0	2.9	7.6

To quantify the orientation error, the rotation quaternion from leader to follower was calculated at every time step and converted to its axis-angle representation to find the error as a single angular value in degrees. This is plotted for trial 1 in Fig. 10. The mean angular displacements between leader and follower were 5.87° and 6.89° for trial 1 and 2 respectively. Ignoring the high peaks above 12° where the orientation was suddenly changed more dramatically and the follower had not yet reacted, these errors are reduced to 5.2° and 5.5° respectively. These represent steady-state errors. As expected, the mean non-steady-state error in trial 2 is larger because the motions were sharper.

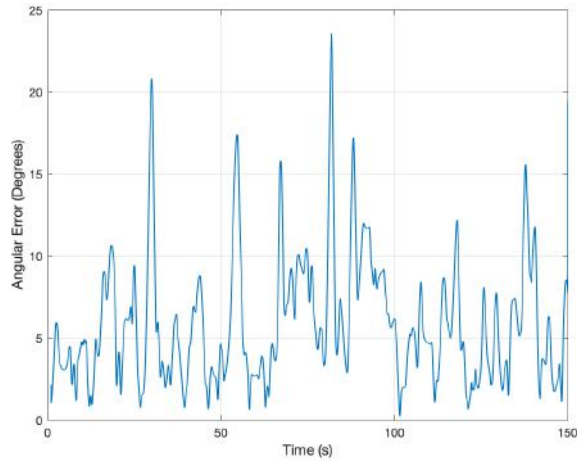


Fig. 10. Angular displacement between follower and leader vs. time, representing the orientation error. The sharp peaks are where the orientation was suddenly changed and the follower had not yet reacted.

In summary, the mean tracking error was measured to be 7.1 ± 0.3 mm and $6.3 \pm 0.5^\circ$ for general teleoperation, and smaller in smoother, slower motions as experienced in ultrasonography.

D. Latency:

Using the same measurements, it is possible to approximate the average latency of the teleoperation by determining the time delay between the leader and follower position signals. This is calculated by applying a varying time delay to the leader signal and maximizing the absolute value of the resulting normalized cross-correlation between the signals as a function of time delay. The approximate teleoperation latencies in the three positional axes are given in Table IV. On average, the total teleoperation latency from both the communication system and follower response time is 0.27 seconds.

TABLE IV
MEAN TELEOPERATION LATENCIES DETERMINED BY FINDING THE TIME DELAY IN THE LEADER POSITION SIGNAL THAT MAXIMIZED THE ABSOLUTE NORMALIZED CROSS-CORRELATION BETWEEN THE SIGNALS.

	X	Y	Z	Mean
Trial 1 Latency (sec)	0.34	0.27	0.31	0.31
Trial 2 Latency (sec)	0.40	0.30	0.01	0.24

E. Procedure Efficiency:

While the previous tests establish the efficacy of the human teleoperation architecture in general, it remains to be shown that this concept is useful in teleultrasound specifically. One of the primary benefits of this control scheme is that it should make remote ultrasound procedures faster and more precise by improving the efficiency of the communication through direct teleoperation. In order to verify that this is indeed the case, two procedures were carried out on two patients each, first directly by an expert sonographer, then by inexperienced subjects guided verbally by the expert, and finally by different

inexperienced subjects guided through human teleoperation by the expert.

The first test establishes the ground truth for the measured values and the time taken to complete the procedure. The second and third tests form a comparison between currently commercially available tele-guidance systems such as Clarius versus human teleoperation. The two procedures involved specific, quantitative endpoints so the effectiveness of the method could be quantified by comparison of the measured values, and the time taken to complete the procedure was well defined. The procedures were (1) measurement of the kidney size (length and width) and (2) measurement of the vena cava diameter. Each subject was teleoperated on one procedure and verbally guided on the other to avoid learning the procedure and thus introducing bias into the experiment. Procedure times and values differ between the patients due to differences in anatomy. However, these differences should cancel out when studying the percent changes in the metrics between tests on a given patient. Additionally, though one follower may be a faster learner than another, each follower participates in one test of each teleguidance method, so again no bias is introduced. The results are outlined in Table V. They show a clear improvement in both speed and precision using human teleoperation over existing systems.

