Mixed Reality Tele-ultrasound over 750 km: a Clinical Study

Category: Research

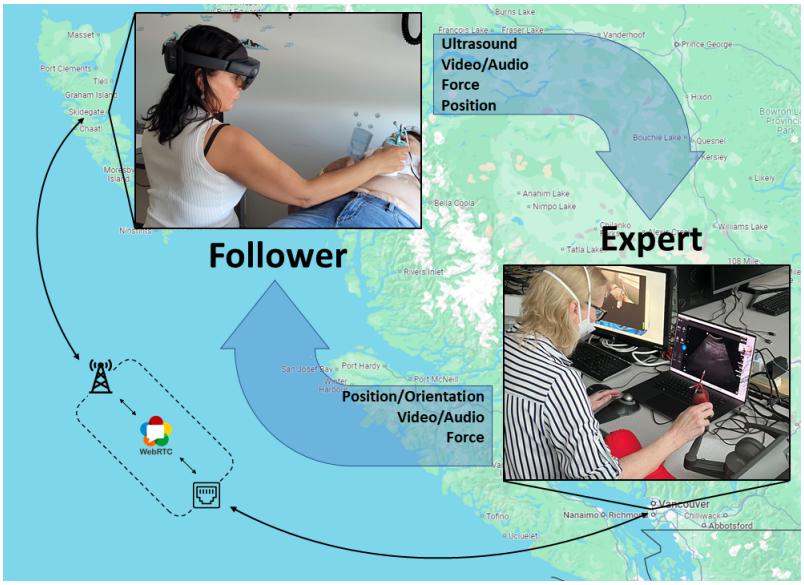


Figure 1: High-level system diagram illustrating the mixed reality teleultrasound system being used over a large distance.

Abstract

Ultrasound is a hand-held, low-cost, non-invasive medical imaging modality which plays a vital role in diagnosing various diseases. Despite this, many rural and remote communities do not have access due to the lack of local experts trained to perform ultrasound scanning. To address this challenge, we built a mixed reality and haptics-based tele-ultrasound system to enable an expert to precisely guide a novice remotely in carrying out an ultrasound exam. The precision and flexibility of our solution makes it more practical than existing tele-ultrasound solutions. We tested the system in Skidegate on the islands of Haida Gwaii, BC, Canada, with the experts positioned 754 km away at the University of British Columbia, Vancouver, Canada. We performed 11 scans with 10 novices and 2 experts. The experts were tasked with acquiring 5 target images and measurements in the epigastric region. The novices of various backgrounds and ages were all inexperienced in mixed reality and were not required to have prior ultrasound experience. The captured images were evaluated by two radiologists who were not present for the tests. These results are discussed along with new insights into the human computer interaction in such a system. We show that human teleoperation is feasible and can achieve high performance for completing remote ultrasound procedures, even at a large distance and with completely novice followers.

Index Terms: Mixed reality, teleultrasound, ultrasound, teleoperation, human computer interaction

1 Introduction

Ultrasound (US) is a medical imaging modality widely used for the diagnosis of various diseases by allowing visualization of soft tissue structures under the skin. It offers many benefits due to its non-invasiveness, low-cost, and portability [20]. Unlike other medical imaging modalities, US is hand-held, making the acquired image quality highly dependent on the operator. As such, a trained sonographer is required to perform the imaging. This requirement leads to a lack of access to US imaging in remote communities. A qualitative study highlighted geographical isolation as a central barrier to US imaging for remote, Indigenous communities in Canada. Individuals in these communities must either wait for an itinerant sonographer who comes once per month, or travel up to 1040 km to a larger city for access to US imaging [1]. For instance, individuals living in Skidegate, a small community on the islands of Haida Gwaii, BC, Canada, must take an 8 hour ferry to Prince Rupert where the closest major hospital is located to obtain a US exam. The resulting delay in diagnosis can lead to increased health complication rates and the travel can be a significant burden.

Tele-ultrasound (tele-US) aims to solve this problem by enabling an expert to remotely guide or perform US imaging. Tele-mentored US is a form of tele-US which uses audiovisual video conferencing to enable a sonographer to guide a novice in real-time to perform US imaging [9]. Robotic US is another form of tele-US in which an expert teleoperates a robotic arm to acquire the US images. A recent review of robotic US imaging can be found in [19]. Robotic US gives sonographers precise and responsive control and has been demonstrated in multiple clinical studies [11, 22, 14, 2]. Yet, there has been minimal commercial success which can likely be attributed to the high cost, large footprint, and complex maintenance, set-up and operation of a telerobotic system. Recently, there has been increased interest and advancements towards autonomous robotic US [29, 26, 27], but concerns regarding safe human-robot interaction limit its feasibility.

Tele-mentored US is commonly used together with portable point of care US (POCUS) devices such as the Butterfly iQ3 (Butterfly Network, Burlington, MA) or Clarius C3HD3 (Clarius Mobile Health, Vancouver, BC). This makes the solution cheaper and easier to set up compared to robotic US. Despite this, these systems are often used for mentoring novice sonographers or clinicians with some US experience rather than a completely inexperienced novice. This is likely due to the lack of precision and the high latency of the interaction. Additionally, there are few studies of good methodological quality that demonstrate the efficacy of tele-mentored US [15].

