Augmented Reality as a Telemedicine Platform for Remote Procedural Training

by

© Shiyao Wang

A thesis submitted to the

School of Graduate Studies

in partial fulfilment of the

requirements for the degree of

Master of Science

Department of Computer Science

Memorial University of Newfoundland

May 2017

St. John's Newfoundland

Abstract

Traditionally, rural areas in many countries are limited by a lack of access to health care due to the inherent challenges associated with recruitment and retention of healthcare professionals. Telemedicine, which uses communication technology to deliver medical services over distance, is an economical and potentially effective way to address this problem. In this research, we develop a new telepresence application using an augmented reality (AR) system. We explore the use of the Microsoft HoloLens to facilitate and enhance remote medical training. Intrinsic advantages of AR systems enable remote learners to perform complex medical procedures such as Point of Care Ultrasound (PoCUS) without visual interference. This research uses the HoloLens to capture the first-person view of a simulated rural emergency room (ER) through mixed reality capture (MRC) and serves as a novel telemedicine platform with remote pointing capabilities. The mentor's hand gestures are captured using a Leap Motion and virtually displayed in the AR space of the HoloLens. To explore the feasibility of the developed platform, twelve novice medical trainees were guided by a mentor through a simulated ultrasound exploration in a trauma scenario, as part of a pilot user study. The study explores the utility of the system from the trainees, mentor, and objective observers' perspectives and compares the findings to that of a more traditional multi-camera telemedicine solution. The results obtained provide valuable insight and guidance for the development of an AR-supported telemedicine platform.

Acknowledgements

I would like to express my sincere thanks to my supervisors, Dr. Oscar Meruvia-Pastor, Dr. Minglun Gong and Dr. Andrew Smith. They have been providing me the possible research direction of AR-aided telemedicine, detailed discussions, and encouraging innovative ideas and critical thinking. All these are the key points guiding me to be an independent researcher.

I would like to acknowledge that Ignite R&D program, which is managed by Research and Development Corporation (RDC) of Newfoundland and Labrador, and provided me continuous financial support [RDC Project #5404.1890.101].

Additionally, I would like to give my special thanks to the researchers for details discussions related to the HoloLens from Microsoft and other organizations. They provided professional support of holographic programming, network programming, and user inteface design.

Last but not the least, I would like to thank my mother who have been supporting me from all sides.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Rural Healthcare Problems

Frequently, the provision of healthcare to individuals in rural areas represents a significant logistical challenge resulting from geographic, demographic and socioeconomic factors. Recruitment and retention of healthcare providers (HCP) to rural locations continues to be a significant problem [13]. Research focused on addressing the problems associated with the provision of rural healthcare is a top priority in many countries [14].

Newfoundland and Labrador population is 41% rural according to the 2001 Census [15], making rural healthcare a very important problem in our province. Several provinces have an even higher proportion of rural population (Table 1.1), which makes rural health care provision a national issue in Canada. In fact, this is a global problem, particularly in developing countries.

An economical and effective solution to the lack of HCP in rural areas is telemedicine,

Table 1.1: Rural Population Statistics, Canada 2001 [1]

	RURAL POPULATION	PERCENT(%)
Canada	6,098,883	20.3
Yukon	11,831	41.3
Northwest Territories	15,529	41.6
Newfoundland and Labrador	216,734	42.3
Nova Scotia	400,998	44.2
New Brunswick	361,596	49.6
Prince Edward Island	74,619	55.2
Nunavut	18,056	67.5

which uses information technologies to deliver health care services over both large and small distances [14, 16]. Telemedicine has many advantages, such as improved access to primary and specialized health services, improved continuity of care, increased availability of patient information, and decreased frequency of patient visits to health care specialists [17]. It also has been shown to increase patient empowerment and patient/provider satisfaction [16], while decreasing unnecessary referrals, travel and wait times, as well as the associated costs for both patients and providers [18].

1.2 Current Limitations

Teleconferencing is one of the main applications within telemedicine, enabling healthcare providers to interact with patients or colleagues on a regular basis. Current "talking head" interfaces (Figure 1.1) used in traditional teleconferencing systems may be adequate for supporting one-on-one communication between a doctor and a patient, or even a group of doctors, but may be unsuitable in a more chaotic environment such as an emergency room (ER). Mobile robot systems have been deployed in rural settings such as "Rosie the Robot" in Nain, Labrador [4] (Figure 1.2), in an attempt to address these problems. However, they remain quite expensive. Real-time consultation and support during low-frequency, high-stakes scenarios has the potential to enhance acute medical care. A system that provides a better immersive experience coupled with real-time consultation could improve performance during complex life-saving medical procedures.

Figure 1.1: "Talking head" Telepresence [3]

Figure 1.2: Rosie the Robot [4]

1.3 Research Focus

This research aims at the question of how to take advantage of the HoloLens within a telemedicine AR application. The potential of AR technology has always been significant [19, 20]. Even though researchers can nowadays immerse themselves in more complex virtual environments and realistic simulations, the concept of using a computer-mediated reality system in a hospital without a dedicated technician remains a hurdle as these systems are still subject to inherent technical limitations. For example, Google Glass lacked a 3D display, environment recognition ability, and had a very small field of view to be of practical use. Since the introduction of immersive VR HMDs, such as the Oculus Rift and the HTC VIVE, VR has become more accessible as a viable option. However, these and similar devices are still tethered to workstations or have limited computing power. In this sense, the HoloLens has some particular advantages, since it has adequate computing power, does not require any tethering and does not occlude the users' field of view. In spite of these advantages, significant efforts and multi-disciplinary cooperation is still required to assess the suitability of this and similar tools for practical use in telemedicine. Our goal in this research will be building a telemedicine platform with the help of HoloLens and Leap Motion controller (see Section 1.7). The mentor can use Leap Motion for hand tracking, and then position a virtual hand in the view of a remote trainee wearing the HoloLens.