IV. DISCUSSION

This paper introduces the concept of human teleoperation for the broad range of applications where the control system in Fig. 1 could be applied. To better understand the implementation challenges, performance, limitations, and efficacy of the concept, a prototype system was developed for teleultrasound. Through various tests it was shown that the teleoperation error is small: approximately 7 mm and 6° . While a human hand supported at the forearm can achieve accuracy up to 0.34 ± 0.16 mm, it is expected that an unsupported arm on a slippery surface like in ultrasonography has much lower accuracy [37]. Hence, the precision of the teleoperation system is approximately on the same order of magnitude as that of the human hand itself, which shows good performance. The latency is 0.27 sec on average, and the concept outperforms existing, commercially available teleguidance methods in both precision and speed.

While it has not been compared directly to robotic systems, the measured teleoperation precision and latency can be contrasted with the literature. For example, the robotic teleultrasound system described in [38] had a rise time of about 0.08 seconds. Stable teleoperation under time delays in various conditions has been studied in detail [39][40], though the delay can degrade performance. The most realistic model for the network-induced communication delays in this system is one of asymmetric, time-varying delays, which as shown in [41], can be teleoperated stably if the delays are less than 1 second. Thus, the 0.27 second latency of our system is well within the safe bounds and can enable a performant control system. Beyond these performance characteristics are important practical factors such as cost, portability, and setup time. Many existing robotic teleultrasound systems have used

TABLE V

RESULTS FROM TESTING WITH FOUR PATIENTS, FOUR INEXPERIENCED FOLLOWERS, AND ONE EXPERT. EACH PROCEDURE WAS CARRIED OUT DIRECTLY BY THE EXPERT, THEN USING VERBAL TELEOPERATION ON A CLARIUS SYSTEM, AND FINALLY USING HUMAN TELEOPERATION. SETUP TIME WAS NOT CONSIDERED. THIS TOOK LESS THAN 1 MINUTE FOR THE TELEOPERATION. ON AVERAGE, THE TELEOPERATION IS ONLY SLIGHTLY SLOWER THAN THE CONTROL, AND SUBSTANTIALLY FASTER THAN THE VERBAL METHOD. THE ACCURACY IS SIMILAR BUT ALSO SLIGHTLY BETTER IN THE TELEOPERATION. THE VARIATION IN THE VERBAL TIME RESULTS IS DISCUSSED IN THE DISCUSSION SECTION.

Procedure	Patient #	Control		Verbal		Teleoperation	
		Time	Value	Time	Value	Time	Value
Kidney	1	1:13	113 × 49mm	7:01	110 × 59mm	1:20	111 × 54mm
Kidney	2	1:43	118 × 50mm	1:25	123 × 46.5mm	1:52	112 × 46.9mm
Vena Cava	3	0:45	18.2mm	4:20	17.3mm	0:50	16.8mm
Vena Cava	4	0:39	17.4mm	3:30	21mm	0:47	15.9mm
<i>Averages</i>		Time	Error	Time	Error	Time	Error
Kidney		1:28 ± 0:21	0.0mm	4:13 ± 3:58	4 × 12mm	1:36 ± 0:23	4 × 4mm
Vena Cava		0:42 ± 0:04	0.0mm	3:55 ± 0:25	2.3mm	0:49 ± 0:02	1.5mm

large industrial manipulators [42][43][44][45], which are expensive and not portable. The follower first has to move the heavy robot into position, prepare it for use, and home it on the patient, making for a slow and potentially challenging setup. Custom, lighter-weight robotic ultrasound systems have also been developed [46][47] which are smaller but more complex to operate and likely very expensive. Conversely, in our system the follower simply puts on the HoloLens 2 and drags the virtual bounding box into position as shown in Fig. 3. The primary expenses for human teleoperation are the HoloLens 2 and Phantom Omni, which together cost a fraction of an industrial robot. Further, having a human follower rather than a robot is safer as human actuations are inherently passive [48]. Thus, the human teleoperation concept has multiple advantages over existing robotic systems as well as teleguidance methods.