To address the need for a low-cost, flexible and precise tele-US solution, the concept of mixed reality (MR) "Human Teleoperation" was introduced in [5] and expanded on in [4]. Within healthcare, MR has been applied in surgical training and planning to enable 3D visualization of anatomy [28], and for enhancing navigation in ultrasound-guided needle biopsies [16]. In human teleoperation, rather than rendering static overlay images or pointers, MR is used to render a dynamic virtual guiding US transducer controlled in real time by a remote expert. By having the local novice (referred to as the "follower") follow the virtual guiding transducer, the remote expert can control them like a flexible, cognitive robot. This tightly coupled hand-over-hand control is more precise than tele-mentored US while being more accessible and portable than robotic US.

In [6], human participants ($n = 11$) were evaluated for their ability to track pose and force through MR cues in the human teleoperation system. The results demonstrated the feasibility of using this concept for accurate and quick teleoperation. Additionally, the system latency was evaluated in [8] which characterized delay from both the communication system and human response time. The system achieved communication round-trip delays of 40 ± 10 ms over 5G and 5.8 ± 3.3 ms over WiFi with typical throughputs. These results demonstrated the system was network agnostic and could achieve the rates required for teleoperation.

The feasibility of using human teleoperation specifically for tele-US has not yet been evaluated. US examinations typically require small but precise motions which were not replicated in [6]. Additionally, communication latency caused by large distances was not captured in [8]. As such, we evaluated the feasibility of using mixed reality human teleoperation for tele-US by performing remote US scans with sonographers and healthy participants separated by a large distance. Two expert sonographers positioned at the University of British Columbia, Vancouver, Canada performed 11 remote ultrasound scans with 10 different followers located 754 km away in Skidegate, Canada using the human teleoperation system. In this paper, we first summarize new developments made to the human teleoperation system, then describe the remote tele-US tests we performed. Finally, the results from follower and expert questionnaires and the quality of the US images are presented and discussed.

2 Methods

2.1 System Design

In this study, we focus on the use of human teleoperation for tele-US. The system details are provided in references [5, 8, 7, 3] and are summarized here and in Fig. 2 for completeness. For this application, a novice follower situated with the patient wears the Microsoft HoloLens 2 (Microsoft, Redmond, WA) mixed reality headset which displays a virtual guiding US transducer in their environment. The follower then superimposes the real US transducer on the virtual transducer and maintains alignment as the virtual transducer is moved. The pose of the virtual transducer is controlled by an expert sonographer located remotely (eg. at a medical center). The expert controls the virtual transducer using the Touch X haptic device (3D Systems, Rock Hill, SC) while seeing the real-time feed of the US image and point-of-view video from the follower. The haptic device is a desktop 6-degree-of-freedom (DOF) serial manipulator able to apply 3-DOF forces on the handle. As the sonographer moves the handle of the manipulator, its pose is measured and sent to the HoloLens 2 to be replicated by the virtual transducer. The haptic device also renders forces to the sonographer to mimic the sensation of a real US scan.

The communication between the follower side and expert side of the system was built using the Web Real Time Communication (WebRTC) peer-to-peer architecture. This is described in detail and evaluated in [8, 4]. The measured transducer force and pose, and the point-of-view video feed of the follower are sent to the expert

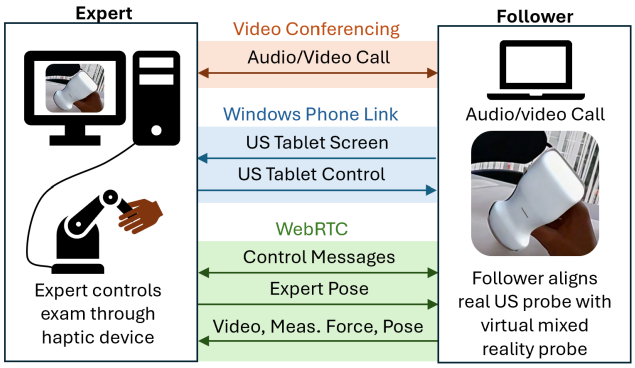


Figure 2: Diagram of human teleoperation system, showing the method and direction of communication.

while the expert's haptic device pose is transmitted to the follower, all over WebRTC. An additional WebRTC channel is used for exchanging control commands between the two sides of the system. The follower and expert communicate over an external video call. The US image feed is shared over the Windows Phone Link app to allow the expert to view the US image and control the image parameters in real-time.

The pose of the US transducer is measured by attaching a known arrangement of infrared (IR) reflective markers and using the IR camera and time-of-flight depth camera on the HoloLens 2. The specific algorithms used to determine the pose and the performance of this implementation are described in [7]. To measure the forces applied on the US transducer, we use differential magnetic force sensing which requires attaching a low-profile shell around the US transducer. This design is described and evaluated in [3].