1.4 Explosion of Computer-Mediated Reality

Virtual, Mixed and Augmented Reality (VMAR), together with unmanned aerial vehicles, autonomous cars, smart homes, as well as High Dynamic Range (HDR) Imaging are listed as top trends at the CES (Consumer Electronics Show) in 2016.

Google released its Glass project in 2013, a technology that enabled users to connect a wearable camera and heads-up display to mobile phones via Wi-Fi. This was followed up with Google's release of Cardboard [21], a simple cardboard box capable of transforming the ubiquitous smartphone into a virtual reality (VR) Head Mounted Display (HMD). Cardboard was instrumental in generating global interest and development of VR applications due to its broad appeal and accessibility. Attention then shifted to the Oculus Rift and HTC Vive, commercial immersive VR HMD systems connected to full computer workstations for increased performance and graphics power. More recently, Microsoft released the HoloLens in 2015. HoloLens was the first AR HMD capable of spatial capture of its environment [22]. Apart from these products, 360 degree cameras such as RICOH Theta S [23] and depth sensors such as Leap Motion controller (see Section 1.7) were all playing an important role in the field of VMAR. All of these products demonstrate the incremental and steady progression towards immersive AR/VR platforms and mass-market appeal (Table 1.2). That being said, computer-mediated systems are still relatively immature, with related techniques and applications waiting to be implemented and explored.

Table 1.2: Low cost VR/AR headsets and select peripherals [2]

Device	Type	Required Hardware	Price(USD)
Google Cardboard	VR Headset	Mobile	$15
Samsung Gear VR	VR Headset	Mobile	$99
Oculus Rift	VR Headset	PC	$599
HTC Vive	VR Headset	PC	$799
Microsoft HoloLens	MR/AR Headset	None	$3000
RICOH Theta S	360 Degree Camera	PC	$350
Leap Motion	Depth Sensor	PC	$79.99

1.5 What is Microsoft HoloLens?

The HoloLens (Figure 1.3) is a wearable computing device issued by Microsoft for augmented reality. It has several key elements:

Figure 1.3: Microsoft HoloLens

- It is an augmented reality (AR) product, which simultaneously presents computer-generated images together with the real world to the user. There are several AR products such as Google Glass which projects the image to the user, as well as mobile AR apps which renders additional objects on top of the camera outputs.

- It has an independent processing unit, including CPU, and GPU/HPU. It requires no external computers, which is vital for an AR helmet that emphasizes environment capture and processing. The CPU is an Intel Atom processor, and part of the Airmont family (Braswell). That processor runs at 1.04GHz, and while it is 64-bit capable, the OS itself is only 32-bit. In this case, the RAM is only 2GB on board.

- The GPU, or HPU on the HoloLens, is created by Microsoft. The full name is Holographic Processing Unit, which is responsible for handling holographic rendering and display. The HPU is a dedicated Application-specific integrated circuit (ASIS), custom designed for the HoloLens. This Microsoft's custom coprocessor (HPU) can reportedly processes terabytes of information from all of the HoloLens sensors in real time [24].

1.6 The HoloLens vs. Google Glass

Google Glass was one key AR product which provides non-occluded vision. Compared to Google Glass, the HoloLens has three-dimensional depth perception, enabling it to perform environment reconstruction and relocation. This is significantly different from Google Glass which only deals with RGB information.

In addition, the HoloLens has three-dimensional rendering ability. The HoloLens display is a set of transparent screens placed in front of the users' eyes. Each screen lets some background light through and simultaneously shows digital content just like a see-through display. Two screens each show a slightly different image independently to one side of your eye, creating a stereoscopic illusion like 3D glasses do when watching 3D movies. However, Google Glass just has a one-eye display.

Finally, the HoloLens has a human-computer interactive system containing gesture and voice recognition, enabling the user to control the HoloLens in multiple ways.

1.7 Leap Motion Controller

The Leap Motion controller is a small USB peripheral device as shown in Figure 1.4, which is designed to be connected to a computer. For virtual reality headsets connected to a computer such as Oculus Rift, the Leap Motion can also be mounted onto the headset. The cable of the Leap Motion should still be connected to the computer, and hand motion will be interpreted from the computer to the headset. It is commonly used to capture hand movements and gestures [25, 26].

Figure 1.4: Leap Motion Controller

1.8 Research History at Memorial University

Memorial University has a long history of contributing to ground-breaking advances in Telemedicine [27]. A self-sufficient telemedicine center had been created, which prevented the discontinuity commonly found in the programs after the grand funds ended. Under the leadership of Dr. Max House, the founder of the center, Memorial was involved in several Telemedicine projects, many of which became ongoing services [27]. This multidisciplinary initiative represented a significant step towards positioning Memorial as an innovator in Telemedicine, with world-class expertise in the provision of rural and remote healthcare. The Research and Development (R&D) focus of this and other research projects will support increasing collaborations with Eastern Health and the Newfoundland and Labrador Centre for Health Information as we work towards a more sustainable model of health care in the province.

