# A Mixed Reality System for Human Teleoperation in Tele-Ultrasound

D. G. Black[1] and S. E. Salcudean[1]

[1]*Robotics and Control Lab, University of British Columbia*
*dgblack@ece.ubc.ca*

## INTRODUCTION

Many applications including telemedicine, manufacturing, and maintenance profit from remote guidance. Existing approaches to tele-ultrasound (US) include robotic teleoperation as well as multimedia applications that combine verbal and graphical guidance. Robotic US systems can provide high precision, low latency, and haptic feedback [1][2][3]. One system has demonstrated clinical utility in trials [4], and much recent work has focused on autonomous robotic US [5]. However, the issues of safe human-robot interaction and guaranteed robust autonomy remain difficult, especially from a regulatory perspective. Further limitations include restricted workspaces, time consuming set-up, large physical size that prevents use in ambulances, and cost, especially compared to inexpensive US systems.

Conversely, systems sold by Clarius Mobile Health Corp. and Butterfly Network use a wireless US probe with images and video conferencing available via a cloud interface on a mobile phone application. Though inexpensive and flexible, the desired probe pose and force are given verbally or with some overlays of arrows or pointers on the US image, which is very inefficient, leading to high latency and low precision.

We present a novel concept of "Human Teleoperation" through mixed reality which bridges the gap between these two methods. In this control framework, the human follower is controlled as a flexible, cognitive robot such that both 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, intuitive, and low latency than verbal guidance, yet more flexible, inexpensive, and accessible than robotic teleoperation. This short paper summarizes the concepts and introduces a new design of the communication system to use a secure, high-speed, network-agnostic WebRTC interface. More details and results can be found in [6].

## MATERIALS AND METHODS

The tele-US system consists of the follower side and the expert side, which communicate wirelessly over the Internet. The follower wears a mixed reality (MR) headset (HoloLens 2 in our implementation) which projects a virtual US transducer into the follower's scene. The expert controls the virtual probe using a haptic controller (Phantom Desktop, 3D Systems, Inc.) to input the desired pose and force. The follower tracks the virtual probe with his/her real probe, as seen in Fig. 1. Thus the expert, in real time, receives the US images, a video stream of the patient with the virtual and real probes in position (called an MR capture), and is in verbal communication with the follower. Additionally, the follower can send a spatial mesh of the patient, generated by the HoloLens 2, to the expert. The mesh is rendered haptically as a virtual fixture for the Phantom Desktop, giving the expert the sensation that they are physically interacting with the tissue. This spatial mapping also provides the expert-to-follower coordinate transform. The system is shown in Fig. 2.

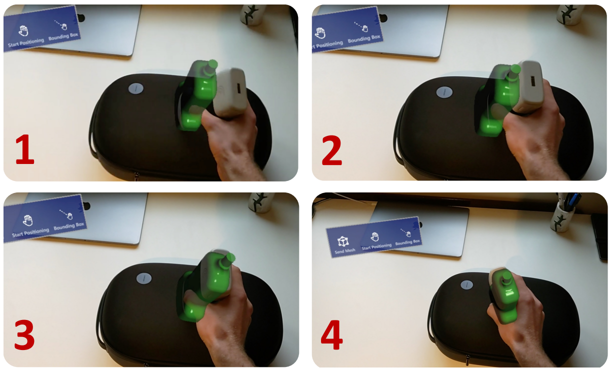


**Fig. 1** The teleoperation concept. Frames 1-3: follower matches the virtual probe precisely, starting in a random pose. Frame 4: the expert moves the virtual probe, and the real one quickly tracks the motion.

The effectiveness of the approach was demonstrated first using a WebSocket server and Robot Operating System (ROS) on a local wireless network (WLAN). Refer to [6] for details. We measured the data latencies, the actual end-to-end teleoperation latency, and the positional precision. Additionally, two procedures were carried out on two patients each to compare Human Teleoperation to verbal guidance and to direct US by the expert. These procedures involved quantitative endpoints so the numerical measurements (inferior vena cava diameter and kidney length) could be compared between methods. However, the US images, video feed, and spatial meshes require a large bandwidth while haptic feedback and MR teleoperation necessitate very low latencies for stable and transparent teleoperation. A Web Real Time Communication (RTC)-based system is more suitable to meet these requirements and support tele-US at large distances. This framework provides a direct peer-to-