In this study, we focus on assessing the feasibility of using human teleoperation for remote abdominal US scans. To produce accurate haptic feedback to the expert, we estimate the patient's abdominal surface with an ellipsoid and compute the forces based on the interaction between this ellipsoid and the virtual transducer. The points used to fit the ellipsoid are collected by pressing the US transducer on the patient's abdominal surface. Transducer forces are measured and when they exceed a threshold as a result of pressing into the patient, the tracked pose is used to determine the transducer's tip position. This provides a very intuitive way for the novice to collect points at the beginning of a US scan. Several assumptions allow us to minimize the number of points required to fit the ellipsoid. Firstly, we assume the ellipsoid axes are aligned with the patient's longitudinal, sagittal and frontal axes respectively. This assumption holds as long as the patient is lying in supine position which is the case for the abdominal scans we perform in this study. Second, we set the longitudinal semi-axis (which we denote c) of the ellipsoid to a very large value since the accuracy of this dimension is not necessary. Lastly, we assume the ellipsoid is tangent to the bed surface. With these assumptions we can determine the ellipsoid center (z_c, y_c, x_c) and remaining two semi-axes lengths (a and b) with the following four points:

1. Patient's xiphoid process just below the sternum (z_1, y_1, x_1)
2. Patient's extreme left (z_2, y_2, x_2)
3. Patient's extreme right (z_3, y_3, x_3)
4. Bed level (z_4, y_4, x_4)

These points are illustrated in Fig. 3 which shows a transverse view of the ellipsoid. The fourth point ("Bed level") can be collected anywhere on the surface of the bed as only the height component (y_4) is used. The parameters can then be computed efficiently

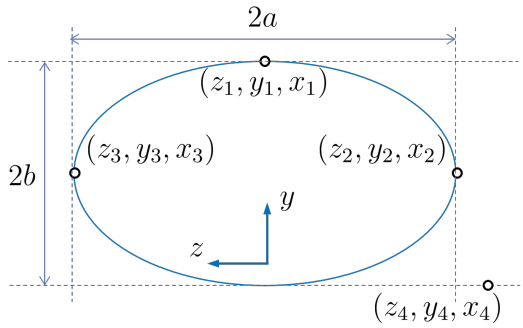


Figure 3: ZY (transverse) view of ellipsoid illustrating where points are collected to fit the ellipsoid.

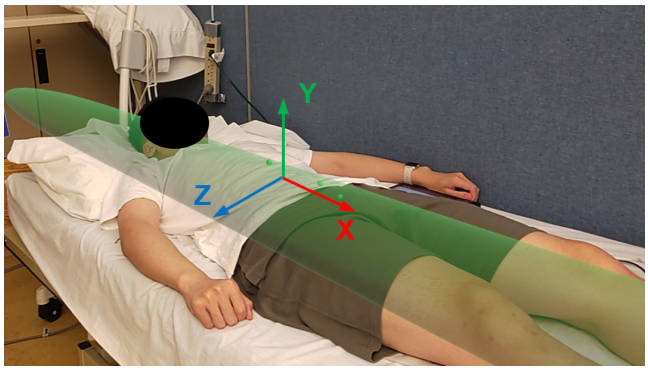


Figure 4: Fitted ellipsoid and coordinate system rendered over the patient's body.

by the set of equations

$$a = (z_3 - z_2)/2 \tag{1}$$
$$b = (y_1 - y_4)/2 \tag{2}$$
$$z_c = (z_3 + z_2)/2 \tag{3}$$
$$y_c = (y_1 + y_4)/2 \tag{4}$$
$$x_c = x_1 \tag{5}$$

We can now specify the ellipsoid by

$$\frac{(z-z_c)^2}{a^2} + \frac{(y-y_c)^2}{b^2} + \frac{(x-x_c)^2}{c^2} = 1 \tag{6}$$

The resulting ellipsoid, rendered visually, is shown in Fig. 4. Using Eq. (6) we determine when the expert's virtual transducer position ($\boldsymbol{x}_e = [x_e,\ y_e,\ z_e]^\top$) penetrates the ellipsoid and render a force proportional to the penetration distance with a spring and damping constant and normal to the ellipsoid surface. To detect penetration into the ellipsoid, consider an ellipsoid that is uniformly scaled down from the original one such that $\boldsymbol{x}_e$ lies on its surface. The surface normal vector at $\boldsymbol{x}_e$ is then

$$\boldsymbol{n} = \begin{bmatrix} \frac{(x_e - x_c)}{c^2} \\ \frac{(y_e - y_c)}{b^2} \\ \frac{(z_e - z_c)}{a^2} \end{bmatrix}$$
$$\hat{\boldsymbol{n}} = \boldsymbol{n}/\|\boldsymbol{n}\| \tag{7}$$

To find the penetration distance, d, we find where $\boldsymbol{x}_e + d\hat{\boldsymbol{n}}$ intersects the ellipsoid by solving the following quadratic:

$$(\boldsymbol{x}_e + d\hat{\boldsymbol{n}} - \boldsymbol{x}_c)^\top P^{-1} (\boldsymbol{x}_e + d\hat{\boldsymbol{n}} - \boldsymbol{x}_c) = 1 \tag{8}$$

Where P is a diagonal matrix with elements a^2, b^2, and c^2. The quadratic gives two solutions, the near and far-side intersections with the ellipse, so we choose the smaller of the two. If $d > 0$, the haptic device is intersecting the ellipsoid and we apply a force

$$\boldsymbol{f} = dK_p\hat{\boldsymbol{n}} - K_d\boldsymbol{v} \tag{9}$$

Where K_p, K_d are diagonal gain matrices and $\boldsymbol{v}$ is the velocity. Due to the large round-trip time delay between expert motion and measured force caused by the communication and follower lag, directly feeding back forces was found to be impractical. The ellipsoid method, however, gave the expert a stable and relatively accurate haptic representation of the patient on which to move and rest their hand while guiding the scan.