1.9 Point of Care Ultrasound

The use of portable ultrasound by clinicians outside of the radiology suite is rapidly increasing throughout the field of medicine. Its portability and reduced cost, combined with the fact that scans can be performed rapidly, repeatedly and without exposure to ionizing radiation, has enabled its widespread acceptance [28, 29]. Performance of Point of Care Ultrasound (PoCUS) is a complex resource-intensive task similar to other medical procedures that require repetition under expert supervision over time. Telemedicine has been successfully used to support remote training both in rural, remote and even extra-terrestrial environments [30, 31].

Related research of VR/AR and related technologies in medical field will be discussed in the next chapter (Chapter 2), the history of telemedicine and the medical research about Google Glass and Microsoft HoloLens will also be present.

Chapter 2

Background and Related Work

2.1 Telemedicine

In 1924, the cover of the magazine "Radio News" showed a radio doctor (Figure 2.1), which was considered to be "the first exposition of Telecare" [5]. There has been a massive boom of telemedicine programs all over the United States since the 1960's [6]. On December 7th, 1966, NASA launched the ATS-1 Satellite (Figure 2.2), the first satellite used for civic purposes, such as education, and telemedicine [6].

Canada was one of the first countries to use tele-communications to assist in the delivery of health services. In the late 1950s, Dr. Albert Jutras, a radiologist in Montreal, used closed-circuit television to transmit medical images [32]. Communication satellites were used as early as the 1960s to send electroencephalograms (EEGs) over large distances [33]. In 1976, the launching of the Hermes-CTS communications satellite permitted Canada and the United States to embark on extensive telemedicine experiments and pilot projects to reach the far northern regions and to test wider applications [33].

The Canadian government has continued to support the concept of using communications technologies to provide medical care and education at a distance due to its climate and geography. In the early 1990's, the Internet went public and telemedicine expanded exponentially [6].

Figure 2.1: "Radio News" Magazine [5]

Figure 2.2: ATS-1 Satellite [6]

2.2 Virtual Reality (VR) History

It is well believed that the first attempt at Virtual Reality (VR) came in the 1860's, as artists began to create three-dimensional, panoramic murals [7]. In 1957, Morton Heilig invented the Sensorama (Figure 2.3), a simulator with 3D images along with smells, wind and sound, to create the illusion of a virtual reality [7]. Jaron Lanier, founder of VPL Research and creator of the DataGlove and the EyePhone, was credited with coining the term “Virtual Reality” in 1987 [7]. VR was used for military training and as a form of therapy since 1990’s [34].

In 2012 the company called Oculus turned to fundraising platform KickStarter to finance the Oculus developer kit, which was meant to get the Oculus Rift to developers who could then integrate the VR device into their games. Oculus received a lot of attention from the media and, in April 2014, the Federal Trade Commission approved Facebook’s purchase of Oculus for nearly USD $2 billion [35]. Today’s VR is perhaps best known for its use in the gaming world.

Figure 2.3: Sensorama [7]

2.3 Augmented Reality (AR) History

The first related idea of AR is believed to have come in 1901 when L. Frank Baum, in his novel "The Master Key: An Electrical Fairy Tale" described a set of electronic glasses that provided insight into a person's character [8]. The term "Augmented Reality" was coined in 1990 by Boeing researcher Tom Caudell [8]. Louis Rosenberg developed Virtual Fixtures (Figure 2.4), one of the first functioning AR systems, for the Air Force [8]. This allowed the military to work in remote areas. Wearable AR made headlines in 2014, mostly thanks to Google Glass. However, other companies, like Epson, had also developed their own smart glasses. The start-up company Innovega was then taking smart glasses one step further, introducing AR contact lenses [36]. Looking ahead, the development of the HoloLens system by Microsoft in 2015 and its native support in Windows 10 could facilitate the development of a new range of training and educational applications.

Figure 2.4: Virtual Fixtures [8]

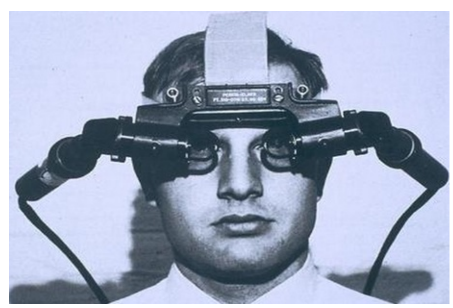

Figure 2.5: Ultimate Display [9]

2.4 Prototype of VR/AR headset

Both VR and AR headsets come from the same general model. Ivan Sutherland devised a machine called the Ultimate Display (Figure 2.5) in 1965 [9]. Using a head-mounted display (HMD) connected to a computer, users could see a virtual world. The system was suspended from the ceiling. The computer-generated graphics users saw were just simple wireframe drawings. He was thus credited with creating the first HMD system.

2.5 VR Research in Medicine

An early application of VR for health care provision started in the early 90's intended to visualize complex medical data during surgery or surgery planning [37]. Surgery-related applications of VR were divided into two categories: surgery training and surgery planning [38]. VR applications in medicine have broadened to a range of disciplines including neuropsychological assessment and rehabilitation, such as the treatment of anxiety disorders, post-traumatic stress disorders, eating disorders and obesity [39].

From the clinical practice perspective, VR was considered a "closed" experience, produced and lived in the therapist's office only, separated from the patient in the real world. To overcome this issue, a critical advancement can be exemplified by a new technological paradigm, Interreality [40, 41] a hybrid, closed-loop empowering experience that uses smartphones and wearable devices to bridge physical and virtual worlds.