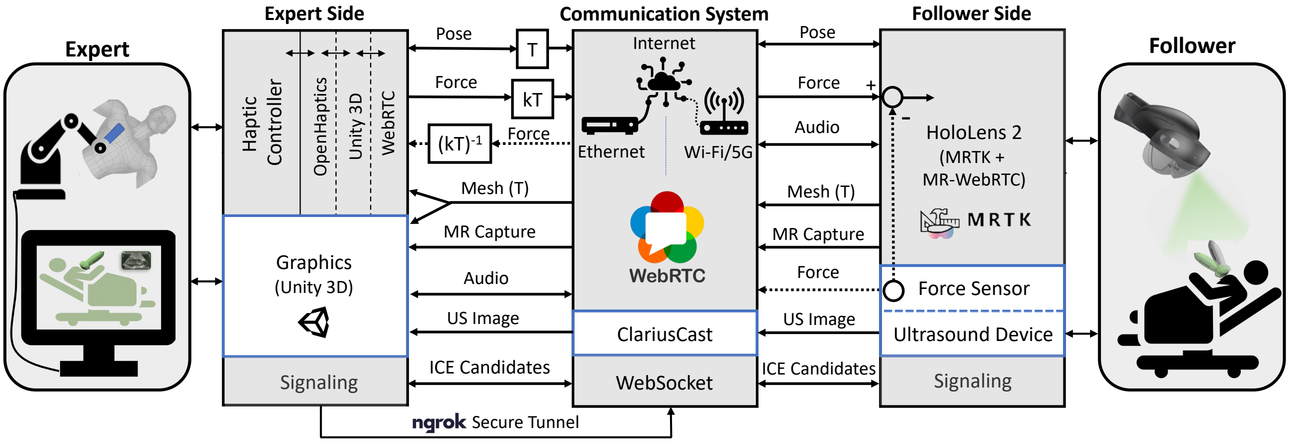


**Fig. 2** System Architecture. $k$ is a scaling factor for the force while $T$ is the transform from expert to follower coordinates, obtained from the mesh. The force feedback (dotted lines) has not yet been implemented.

peer connection over UDP between the expert and follower, thus removing server-related delays. Any dropped packets are quickly replaced with new information, and local consistency checks are in place, so the higher speeds of UDP are preferable to the reliability of TCP. An implementation of this system is in place, and in collaboration with Rogers Communications, a Canadian telecommunications company, the communication is set up to run over the 5G radio access network (RAN). The 5G network holds promise for achieving the required bandwidth and latency, and this will be tested using the sub-6Ghz and mm-wave bands. Latency results from initial tests with WebRTC are presented for comparison to the WebSocket-based communication.

## RESULTS

Our latency tests on the WebSocket implementation showed on average 11.4 ms delay for pose and force transmission. In contrast, the preliminary WebRTC implementation has delays of only 5.4 ms on average for the same local network and it can run over the Internet. In addition, the latency of the MR capture over WebRTC is 160ms whereas with Windows Device Portal it was found to be $\geq 4$ seconds, and with ROS it was infeasibly slow. The mean end-to-end teleoperation latency was measured through two trials to be 270 ms, with mean error in pose tracking of 7 mm and 6°.

While no patient tests have yet been carried out with the WebRTC implementation, our WebSocket results show average time taken to complete the procedures to be 65 seconds for direct US, 214 seconds for verbal guidance, and 73 seconds for Human Teleoperation. The average measurement error compared to direct US was 6 mm for verbal guidance and 3 mm for Human Teleoperation.

## DISCUSSION

The teleoperation latency is greater than that of any of the data channels individually because it is limited by the response time of the human follower. However, the latency of 270 ms is permissible even for haptic feedback without greatly degrading the user experience. Additionally, the new WebRTC implementation is shown to be much faster and more appropriate for teleultrasound at large distances. The effectiveness of the concept is shown most clearly in the patient tests, where it greatly outperformed verbal teleguidance in efficiency and precision. These tests were preliminary and had a small sample size, but nonetheless revealed a clear difference between methods. No quantitative comparison to robotic teleoperation has yet been made, but the benefits of this system compared to robotics lie in cost, ease of deployment, acceptance in communities, and flexibility.

Ongoing and future work includes utilizing 5G to perform remote tests, exploring how to improve the human-computer interaction, and investigating reinforcement learning for autonomous US guidance. We also plan on integrating force sensing on an US probe to allow the study of stable and transparent force reflection in bilateral teleoperation under time delays imposed by the human response time.

## REFERENCES


[1] 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.

[2] M. Akbari, J. Carriere, T. Meyer, R. Sloboda, S. Husain, N. Usmani, and M. Tavakoli, "Robotic ultrasound scanning with real-time image-based force adjustment: Quick response for enabling physical distancing during the covid-19 pandemic," *Frontiers in Robotics and AI*, vol. 8, 2021.

[3] C. Delgorge, F. Courrèges, L. A. Bassit, C. Novales, C. Rosenberger, N. Smith-Guerin, C. Brù, R. Gilabert, 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.

[4] P. Vieyres, G. Poisson, and et al., "A tele-operated robotic system for mobile tele-echography: The otelo project," in *M-health*. Springer, 2006, pp. 461–473.

[5] S. Wang, J. Housden, Y. Noh, D. Singh, and A. Singh, "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, pp. 27–38.

[6] D. Black, Y. O. Yazdi, A. H. H. Hosseinabadi, and S. Salcudean, "Human teleoperation - a haptically enabled mixed reality system for teleultrasound," *TechRxiv*, /08/19 2021.