2.2 User Study

With this setup, we conducted a user study to assess the feasibility of using human teleoperation for remote US imaging. We performed 11 abdominal US scans on healthy volunteer patients. Participants ($n = 10$) of various ages (ranging from 20–60+ years) and backgrounds, and with no prior experience using the human teleoperation system, were recruited as the followers. Two certified expert sonographers took turns performing the scans by guiding the followers from the expert side. The two sonographers both have 30+ years of experience in US scanning. The followers and patients were situated in Skidegate, Haida Gwaii while the expert sonographers were situated 754 km away at the University of British Columbia, Vancouver, Canada. The expert side system was connected to the Internet via Ethernet while the follower side system was connected via WiFi. For each scan, the sonographers were tasked with acquiring the following five target images and measurements:

1. Proximal aorta with the anteroposterior (AP) diameter
2. Longitudinal view of the infrarenal segment of the inferior vena cava (IVC)
3. Long axis view of the left lobe of the liver
4. Transverse view of the left lobe of the liver
5. Transverse view of the right portal vein

As done in a regular US exam, the sonographers froze the image and made annotations as necessary for each of the targets before saving the images. Prior to starting any scans, the two sonographers were introduced to the system and each performed a trial scan at the University of British Columbia to gain familiarity using the system. Before each scan, the follower performing the scan was given instructions through a presentation with short video demonstrations.

After each scan, the follower completed a questionnaire which contained a NASA TLX survey [17] and a modified system usability scale. The sonographer also completed a questionnaire after each scan to assess their confidence in the images captured and interaction with the follower. Once all 11 scans were completed, both sonographers filled out a questionnaire assessing the usability of the system. The images obtained were also individually evaluated by two radiologists for their quality. Each image was scored on a six point scale based on the image generation assessment tool introduced and validated in [24, 23]. In these two studies, the assessment tools were designed specifically for cardiac and thoracic POCUS, but both used the same scale for image generation which is shown in Tab. 1. This scale demonstrated good inter-rater reliability and was able to detect change in learner performance. An additional binary scale for image interpretation was also employed, but its reliability was found to be lower than that of the image generation scale. Moreover, the image interpretation scale was used to assess

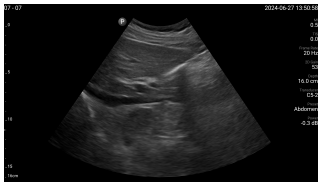
the ability to diagnose specific pathologies from the set of images obtained, rather than evaluating individual images. For this reason, we opted to use only the image generation scale in our study.

Table 1: Image quality scores and definitions [24, 23]. The radiologists provided a score for each individual image.

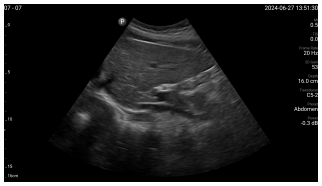
Score Definition	Score
Not obtained	0
Image quality is too poor to permit meaningful interpretation	1
(Not defined)	2
Suboptimal image quality, but basic image interpretation is possible	3
(Not defined)	4
Good image quality, meaningful image interpretation is easy	5

3 Results

(a) Proximal aorta with AP diameter.



(b) IVC - longitudinal.



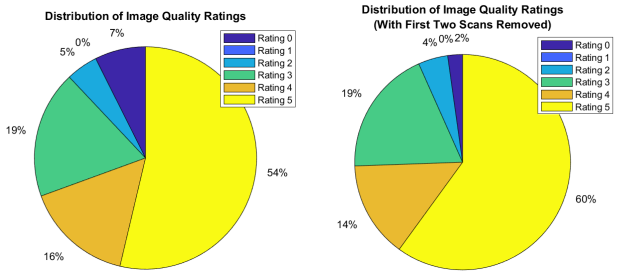
(c) Left lobe of liver - long axis.

(d) Left lobe of liver - transverse.

(e) Right portal vein - transverse.

Figure 5: Example of the target US images acquired by the sonographers using the human teleoperation system.

The sonographers completed 11 abdominal US scans, each with 5 target images and measurements for a total of 55 targets. An example image of each of the targets acquired during these tests is shown in Fig. 5. After completing the scans, two radiologists scored the images based on quality, including identifying targets that could not be seen or were not captured (which were given a score of 0). The distribution of these scores is illustrated in Fig. 6. The first radiologist identified 4 out of the 55 targets as not visible, while the second radiologist identified 6 out of 55 as not visible. Combined, this accounted for 7 unique targets that at least one radiologist considered missing. Of these missing targets, three were not captured due to large amounts of bowel gas and body habitus while one was seen but the sonographer did not capture and save the image. When excluding all the missing targets, the images obtained a mean score of 4.28 ± 0.95 out of 5 and 91.7% of the images were scored 3 or higher by both radiologists. A score of 3 or higher indicated the image quality was sufficient for basic image interpretation. 31.3% of the images were scored 5 by both radiologists, indicating the image quality was good and meaningful image interpretation was easy.



(a) All scans.

(b) Excluding first two scans.

Figure 6: Distribution of image quality scores. A score of 0 was given if the target was not captured or could not be seen and a score of 5 was given if the image quality was good and meaningful image interpretation was easy.

A summary of the NASA TLX results is shown in Fig. 7. Four extreme outlier scores on individual criteria were removed as they contradicted the scores given in other categories and were likely a result of human error during survey completion. Overall, the NASA TLX and system usability scale results showed low workload and high usability. The mental demand, effort, and frustration subscales had mean ratings of 23.3, 18.9 and 16.5 respectively out of 100, where a higher score indicated higher workload. The performance subscale had a mean rating of 37.0, where a lower score indicated better performance.