2.6 AR Research in Medicine

Doctors can use AR as a visualization and mentoring aid in open surgery, endoscopy, and radiosurgery [38]. It has also commonly been used in orthopedic surgery, neurosurgery and oral maxillofacial (OMF) surgery [42], enabling the surgeon to visualize the proper positioning of their surgical instruments. AR is also useful when operating in a confined space and in close proximity to delicate and sensitive anatomical structures [43]. Many studies suggest that AR-assisted surgery appears to improve accuracy and decrease intraoperative errors in contrast to traditional, non-AR surgery [43, 44, 45, 46, 47]. However, further technological development and research is needed before AR systems can become widely adopted. General medical visualization is another task for AR to access and display types of necessary data simultaneously virtually in the surgical suite [38]. AR has the potential to support the fusion of 3D datasets of patients in real time, using non-invasive sensors like magnetic resonance imaging (MRI), computed tomography scans (CT), or ultrasound imaging. All information could then be rendered with a view of the patient at the same time, like "X-ray vision" (Figure 2.6 and 2.7) [38]. For medical training and education, AR can play an important role [48, 49]. However, gesture interaction in AR has been found to be too complicated, for both trainees and mentors [50, 51].

2.7 Augmented Reality Research in Telemedicine

The early research mentioned in the previous section provided relevant directions and presented valuable solutions in the medical field. More advanced systems have

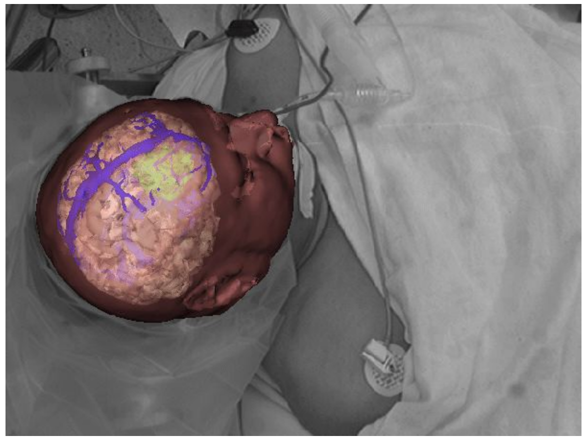

Figure 2.6: Surgeon X-ray Vision [10]

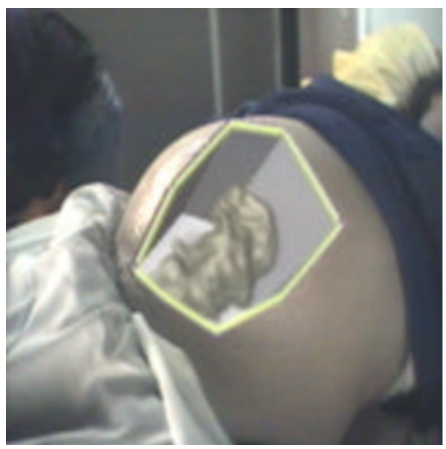

Figure 2.7: Virtual fetus of pregnant patient [11, 12]

been created as the technology has evolved. Ruzena Bajcsy et al. [52, 53] collected patient's depth maps through Microsoft Kinect and then reconstructed a virtual patient in an AR device. Using telemedicine, the mentor could then provide consultation based on the 3D model at a distance as shown in several previously developed teleconsultation applications [39, 54]. However, the application required massive fund and setup[39, 52, 53, 54]. Marina Carbone et al. [55] and Mahesh B Shenai et al.

[56] created AR-assisted telemedicine applications. However, their AR system still required significant setup in both sides and had some shortcomings. It combined video from a computer-generated image and a camera-captured video, which is not as realistic as the combination of the HoloLens see-through stereoscopic vision and 3D graphics imagery. Their systems were not validated through a comparison with other more traditional telemedicine setups. Telemedicine has been proposed to solve the lack of HCP in remote locations, however, if the telemedicine application requires significant setup and even requires technical professionals in rural locations, it would lead to a new problem regarding the lack of technicians. All previous systems have this problem, while our system only requires the trainee to wear the HoloLens, which is a self-contained solution specially suitable for telemedicine. Our research attempts to overcome the limitations in previous works by designing a new telemedicine architecture using the latest telecomunication protocols and the Microsoft HoloLens. This work also provides insight into how our solution compares to more traditional telemedicine solutions.

2.8 Advantages of the HoloLens

One of the main strengths of the HoloLens as a telemedicine platform is that it is untethered - a feature valuable for chaotic environments such as the ER or operating room. It is a non-occluding AR system, in that it complements the actual scene with a relatively small amount of computer-generated imagery using a semi-transparent HMD. Furthermore, it enables a first-person view of the remote environment to be relayed and represented locally to expert observers at a remote location through a camera mounted in the middle of the HMD. Such a telepresence interface [57, 58] has the potential to enhance the observers' sense of presence, enabling them to better understand crucial circumstances and provide better guidance at critical times. From the remote learners' perspective, the HoloLens enables recipients to participate in real-time mentoring as well as 'just-in-time' learning during extremely stressful situations. The ability to receive visual guidance and instructions during an infrequently performed complex medical procedure represents a significant advance for emergency personnel. A final feature is the HoloLens' intrinsic depth-sensing and relocation ability which can be used to support remote pointing and enhance procedural guidance. This last element is the main subject of this research. The HoloLens can simultaneously interpret its own location inside a room when the user is moving. Therefore, we can render a virtual object statically in the environment and keep it stable. With the help of this, the remote presenter can steadily position a virtual object to a real location.