Notice also that the only application specific aspects of the system described above are 1) the use of a virtual ultrasound probe, and 2) the transmission of ultrasound images. Thus, the teleultrasound system presented here can be extended easily to applications such as remote maintenance or manufacturing by replacing the virtual transducer with other virtual tools or devices that the follower is to use. The expert can also switch between a library of different virtual tools on demand, thus guiding not only the exact motion and force, but also which tool is being used. This would in fact be a simplification of the system presented here, as no ultrasound data would have to be transmitted. Further communication channels can trivially be added to the system by creating a new topic in the rosbridge network from any device such as a Raspberry Pi or PC that is connected to WiFi. In this way, further sensors and devices can be integrated as the application requires.

Though the results are promising, the implemented system also has certain limitations, which are discussed here. First, the tele-ultrasound system was implemented on local networks to allow rapid prototyping and development. However, to be truly useful in the real world, it would have to be expanded to run on external networks. With the advent of 5G, the required bandwidths outlined in Table I can easily be supported.

In addition, our system relies partly on the patient mesh to provide real-time, 3D positional and force feedback to the expert. However, though the mesh captured by the HoloLens 2 is sufficiently accurate to create a haptic surface of larger

anatomies for the expert to interact with, it does not closely resemble a person, in part because it captures only the broad shape, not the fine details of the patient, nor the color or texture (as seen in Fig. 7). It would therefore be of interest either to improve the resolution of the 3D perception to better capture the details of the patient, or to overlay the existing mesh with registered and deformed MR capture of the patient. Both would improve the expert's ability to recognize features of the patient to give anatomical context, and both could potentially be achieved using Microsoft's Research Mode APIs [49] using known methods for deformed registration and overlay [50][51]. This could also be useful in other fields such as manufacturing where a more precise mesh might be required.

A final limitation and area for further research is the haptics aspect of the system. As explained in Section II-D, the force control is currently almost entirely open-loop, with no force sensing at the ultrasound probe. This is in part because the choice of a force sensing method is very application dependent and may differ widely between teleultrasound and other applications of the human teleoperation concept such as manufacturing. To demonstrate the fundamental capabilities of this concept, therefore, rather than focusing too much on a specific application, the force sensing was predominantly left for a future publication.

However, to improve the reliability, accuracy, and transparency of the control system, the forces applied by the ultrasound probe should be determined [52]. This can be achieved by instrumenting the probe itself with a force sensor [53][54], or by estimating the forces visually using the HoloLens through recurrent neural networks [55] or with a model-based approach, looking at tissue deformation [56]. In this way, more complex force teleoperation architectures can also be implemented, including 4-channel teleoperation for optimal transparency [57]. Here the expert would not have a virtual fixture to interact with, but rather would have the exact forces applied by the follower on the patient reflected through the haptic controller. Future work would focus on stable and transparent force reflection for bilateral teleoperation under time delays. This has been studied extensively in the context of robotics, for example using passivity and scattering theory [58], wave variables [59][60], μ -synthesis [61], and an input-to-output stability small gain approach [62]. However, in this system the communication delays are imposed by the human

response time in the actuations, so this would constitute an interesting bridge between control theory and human teleoperation. In addition, the forces could be scaled down at the expert side to reduce fatigue and stress-related injuries common in ultrasonographers [63].