The follower usability questions and mean scores are shown in Tab. 2. The overall scores were all positive. All 10 followers agreed or strongly agreed the system was easy to use. All 10 followers also agreed or strongly agreed that most people would learn to use the system very quickly. In the usability questionnaire completed by the sonographers, both indicated the expert side of the system was easy to use and the control of the US transducer using the haptic device was intuitive. The mean completion time to acquire all target images and measurements was 645±116 seconds, excluding the first scan which was an outlier and took twice as long to complete due to technical issues.

Table 2: Mean scores from the modified system usability scale completed by the followers (0 = strongly disagree, 4 = strongly agree).

Usability Prompt/Question	Mean $\pm$ SD
I thought the overall system was easy to use	3.6 ± 0.52
I found it easy to follow the virtual transducer	3.6 ± 0.52
I found the point/dot fitting at the beginning easy to do	3.7 ± 0.48
I found it easy to follow the sonographer/expert's instructions	4.0 ± 0
I needed to learn a lot of things before I could get going with the system (lower is better)	1.4 ± 0.97
I would imagine that most people would learn to use this system very quickly	3.5 ± 0.53
I found the system very cumbersome to use (lower is better)	1.2 ± 1.23
I felt very confident using the system	3.4 ± 0.70

After completion of the scans, the followers were also asked to provide verbal feedback on their experiences. This provided additional insights into the human computer interaction aspects of this system not captured by the questionnaires. Many followers indicated it was easiest to follow the virtual transducer by looking at

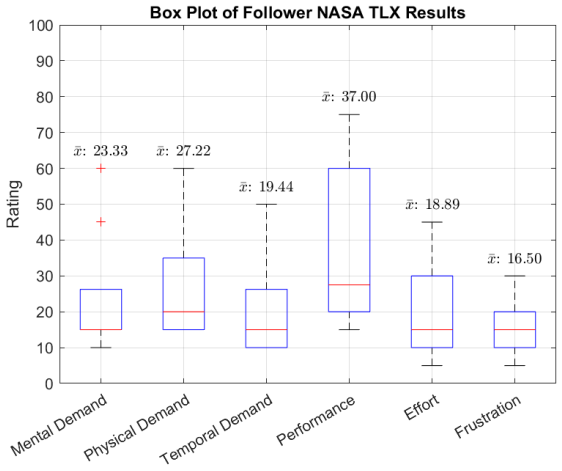


Figure 7: Boxplot and mean scores (out of 100, lower is better) from the NASA TLX survey completed by the followers ($n = 10$).

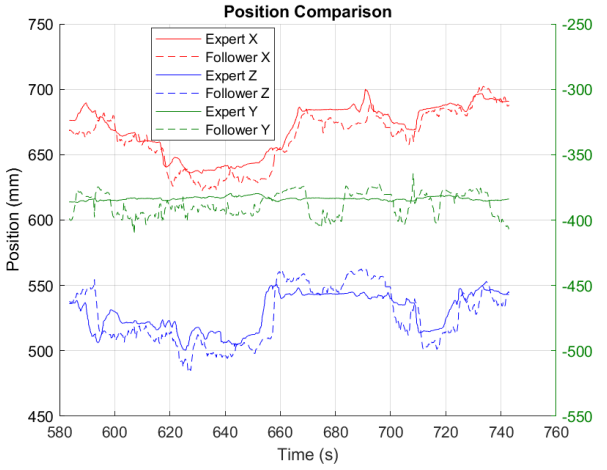


Figure 8: Example trajectory of follower's transducer (dotted lines) following the virtual transducer's trajectory controlled by the expert (solid lines).

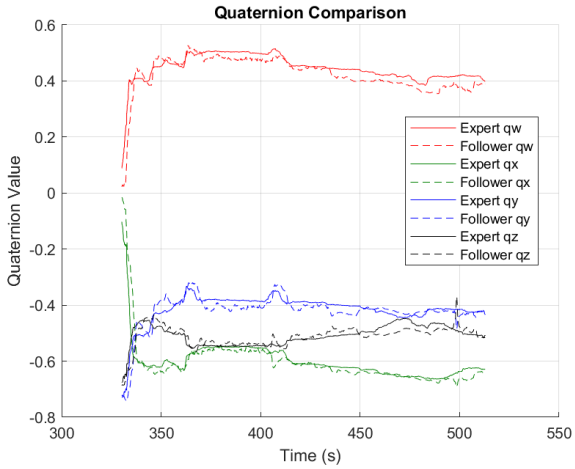


Figure 9: Example quaternion trajectory of follower's transducer (dotted lines) following the virtual transducer's quaternion trajectory controlled by the expert (solid lines).

Table 3: Position and orientation RMSE for each scan. The normalized RMSE was computed from the error values with the mean error subtracted.

Scan Number	Position RMSE (mm)	Normalized Position RMSE (mm)	Orientation RMSE (°)	Normalized Orientation RMSE (°)
1	31.9	14.2	18.3	5.70
2	41.3	6.99	11.5	3.82
3	29.8	8.88	8.04	3.73
4	30.8	8.09	9.21	3.72
5	44.3	18.5	15.2	4.19
6	32.7	11.3	11.5	10.9
7	22.7	9.13	6.46	2.63
8	51.8	11.6	15.1	13.2
9	19.3	5.14	8.89	5.16
10	22.7	8.90	11.5	6.94
11	24.0	7.84	9.59	3.96
Mean $\pm$ SD	31.9 ± 10.2	10.1 ± 3.7	11.4 ± 3.55	5.81 ± 3.33

the transducer body itself rather than the markers on top. It was also often mentioned that the geometry and size of the transducer's force sensor shell made it difficult to grasp and rotate the transducer comfortably.