2.9 Disadvantages of the HoloLens

Even though Microsoft manufactures the HoloLens with a decent 120 degrees field of view (FOV), it is still not comparable to a fully immersive HMD [59]. The weight of the HoloLens is also a problem. Discomfort and pain reports can easily be found in the literature regarding to the HoloLens [59, 60]. In addition, the ergonomics of the HoloLens are described as disappointing in various aspects, including the "airtap" gesture, weight, vision and comfort [59, 60]. The HoloLens is also significantly lower-resolution [59] than full HD monitors. Furthermore, the battery of the HoloLens could only last for approximately 100 minutes when running an application before having to be charged again. Another issue was that the HoloLens would sometimes kill an application in order to protect itself, due to the limited memory size [59]. A further limitation is that it has been designed to be exclusively as an indoor device, designed to capture its surroundings in closed environments, such as laboratories, offices and classrooms. Other disadvantages will be described in more detail in the following Chapters.

2.10 Stereo Vision and Simultaneous Localization And Mapping (SLAM)

Part of the "reality" to the HoloLens is that it is a see-through display that allow viewers to see the real world without any cameras used as mediators. The other part of the reality is the data captured about the environment where the user is located. This is captured from multiple sensors, a.k.a., cameras. The HoloLens contains four cameras designed to enhance environmental understanding, two on each side, providing support for stereoscopic vision (stereo vision for brevity). Through real-time analysis of four video streams, the HoloLens has a 120 degree Field of View (FOV) horizontally and vertically.

Stereo vision (Figure 2.9) is an essential part of computer vision and is used to obtain the depth map from images captured by multiple cameras, by focusing on recovering the distance between all the objects in the captured environment and the device, as shown in Figure 2.8.

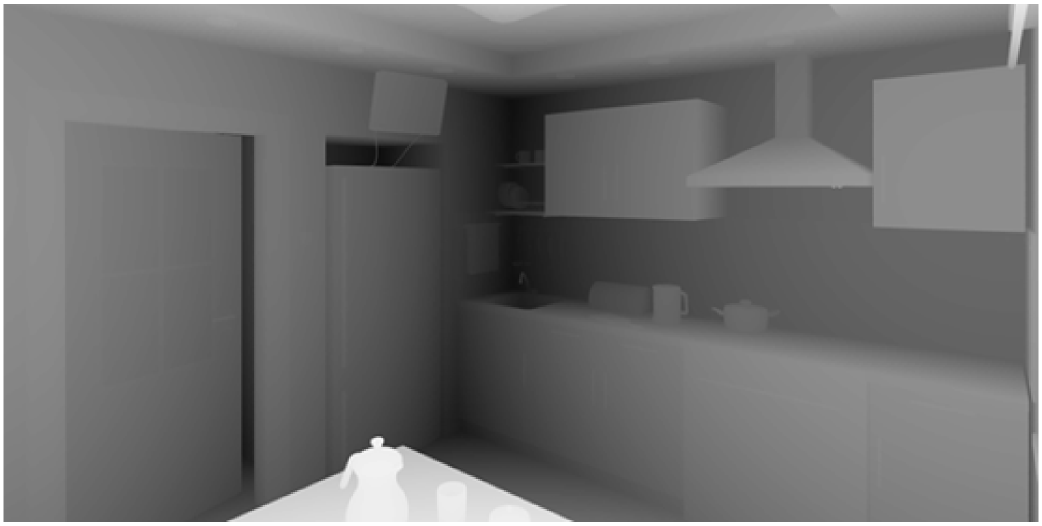

Figure 2.8: Depth Map

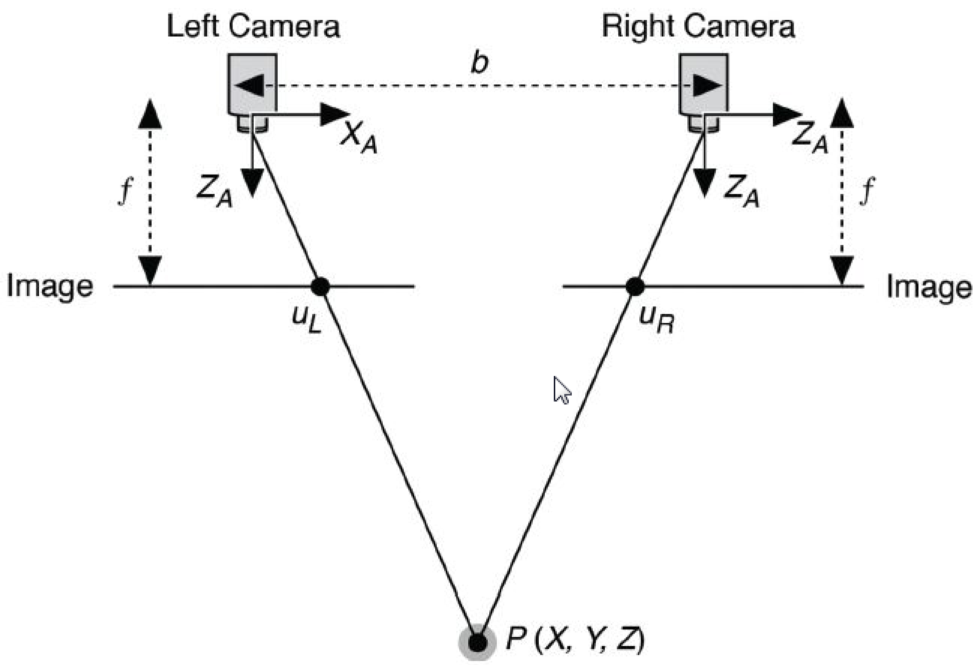

Figure 2.9: Stereo Vision

Stereo vision begins with a procedure designed to eliminate the distortion of a camera using a pre-calibration process. In the next step, rectification is used to identify the overlapping portions of each image followed by the creation of a disparity map. Finally, a re-projection function is used to calculate the final depth map [61].