In order to realize these improvements, a more capable haptic device is required, as explained in the Section II-D. With the ability for the expert to input a precise force vector rather than a binary more/less, the rendering of the haptic feedback at the follower must be adapted as well. A continuous spectrum of colors can be used to indicate force magnitude, and an arrow for direction. Alternatively, a second virtual transducer could be positioned with a slight offset from the original, where the direction of the offset indicates the direction of desired force and the magnitude of the offset conveys the magnitude of the commanded force, proportional to some stiffness parameter. For example, to increase the pressure, the second virtual probe could be positioned further into the patient. Then the follower would push their probe harder into the patient to reach the second probe, thus increasing the force in that direction to equal the desired force.

While these are all implementation details and not fundamental limitations of the human teleoperation concept, the reaction time latency is the primary systematic limitation that affects the concept itself. The latencies presented in Section III represent close to the minimum possible response times because they are limited by the reaction time of the follower. Thus, this system can never achieve robot-level latency. However, this was clear from the start, and as the results show, the 0.27 second latency is relatively small, is much faster than alternative teleguidance techniques, and is well below the cutoff time delay for stable teleoperation given in [41]. Furthermore, the tests of pose error and latency studied unconstrained motion in three dimensions while in an ultrasound procedure the transducer is approximately constrained into two dimensions on the surface of the patient, so the error would likely be lower. On the other hand, the measured 0.27 second latency value will likely vary between followers and can be affected by external influences such as stress, fatigue, and distractions, which is unlike a robotic system.

In the expert ultrasound tests, the standard deviations in timing were large, and in one case the verbal communication was faster than the direct measurement. This instance was an outlier where the follower coincidentally set the initial pose so that little adjustment was necessary to obtain the image. This outlier, however, does not affect the outcome which shows improved precision and speed in human teleoperation compared to existing methods. Indeed, while the tested procedures were very simple, it is expected that the teleoperation will prove even more beneficial when used in longer and more involved procedures, for example with multiple measurements or with a qualitative aspect where the expert's judgement is needed. This is because the teleultrasound system provides the ability for them to consistently have the ultrasound exactly where they want it over an extended period of time, and it offsets the initial setup time which was never above 1 minute in testing.

V. CONCLUSION

In this paper, we presented a novel concept of "human teleoperation" through haptically-enabled mixed reality which bridges the gap between robotic and verbal methods of teleguidance. In this control framework, both the input and the actuation are carried out by people, but with near robot-like latency and precision. This allows teleguidance that is far more precise, intuitive, and low latency than verbal guidance, yet it is more flexible, inexpensive, and accessible than robotic teleoperation. A prototype system was implemented in the context of teleultrasound which shows the efficacy of the concept for a variety of potential applications including telemedicine, remote manufacturing, maintenance, and teaching. The system was subjected to a number of tests that show its effectiveness, including teleoperation latencies of 0.27 seconds on average, and error in the pose tracking of 7mm and 6°.

A range of additional research is possible for the human teleoperation concept, including applying it to other domains to see its efficacy there, and exploring the generalization of aspects of robotic control theory to human teleoperation. This includes for example studying stable and transparent force reflection in bilateral teleoperation under time delays imposed by the communication system and human response time.

VI. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support and guidance of the Engineering Physics Project Lab, University of British Columbia, as well as the valuable feedback of expert sonographers Vickie Lessoway and Jan Reid. We would like to thank Dr. Peter Black MD as well for performing the many expert ultrasounds during testing. Professor Salcudean gratefully acknowledges infrastructure support from CFI and funding support from NSERC and the Charles Laszlo Chair in Biomedical Engineering.