The followers' tracking performances were evaluated by comparing the position and orientation trajectories of the real transducer to the expert controlled virtual transducer. An example of this is illustrated in Figs. 8 and 9 and the root mean squared error (RMSE) values of each scan are presented in Tab. 3. Overall, the followers achieved a mean position RMSE of 31.94 ± 10.17 mm and mean orientation RMSE of $11.39 \pm 3.55°$. After subtracting the mean error from each sample, the mean normalized position and orientation RMSE was 10.06 ± 3.74 mm and $5.81 \pm 3.33°$ respectively. No statistically significant correlation was observed between tracking performance (both position and orientation) and image quality.

4 Discussion

The results of this study demonstrated the feasibility of using human teleoperation for remote tele-US. The sonographers were able to perform scans and capture pre-specified target images and measurements, with 91.7% of the acquired images considered usable for interpretation by both radiologists evaluating them. Of the 7 unique missing targets (1 identified by the first radiologist, 3 by the second radiologist, and 3 by both), 5 occurred in the first two scans. The sonographer performing the second scan noted the patient was more difficult to scan due to overlying bowel gas and body habitus. The first two scans were also occasionally interrupted by technical issues such as system disconnections which were resolved in the later scans. These various factors likely contributed to the fact that the majority of the missing targets occurred in these first two scans. This is illustrated by comparing the distribution of image quality scores in the two plots in Fig. 6. Without the first two scans, there is a noticeable increase in the proportion of image quality scores greater than 2. Any image given a score greater than 2 were an indication that it was usable by the radiologists for interpretation.

Comparing the image quality scores to the corresponding RMSE, there was no observable or statistically significant correlation. This is illustrated in Fig. 10. Overall, the second scan received the lowest mean image quality score of 2.1 while the eighth scan received the

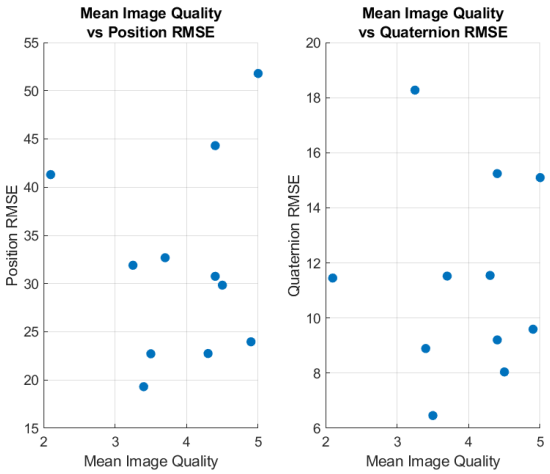


Figure 10: Comparison of mean image quality scores with position and quaternion RMSE.

highest mean image quality score of 5.0. The sonographers had noted the second patient was difficult to scan due to body habitus while the eighth patient was ideal and easy to scan. This suggests image quality was not affected by the performance of the follower but rather the patient's body habitus. This is something that would affect image quality even in a standard US exam. In future work, we will aim to obtain a better understanding of how tele-US with human teleoperation compares to standard US. This will be done by directly comparing the image quality and measurements obtained using human teleoperation with those obtained through standard US on the same patient with minimal temporal separation.

The NASA TLX ratings from our followers were compared to the reference values reported in [18], which were based on the mean scores from 556 studies. These reference values encompass various domains, including healthcare and virtual reality technologies. Our results consistently showed lower (i.e. better) scores across all categories compared to the healthcare reference values. Our results also surpassed those for virtual reality technologies in all categories except physical demand, where the difference was a negligible one point. Compared to other subscales, performance had a higher mean and greater variability. This aligned with the feedback provided by many of the followers who indicated being unsure of whether their alignment with the virtual transducer was accurate. This lack of confidence likely led to the lower self-reported performance scores. Similarly, the confidence question in the usability questionnaire had a slightly lower mean score and higher variance compared to other questions (particularly the questions related to ease of use).

The system usability results showed it was easy for complete novices from various background to use the system as a follower. All of the followers indicated the brief instructions presented at the beginning were sufficient for the task. All the followers also strongly agreed the instructions provided by the sonographer throughout the scans were easy to follow. In fact, many followers noted that the instructions helped them achieve better tracking of the virtual transducer. The sonographers often verbalized their next move (such as stating they will rotate the transducer) which helped the followers anticipate and track the motion more easily, though we did not measure this quantitatively. Several followers also expressed uncertainty about their alignment with the virtual transducer, which was reflected in the performance category of the NASA TLX results, showing relatively high variance as depicted in Fig. 7. Without visual or other feedback on their alignment, followers may make excessive micro-adjustments, resulting in a less stable image. Future work will focus on addressing this issue by exploring methods to provide alignment feedback, such as incorporating simple visual cues.