One single depth map simply represents a snapshot of the environment from a single location and point in time. In order to reconstruct the whole scene, we need depth information from multiple samples. Therefore, a series of depth maps are used to reconstruct the environment.

SLAM (Simultaneous Localization and Mapping) is a solution to reconstruction and relocation, meant to address two fundamental problems: i) where is the user or robot located within the environment, and ii) where the user or robot should go. SLAM has multiple implementations, including Microsoft's Kinect Fusion [62, 63] and open source libraries such as Point Cloud Library (PCL) [64].

Kinect Fusion (Figure 2.10 [62, 63]) captures different angles of the indoor environment from a moving Kinect and calculates the room model from those depth maps. In order to obtain a raw depth map, a Depth Map Conversion should be performed to convert the data into meters. After that, we can get the coordinates of vertices and normal maps of surfaces. Then, we will be able to track the camera pose and location with these coordinates and normal maps. Using depth maps and camera coordinates from the previous step, we can then use Volumetric Integration to get the 3D model. Additionally, we can also perform 3D Rendering techniques, such as Raycasting, Raytracing and Rasterization to render the 3D scene.

A proprietary and undisclosed simultaneous localization and mapping (SLAM) algorithm is embedded in the HoloLens. It is thought to be a modified Kinect Fusion algorithm ([65]), with the main difference being that Kinect Fusion is open to public. Unfortunately, there is a limited ability to modify the HoloLens depth map/data,

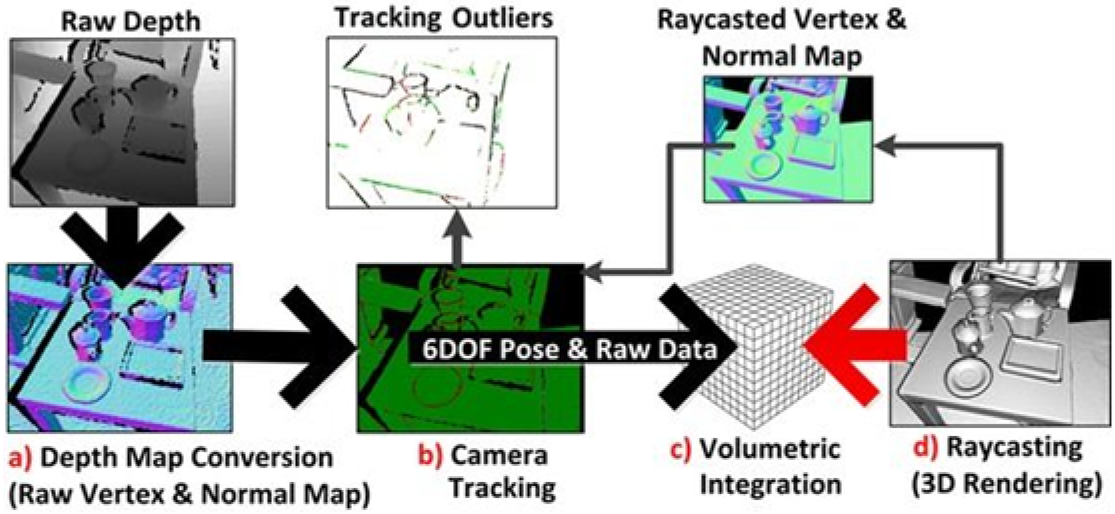

Figure 2.10: Kinect Fusion workflow

and this constrains the ability to research and develop customized solutions. This "black-box" approach does have several advantages from a commercial development perspective, though. No extra knowledge is needed, no complicated programing environment needs to be set, no optimization needs to be performed. The degree of expertise required for this approach is considerably lower. Therefore, if we just need a SLAM solution, the HoloLens is a good option.

2.11 Google Glass and Microsoft HoloLens

Google Glass has been tested in a wide range of medical applications since 2014. Muensterer et al. explored its use in pediatric surgery, and concluded that the Glass had some utility in the clinical setting [66]. However, substantial improvements were needed prior to actual development in the medical field related to data security and specialized medical applications [66]. Other applications include Glass being used for Disaster Telemedicine triage, but no increase in triage accuracy was found [67]. Mentoring was also studied in which recorded videos were played with Google Glass [68], while a similar approach was used to enable telemedicine communication [69]. Research has also explored pre-hospital care, in which Glass acted like a console for transferring patient data [70]. However, Glass could not show any advantage compared to mobile devices in this study.

Due to its novelty, research literature using the Microsoft HoloLens (released in 2015) is still scarce, especially in the medical field. Nan Cui et al. [71] have used it in near-infrared fluorescence-based image guided surgery in which the HoloLens was used to provide vital information such as location of cancerous tissue to the surgeon. Additionally, in [72] the HoloLens was used to elicit gestures for the exploration of MRI volumetric data.

The design and the overview of the system will be presented in Chapter 3.

Chapter 3

System Design and Overview

In order to test the possibility that the HoloLens can be used in the field of remote ultrasound training, we developed several prototypes covering different approaches of telecommunication technologies. Those prototypes demonstrated different shortcomings, which illuminated a feasible solution to the problem. With the help of those prototypes, we proposed a final design to use of the HoloLens in a telemedicine application. Further detail about those prototypes can be found in Appendix B. An important technical aspect of the implementation is the video streaming solution we chose for use with the HoloLens. Appendix C discusses this aspect in more detail.