REFERENCES

- [1] L. S. Dalenogare, G. B. Benitez, N. F. Ayala, and A. G. Frank, "The expected contribution of industry 4.0 technologies for industrial performance," *International Journal of Production Economics*, vol. 204, pp. 383–394, 2018.
- [2] C. Thuemmler and C. Bai, *Health 4.0: How virtualization and big data are revolutionizing healthcare*. Springer, 2017.
- [3] A. G. Frank, L. S. Dalenogare, and N. F. Ayala, "Industry 4.0 technologies: Implementation patterns in manufacturing companies," *International Journal of Production Economics*, vol. 210, pp. 15–26, 2019.
- [4] A. C. Ferreira, E. O'Mahony, A. H. Olliani, E. Araujo Júnior, and F. da Silva Costa, "Teleultrasound: historical perspective and clinical application," *International journal of telemedicine and applications*, vol. 2015, 2015.
- [5] S. Wu, K. Li, R. Ye, Y. Lu, J. Xu, L. Xiong, A. Cui, Y. Li, C. Peng, and F. Lv, "Robot-assisted teleultrasound assessment of cardiopulmonary function on a patient with confirmed covid-19 in a cabin hospital," *Advanced Ultrasound in Diagnosis and Therapy*, vol. 4, no. 2, pp. 128–130, 2020.
- [6] K. S. Boniface, H. Shokoohi, E. R. Smith, and K. Scantlebury, "Teleultrasound and paramedics: real-time remote physician guidance of the focused assessment with sonography for trauma examination," *The American journal of emergency medicine*, vol. 29, no. 5, pp. 477–481, 2011.
- [7] A. Strumia, F. Costa, G. Pascarella, R. Del Buono, and F. E. Agrò, "U smart: ultrasound in your pocket," *Journal of clinical monitoring and computing*, pp. 1–3, 2020.