The tracking RMSE values decreased significantly after the mean error was subtracted from the errors, as shown in Tab. 3. This shows many of the followers likely had a constant offset in their tracking. A likely cause of this issue is the improper wearing of the HoloLens 2 headset, leading to a misalignment between the perceived and actual positions of the virtual transducer. However, this offset is implicitly accounted for by the sonographer, who relies on the ultrasound image to infer the true position of the transducer, and based on this makes relative motions which are tracked by the follower. As a result, the offset does not impair the expert's ability to perform the scans. This is evident in scan 8, where the follower initially had a large RMSE of 51.8 mm, which decreased to 11.6 mm after removing the constant offset. Despite the significant offset, the sonographer noted in the questionnaire that they did not notice any substantial tracking inaccuracy. Moreover, all images acquired during this scan received the highest image quality score from both radiologists. This suggests that having the follower achieve high absolute accuracy in tracking is less critical than ensuring precise and stable tracking. This is one benefit of human teleoperation over telerobotics, in which a precise calibration step is usually required to achieve absolute positioning.

Feedback from the expert sonographers was mostly related to the control of the US imaging parameters. The Windows Phone Link app was used to allow the sonographers to remotely interact with the Philips Lumify Ultrasound app, but the sonographers found it difficult to use the cinescrolling feature, place calipers in the image and make annotations. In future work, we will create a more intuitive dedicated screen sharing and control app to address these issues, as started in [4]. Apart from this, the sonographers also noted it would have helped to have direct feedback of the force applied by the follower. To address this, future work will explore other methods of force rendering such as using a mesh of the patient acquired with the HoloLens 2, which would be more accurate and flexible compared to estimating the patient using an ellipsoid. Another option would be combined position-force feedback using the tracked pose of the transducer and the force sensor attached to the transducer.

While this study demonstrated the feasibility of using human teleoperation for tele-US, additional scans in other abdominal regions would further validate the system's capabilities. Future work will involve evaluating human teleoperation for complete abdominal aorta examinations, where the sonographer will scan through the length of the aorta and capture measurements of its different segments. These quantitative measurements will be ideal for comparison with standard US. Further tests will explore the use of human teleoperation for renal ultrasound, which will require the patient to switch between supine and lateral decubitus positions. This adds complexity, as the scans in this study only required the patient to remain in a supine position. Additionally, this study focused more on understanding the experience of the followers rather than the sonographers. In future work, we will aim to develop a better understanding of the user experience from the experts' perspectives by recruiting a greater number of sonographers to use and evaluate the system.

The data collected through these tests will also present an opportunity to investigate integrating artificial intelligence for autonomous US guidance with mixed reality. The use of learning from demonstration and reinforcement learning for enabling autonomous US has previously been explored [10, 12, 21, 13, 25]. By leveraging the rich, multi-modal data obtained with the human teleoperation system, it may be possible to train an AI agent to replace the expert or augment certain sub-tasks, a solution that would

be especially valuable in areas with limited or no internet connectivity.

5 Conclusion

The study presented in this paper demonstrated the feasibility of using human teleoperation to perform remote abdominal US scans where the sonographer and patient are separated by a significant distance. Expert sonographers were able to use the system to acquire 48 out of the 55 target images and measurements across 11 scans. Of these acquired images, 92% were considered sufficient quality for basic image interpretation by both radiologists who evaluated the images. The 10 novice followers all expressed that the system was easy to use, even with the minimal instructions provided to them. We found no correlation between the followers' tracking errors and the image quality scores given by the radiologists. Ultimately, these results show the potential impact this solution can have for remote communities without local access to US imaging services.

References

[1] S. J. Adams, P. Babyn, B. Burbridge, R. Tang, and I. Mendez. Access to ultrasound imaging: A qualitative study in two northern, remote, indigenous communities in canada. *Int. J. Circumpolar Health*, 80(1):1961392, Dec. 2021. 1

[2] S. J. Adams, B. Burbridge, L. Chatterson, P. Babyn, and I. Mendez. A telerobotic ultrasound clinic model of ultrasound service delivery to improve access to imaging in rural and remote communities. *J. Am. Coll. Radiol.*, 19(1 Pt B):162–171, Jan. 2022. 1

[3] D. Black, A. H. H. Hosseinabadi, N. R. Pradnyawira, M. Nogami, and S. Salcudean. Low-profile 6-axis differential magnetic force/torque sensing. *IEEE Trans. Med. Robot. Bionics*, pp. 1–1, 2024. 2

[4] D. Black, M. Nogami, and S. Salcudean. Mixed reality human teleoperation with device-agnostic remote ultrasound: Communication and user interaction. *Comput. Graph.*, 118:184–193, Feb. 2024. 2, 6

[5] D. Black, Y. Oloumi Yazdi, A. H. Hadi Hosseinabadi, and S. Salcudean. Human teleoperation - a haptically enabled mixed reality system for teleultrasound. *Hum.-Comput. Interact.*, pp. 1–24, June 2023. 2

[6] D. Black and S. Salcudean. Human-as-a-robot performance in mixed reality teleultrasound. *Int. J. Comput. Assist. Radiol. Surg.*, 18(10):1811–1818, Oct. 2023. 2

[7] D. Black and S. Salcudean. Robust object pose tracking for augmented reality guidance and teleoperation. *IEEE Trans. Instrum. Meas.*, 73:1–15, 2024. 2

[8] D. G. Black, D. Andjelic, and S. E. Salcudean. Evaluation of communication and human response latency for (human) teleoperation. *IEEE Trans. Med. Robot. Bionics*, 6(1):53–63, Feb. 2024. 2