3.1 Final Design

For our final design, we took the following observations and requirements into account:

- Latency is an important factor in the quality of the teleconference experience and should be kept to a minimum.

- Voice communication is critical for mentoring. Video conferencing within the AR without two-way voice communication was found generally less valuable.
- Immersive VR HMD for the mentors creates more challenges and requires significant technical development prior to enhancing telemedicine.
- The simplicity and familiarity of conventional technology for the mentor was an important aspect that should remain in the proposed solution.
- Remote pointing and display of hand gestures from the mentor to the trainee would be helpful for training purposes.
- Specific to ultrasound teaching, a hologram with a hand model provided additional context for remote training.

We proposed a design in order to address the requirements above through the following implementation:

1. The Leap Motion sensor was used to capture the hand and finger motion of the mentor in order to project into the AR space of the trainee.
2. Three static holograms depicting specific hand positions holding the ultrasound probe was generated and controlled by the mentor using the Leap Motion.
3. MRC (video, hologram and audio) was streamed to the mentor while the mentor's voice and hologram representations of the mentors' hand(s) was sent to the trainee to support learning.

4. Hand model data captured by Leap Motion was serialized and bundled together with the mentor's audio at a controlled rate to minimize latency while maintaining adequate communications.

3.2 System Overview

3.2.1 The Mentor's End

We implemented an application using the Unity game engine. The final application was run on a laptop PC with a Leap Motion sensor attached to it. The hand gestures were captured and manipulated using the Leap Motion SDK v3.1.3. There are four different ultrasound probe holding postures needed for remote ultrasound teaching. Different hand gestures would be recognized as one of four holding posture from the mentor's side, as shown in Figure 3.1. Buttons that represent different gestures were also displayed for clicking as an alternative to compensate in case of malfunction of gesture recognition. The data from the Leap Motion was sent to the application and then serialized and compressed. We used a Logitech headphone to eliminate the presence of audio echo and to emphasize the remote sounds by keeping the surrounding noise to a minimum. The audio data from the headphone was also captured and encoded using the A-law algorithm. The computer exchanged data with the HoloLens located in a separated simulated ER (details below). The MRC video received from the HoloLens was rendered and played by a free add-on to stream video to texture using VLC media backend [73].

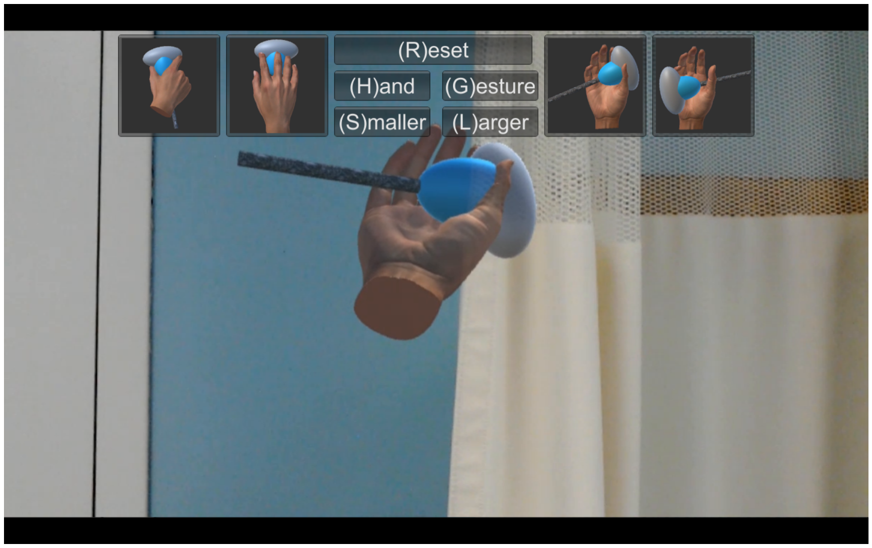

Figure 3.1: Mentor side of view. A live mixed reality view displaying with buttons.

3.2.2 The Trainee's End

We developed another application using the Unity game engine with HoloLens support. The hand models were created based on the Leap Motion Unity asset Orion v4.1.4. Several preliminary Unity 3D objects (cubes, cylinders, spheres) were combined to represent an ultrasound transducer being held in a hand model, as shown in Figure 3.2. The orientation and position of the hand were simulated through the data received from the mentor's side. The audio data was decoded and played. The MRC live video was captured through Microsoft's Device Portal REST API [22, 74].

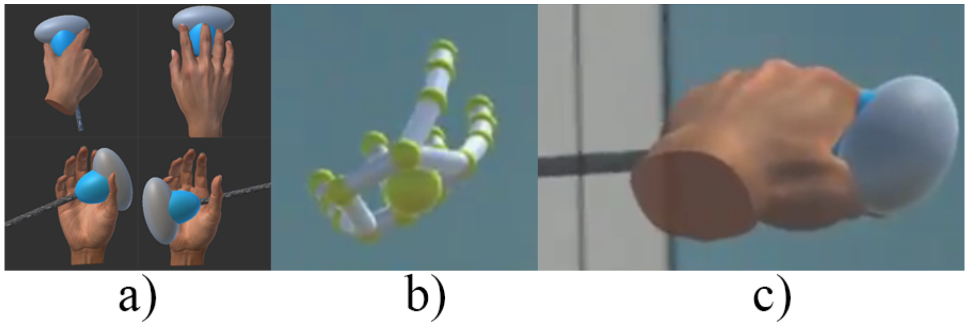

Figure 3.2: Trainee side of view. a) four holograms represent different posture. b) skeletal hand view on the HoloLens. c) real hologram view on the HoloLens.