- [8] C. Delgorge, F. Courrèges, L. A. Bassit, C. Novales, C. Rosenberger, N. Smith-Guerin, C. Brù, R. Gilibert, M. Vannoni, G. Poisson *et al.*, “A tele-operated mobile ultrasound scanner using a light-weight robot,” *IEEE transactions on information technology in biomedicine*, vol. 9, no. 1, pp. 50–58, 2005.
- [9] G. Kontaxakis, S. Walter, and G. Sakas, “Eu-telein vivo: an integrated portable telemedicine workstation featuring acquisition, processing and transmission over low-bandwidth lines of 3d ultrasound volume images,” in *Proceedings 2000 IEEE EMBS International Conference on Information Technology Applications in Biomedicine. ITAB-ITIS 2000. Joint Meeting Third IEEE EMBS International Conference on Information Technol.* IEEE, 2000, pp. 158–163.
- [10] K. Mathiassen, J. E. Fjellin, K. Glette, P. K. Hol, and O. J. Elle, “An ultrasound robotic system using the commercial robot ur5,” *Frontiers in Robotics and AI*, vol. 3, p. 1, 2016.
- [11] V. G. Adriana, J. Troccaz, P. Cinquin, F. Courreges, G. Poisson, and B. Tondou, “Robotic tele-ultrasound system (ter): slave robot control,” *IFAC Proceedings Volumes*, vol. 34, no. 9, pp. 425–430, 2001.
- [12] P. Abolmaesumi, S. E. Salcudean, W.-H. Zhu, M. R. Sirouspour, and S. P. DiMaio, “Image-guided control of a robot for medical ultrasound,” *IEEE Transactions on Robotics and Automation*, vol. 18, no. 1, pp. 11–23, 2002.
- [13] F. Courreges, P. Vieyres, and R. Istepanian, “Advances in robotic tele-echography services—the otelo system,” in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2. IEEE, 2004, pp. 5371–5374.
- [14] P. Vieyres, G. Poisson, F. Courrèges, N. Smith-Guerin, C. Novales, and P. Arbeille, “A tele-operated robotic system for mobile tele-echography: The otelo project,” in *M-health*. Springer, 2006, pp. 461–473.
- [15] F. Courreges, P. Vieyres, R. S. Istepanian, P. Arbeille, and C. Bru, “Clinical trials and evaluation of a mobile, robotic tele-ultrasound system,” *Journal of telemedicine and telecare*, vol. 11, no. 1_suppl, pp. 46–49, 2005.
- [16] S. Wu, D. Wu, R. Ye, K. Li, Y. Lu, J. Xu, L. Xiong, Y. Zhao, A. Cui, Y. Li *et al.*, “Pilot study of robot-assisted teleultrasound based on 5g network: a new feasible strategy for early imaging assessment during covid-19 pandemic,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 11, pp. 2241–2248, 2020.
- [17] S. E. Salcudean, G. Bell, S. Bachmann, W.-H. Zhu, P. Abolmaesumi, and P. D. Lawrence, “Robot-assisted diagnostic ultrasound—design and feasibility experiments,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 1999, pp. 1062–1071.
- [18] P. A. Lasota, T. Fong, J. A. Shah *et al.*, *A survey of methods for safe human-robot interaction*. Now Publishers, 2017.
- [19] M. Victorova, D. Navarro-Alarcon, and Y.-P. Zheng, “3d ultrasound imaging of scoliosis with force-sensitive robotic scanning,” in *2019 Third IEEE International Conference on Robotic Computing (IRC)*, 2019, pp. 262–265.
- [20] I. Aliaga, A. Rubio, and E. Sanchez, “Experimental quantitative comparison of different control architectures for master-slave teleoperation,” *IEEE Transactions on Control Systems Technology*, vol. 12, no. 1, pp. 2–11, 2004.
- [21] B. J. Park, S. J. Hunt, G. J. Nadolski, and T. P. Gade, “Augmented reality improves procedural efficiency and reduces radiation dose for ct-guided lesion targeting: a phantom study using hololens 2,” *Scientific Reports*, vol. 10, no. 1, pp. 1–8, 2020.
- [22] P. Bettati, M. Chalian, J. Huang, J. D. Dormer, M. Shahedi, and B. Fei, “Augmented reality-assisted biopsy of soft tissue lesions,” in *Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 11315. International Society for Optics and Photonics, 2020, p. 113150W.
- [23] D. Ni, W. Y. Chan, J. Qin, Y.-P. Chui, I. Qu, S. S. Ho, and P.-A. Heng, “A virtual reality simulator for ultrasound-guided biopsy training,” *IEEE computer graphics and applications*, vol. 31, no. 2, pp. 36–48, 2009.
- [24] D. Escobar-Castillejos, J. Noguez, L. Neri, A. Magana, and B. Benes, “A review of simulators with haptic devices for medical training,” *Journal of medical systems*, vol. 40, no. 4, p. 104, 2016.
- [25] S. Wang, M. Parsons, J. Stone-McLean, P. Rogers, S. Boyd, K. Hoover, O. Meruvia-Pastor, M. Gong, and A. Smith, “Augmented reality as a telemedicine platform for remote procedural training,” *Sensors*, vol. 17, no. 10, p. 2294, 2017.
- [26] A. Rothberg, M. De Jonge, J. Jia, D. Nouri, J. M. Rothberg, and M. Sofka, “Augmented reality interface for assisting a user to operate an ultrasound device,” U.S. Patent 10702242, December 21, 2017.