[9] N. Britton, M. A. Miller, S. Safadi, A. Siegel, A. R. Levine, and M. T. McCurdy. Tele-ultrasound in resource-limited settings: A systematic review. *Front. Public Health*, 7:244, Sept. 2019. 1

[10] M. Burke, K. Lu, D. Angelov, A. Straižys, C. Innes, K. Subr, and S. Ramamoorthy. Learning rewards from exploratory demonstrations using probabilistic temporal ranking. *Auton. Robots*, 47(6):733–751, Aug. 2023. 6

[11] C. Delgorge, F. Courrèges, L. Al Bassit, C. Novales, C. Rosenberger, N. Smith-Guerin, C. Brù, R. Gilabert, M. Vannoni, G. Poisson, and P. Vieyres. A tele-operated mobile ultrasound scanner using a lightweight robot. *IEEE Trans. Inf. Technol. Biomed.*, 9(1):50–58, Mar. 2005. 1

[12] X. Deng, Y. Chen, F. Chen, and M. Li. Learning robotic ultrasound scanning skills via human demonstrations and guided explorations. In *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, Dec. 2021. 6

[13] R. Droste, L. Drukker, A. T. Papageorghiou, and J. A. Noble. Automatic probe movement guidance for freehand obstetric ultrasound. *Med. Image Comput. Comput. Assist. Interv.*, 12263:583–592, Oct. 2020. 6

[14] B. Duan, L. Xiong, X. Guan, Y. Fu, and Y. Zhang. Tele-operated robotic ultrasound system for medical diagnosis. *Biomed. Signal Process. Control*, 70(102900):102900, Sept. 2021. 1

[15] M. L. Duarte, L. R. Dos Santos, W. Iared, and M. S. Peccin. Telementored ultrasonography: a narrative review. *Sao Paulo Med. J.*, 140(2):310–319, Mar. 2022. 1

[16] L. Groves, N. Li, T. M. Peters, and E. C. S. Chen. Towards a first-person perspective mixed reality guidance system for needle interventions. *J. Imaging*, 8(1):7, Jan. 2022. 2

[17] S. G. Hart and L. E. Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*, Advances in psychology, pp. 139–183. Elsevier, 1988. 3

[18] M. Hertzum. Reference values and subscale patterns for the task load index (TLX): a meta-analytic review. *Ergonomics*, 64(7):869–878, July 2021. 6

[19] Z. Jiang, S. E. Salcudean, and N. Navab. Robotic ultrasound imaging: State-of-the-art and future perspectives. *Med. Image Anal.*, 89(102878):102878, Oct. 2023. 1

[20] J. Law and P. B. Macbeth. Ultrasound: From earth to space. *Mcgill J. Med.*, 13(2):59, June 2011. 1

[21] K. Li, J. Wang, Y. Xu, H. Qin, D. Liu, L. Liu, and M. Q.-H. Meng. Autonomous navigation of an ultrasound probe towards standard scan planes with deep reinforcement learning. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2021. 6

[22] K. Mathiassen, J. E. Fjellin, K. Glette, P. K. Hol, and O. J. Elle. An ultrasound robotic system using the commercial robot UR5. *Front. Robot. AI*, 3, Feb. 2016. 1

[23] S. J. Millington, R. T. Arntfield, R. J. Guo, S. Koenig, P. Kory, V. Noble, H. Mallemat, and J. R. Schoenherr. The assessment of competency in thoracic sonography (ACTS) scale: validation of a tool for point-of-care ultrasound. *Crit. Ultrasound J.*, 9(1), Dec. 2017. 3, 4

[24] S. J. Millington, R. T. Arntfield, M. Hewak, S. J. Hamstra, Y. Beaulieu, B. Hibbert, S. Koenig, P. Kory, P. Mayo, and J. R. Schoenherr. The rapid assessment of competency in echocardiography scale: Validation of a tool for point-of-care ultrasound. *J. Ultrasound Med.*, 35(7):1457–1463, July 2016. 3, 4

[25] G. P. Mylonas, P. Giataganas, M. Chaudery, V. Vitiello, A. Darzi, and G.-Z. Yang. Autonomous eFAST ultrasound scanning by a robotic manipulator using learning from demonstrations. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Nov. 2013. 6

[26] G. Ning, H. Liang, X. Zhang, and H. Liao. Autonomous robotic ultrasound vascular imaging system with decoupled control strategy for external-vision-free environments. *IEEE Trans. Biomed. Eng.*, 70(11):3166–3177, Nov. 2023. 1

[27] D. Raina, S. H. Chandrashekhara, R. Voyles, J. Wachs, and S. K. Saha. Robotic sonographer: Autonomous robotic ultrasound using domain expertise in bayesian optimization. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2023. 1

[28] J. A. Sánchez-Margallo, C. Plaza de Miguel, R. A. Fernández Anzules, and F. M. Sánchez-Margallo. Application of mixed reality in medical training and surgical planning focused on minimally invasive surgery. *Front. Virtual Real.*, 2, Oct. 2021. 2

[29] K. Su, J. Liu, X. Ren, Y. Huo, G. Du, W. Zhao, X. Wang, B. Liang, D. Li, and P. X. Liu. A fully autonomous robotic ultrasound system for thyroid scanning. *Nat. Commun.*, 15(1):4004, May 2024. 1