- [27] —, “Methods and apparatus for tele-medicine,” U.S. Patent 20190261957.
- [28] S. P. Dalvin and M. S. Alkaitis, “Augmented/mixed reality system and method for guidance of a medical exam,” U.S. Patent 20190239850, August 8, 2019.
- [29] W. R. Buras, C. S. Russel, and K. Q. Nguyen, “System and method for three-dimensional augmented reality guidance for use of medical equipment,” U.S. Patent 10636323, April 28, 2020.
- [30] R. Masoni, F. Ferrise, M. Bordegoni, M. Gattullo, A. E. Uva, M. Fiorentino, E. Carrabba, and M. Di Donato, “Supporting remote maintenance in industry 4.0 through augmented reality,” *Procedia manufacturing*, vol. 11, pp. 1296–1302, 2017.
- [31] D. Mourtzis, V. Zogopoulos, and E. Vlachou, “Augmented reality application to support remote maintenance as a service in the robotics industry,” *Procedia Cirp*, vol. 63, pp. 46–51, 2017.
- [32] E. P. Westebring-van der Putten, R. H. Goossens, J. J. Jakimowicz, and J. Dankelman, “Haptics in minimally invasive surgery—a review,” *Minimally Invasive Therapy & Allied Technologies*, vol. 17, no. 1, pp. 3–16, 2008.
- [33] F. Chan, J. Whitehall, L. Hayes, A. Taylor, B. Soong, K. Lessing, R. Cincotta, D. Cooper, M. Stone, A. Lee-Tannock *et al.*, “Minimum requirements for remote realtime fetal tele-ultrasound consultation,” *Journal of Telemedicine and Telecare*, vol. 5, no. 3, pp. 171–176, 1999.
- [34] C. Crick, G. Jay, S. Osentoski, B. Pitzer, and O. C. Jenkins, “Rosbridge: Ros for non-ros users,” in *Robotics Research*. Springer, 2017, pp. 493–504.
- [35] S. Allin, Y. Matsuoka, and R. Klatzky, “Measuring just noticeable differences for haptic force feedback: implications for rehabilitation,” in *Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002*, 2002, pp. 299–302.
- [36] N. Smith-Guerin, L. Al Bassit, G. Poisson, C. Delgorge, P. Arbeille, and P. Vieyres, “Clinical validation of a mobile patient-expert tele-echography system using isdn lines,” in *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine, 2003*. IEEE, 2003, pp. 23–26.
- [37] C. Riviere and P. Khosla, “Accuracy in positioning of handheld instruments,” in *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, 1996, pp. 212–213 vol.1.
- [38] M. W. Gilbertson and B. W. Anthony, “Impedance-controlled ultrasound probe,” vol. 7968, p. 796816, 2011.
- [39] P. Arcara and C. Melchiorri, “Control schemes for teleoperation with time delay: A comparative study,” *Robotics and Autonomous systems*, vol. 38, no. 1, pp. 49–64, 2002.
- [40] G. Niemeyer and J.-J. Slotine, “Stable adaptive teleoperation,” *IEEE Journal of oceanic engineering*, vol. 16, no. 1, pp. 152–162, 1991.
- [41] C.-C. Hua and X. P. Liu, “Delay-dependent stability criteria of teleoperation systems with asymmetric time-varying delays,” *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 925–932, 2010.
- [42] T.-Y. Fang, H. K. Zhang, R. Finocchi, R. H. Taylor, and E. M. Boctor, “Force-assisted ultrasound imaging system through dual force sensing and admittance robot control,” *International journal of computer assisted radiology and surgery*, vol. 12, no. 6, pp. 983–991, 2017.
- [43] J. Carriere, J. Fong, T. Meyer, R. Sloboda, S. Husain, N. Usmani, and M. Tavakoli, “An admittance-controlled robotic assistant for semi-autonomous breast ultrasound scanning,” in *2019 international symposium on medical robotics (ISMR)*. IEEE, 2019, pp. 1–7.
- [44] D. De Cunha, P. Gravez, C. Leroy, E. Maillard, J. Jouan, P. Varley, M. Jones, M. Halliwell, D. Hawkes, P. Wells *et al.*, “The midstep system for ultrasound guided remote telesurgery,” in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)*, vol. 3. IEEE, 1998, pp. 1266–1269.
- [45] E. Degoulange, L. Urbain, P. Caron, S. Boudet, J. Gariépy, J.-L. Megnien, F. Pierrot, and E. Dombre, “Hippocrate: an intrinsically safe robot for medical applications,” in *Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovations in Theory, Practice and Applications (Cat. No. 98CH36190)*, vol. 2. IEEE, 1998, pp. 959–964.
- [46] A. AbbasiMoshaii and F. Najafi, “Design, evaluation and prototyping of a new robotic mechanism for ultrasound imaging,” *Journal of Computational Applied Mechanics*, vol. 50, no. 1, pp. 108–117, 2019.
- [47] F. Najafi and N. Sepehri, “A robotic wrist for remote ultrasound imaging,” *Mechanism and machine theory*, vol. 46, no. 8, pp. 1153–1170, 2011.
- [48] K. Hashtrudi-Zaad and S. E. Salcudean, “Analysis of control architectures for teleoperation systems with impedance/admittance master and

- slave manipulators,” *The International Journal of Robotics Research*, vol. 20, no. 6, pp. 419–445, 2001.
- [49] D. Ungureanu, F. Bogo, S. Galliani, P. Sama, X. Duan, C. Meekhof, J. Stuehmer, T. J. Cashman, B. Tekin, J. L. Schoenberger, B. Tekin, P. Olszta, and M. Pollefeys, “Hololens 2 research mode as a tool for computer vision research,” *arXiv:2008.11239*, 2020.
- [50] D. Thivierge-Gaulin, C.-R. Chou, A. P. Kiraly, C. Chefd’Hotel, N. Strobel, and F. Cheriet, “3d-2d registration based on mesh-derived image bisection,” in *International Workshop on Biomedical Image Registration*. Springer, 2012, pp. 70–78.
- [51] T. Hayashi, F. De Sorbier, and H. Saito, “Texture overlay onto non-rigid surface using commodity depth camera,” in *VISAPP (2)*, 2012, pp. 66–71.
- [52] D. A. Lawrence, “Stability and transparency in bilateral teleoperation,” *IEEE transactions on robotics and automation*, vol. 9, no. 5, pp. 624–637, 1993.
- [53] M. W. Gilbertson and B. W. Anthony, “An ergonomic, instrumented ultrasound probe for 6-axis force/torque measurement,” pp. 140–143, 2013.
- [54] P. Abolmaesumi, S. Salcudean, W. Zhu, S. DiMaio, and M. Sirospour, “A user interface for robot-assisted diagnostic ultrasound,” in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164)*, vol. 2, 2001, pp. 1549–1554 vol.2.
- [55] A. I. Aviles, A. Marban, P. Sobrevilla, J. Fernandez, and A. Casals, “A recurrent neural network approach for 3d vision-based force estimation,” in *2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2014, pp. 1–6.
- [56] S. Giannarou, M. Ye, G. Gras, K. Leibrandt, H. J. Marcus, and G.-Z. Yang, “Vision-based deformation recovery for intraoperative force estimation of tool–tissue interaction for neurosurgery,” *International journal of computer assisted radiology and surgery*, vol. 11, no. 6, pp. 929–936, 2016.
- [57] K. Hashtrudi-Zaad and S. E. Salcudean, “Bilateral parallel force/position teleoperation control,” *Journal of Robotic Systems*, vol. 19, no. 4, pp. 155–167, 2002.
- [58] R. J. Anderson and M. W. Spong, “Bilateral control of teleoperators with time delay,” vol. 1, pp. 131–138, 1988.
- [59] G. Niemeyer and J.-J. E. Slotine, “Telemanipulation with time delays,” *The International Journal of Robotics Research*, vol. 23, no. 9, pp. 873–890, 2004.
- [60] A. Aziminejad, M. Tavakoli, R. V. Patel, and M. Moallem, “Transparent time-delayed bilateral teleoperation using wave variables,” *IEEE Transactions on Control Systems Technology*, vol. 16, no. 3, pp. 548–555, 2008.
- [61] G. M. Leung, B. A. Francis, and J. Apkarian, “Bilateral controller for teleoperators with time delay via mu-synthesis,” *IEEE Transactions on Robotics and Automation*, vol. 11, no. 1, pp. 105–116, 1995.
- [62] I. G. Polushin, P. X. Liu, and C.-H. Lung, “A force-reflection algorithm for improved transparency in bilateral teleoperation with communication delay,” *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 3, pp. 361–374, 2007.
- [63] P. Mirk, N. Magnavita, L. Masini, M. Bazzocchi, and A. Fileni, “Frequency of musculoskeletal symptoms in diagnostic medical sonographers. results of a pilot survey,” *La Radiologia Medica*, vol. 98, no. 4, pp. 236–241, 1999.