3.2.3 Settings

The MRC video from the trainee was captured and broadcasted by a built-in web-server running in the HoloLens. The hand data and audio data from the mentor were transmitted using Unity's built-in multiplayer networking system called UNET. Both the HoloLens and the laptop were connected through a local area network. An overview of the system is shown in Figure 3.3. During the experiment, the mentor and the trainee were in separate rooms to perform a simulated teleconference session.

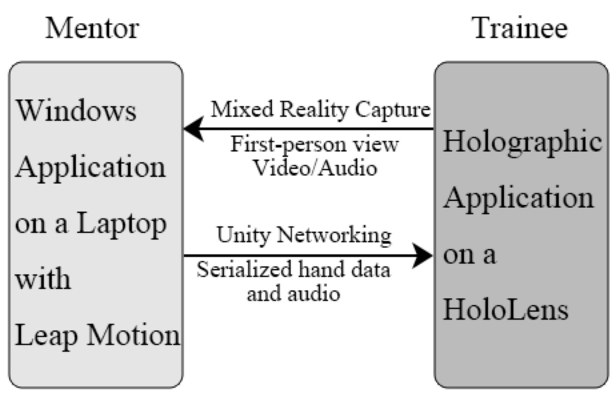

Figure 3.3: Overview of the system pipeline

Chapter 4

Experimental Validation

Point of Care Ultrasound (PoCUS) represents a complex medical procedure usually performed under extremely stressful circumstances. In-person, hands-on training is highly effective, but this remains a significant challenge for rural practitioners seeking initial training or maintenance of skill. The combination of Microsoft's HoloLens and Leap Motion represents an AR platform capable of supporting remote procedural training. In this research, we have performed a pilot user study to explore the feasibility and user experiences of novice practitioners and a mentor using AR to enhance remote PoCUS training and compare the performance to a standard remote training platform.

4.1 Methods

4.1.1 Participants

Twelve first and second year medical students with minimal PoCUS experience were enrolled in the pilot study. Minimal experience is defined as having previously performed 5 or less PoCUS scans. Data from a recent remote PoCUS training study involving a similar study design and cohort but with different participants is used for baseline comparison. Further details about the reference setup are introduced in the next section. One mentor guided all medical students in order to maintain consistency across subjects.

4.1.2 Experimental Control

We compared our solution against one of the configurations most commonly used for telemedicine today, which we refer to as a "full telemedicine setup", and which is used as the experimental control to validate our system. This setup consists of a full overhead view of the whole patient room captured through a PTZ camera near the ceiling and a second view of the patient captured from a webcam placed on the ultrasound machine. Both cameras were live streaming together with the ultrasound screen view from the remote side to the mentor side. VSee (Vsee Lab Inc.) was used for this secure, high-resolution and low-bandwidth video-conferencing task. Both mentor and trainees were wearing a headset to facilitate comunication.

4.1.3 Procedure

Each subject was asked to complete a right upper quadrant Focused Assessment using Sonography in Trauma (FAST) ultrasound examination on a healthy volunteer under the guidance of an experienced mentor (the same as in the study) while wearing the Microsoft HoloLens. In addition to verbal guidance, the mentor provided remotely a physical demonstration of hand position and proper exploration procedures using the Leap Motion. Performance of the trainee was independently observed and graded by a PoCUS expert using a Global Rating Scale (GRS). Participants and the mentor each completed a short Likert survey regarding the utility, simplicity and perceived usefulness of the technology. Cognitive load was assessed using time to perform the task, mental effort and task difficulty rating. A similar approach was used to collect comparison data for remote PoCUS training using standard telemedicine technology during a prior study. Informed written consent was provided prior to participation.

4.1.4 Ethics Approvals

The study design was reviewed and approved by the Health Research Ethics Authority (HREA) at Memorial University, and found to be in compliance with Memorial University's ethics policy (HREA Code: 20161306).

4.1.5 System Setup and Performance

Subjects were asked to wear the HoloLens prior to the start of the procedure. A curvilinear probe (1-5 Mhz) connected to a portable ultrasound (M-Turbo, Sonosite-FujiFilm) was used to perform the FAST examination. The ultrasound was connected to a laptop (Macbook Air, Apple) via a framegrabber (AV.io; Epiphan) and live-streamed over a local-area network via a standard communications software (VSee). Ultrasound streaming was both hardware and network independent from the HoloLens communications. The mentor was asked to wear a Logitech UE4500 headphone connected to a Windows PC. A Leap Motion was attached to the PC. The HoloLens and this PC were connected via a local-area network.

4.1.6 Data and Analysis

The GRS developed by Black et al. was used to objectively assess student performance on the FAST examination [75]. Students and the mentor were surveyed upon completion of the task using both a short Likert survey and open-ended feedback. Cognitive load was assessed using a combination of time taken for task completion and Likert questions. Participants provided a cognitive load and task difficulty measure for each scan, and completed a general information feedback questionnaire for the study. Data were entered into SPSS for analysis. An Independent-Samples T-Test was used for each analysis.

4.2 Results

4.2.1 Trainees

As can be seen in Figure 4.1, the feedback from the 12 participants assigned to use the HoloLens as their telemedicine tool was positive. They felt excited when using this new technology, and considered it useful for the study. Although there was a slight trend toward Full Telemedicine being superior to the HoloLens setup, there wasn't a statistically significant difference between HoloLens and Full Setup for the questions "The technology was easy to use", "The technology enhanced my ability to generate a suitable ultrasound image" and "The technology was overly complex". The numeric values are shown in Table 4.1.

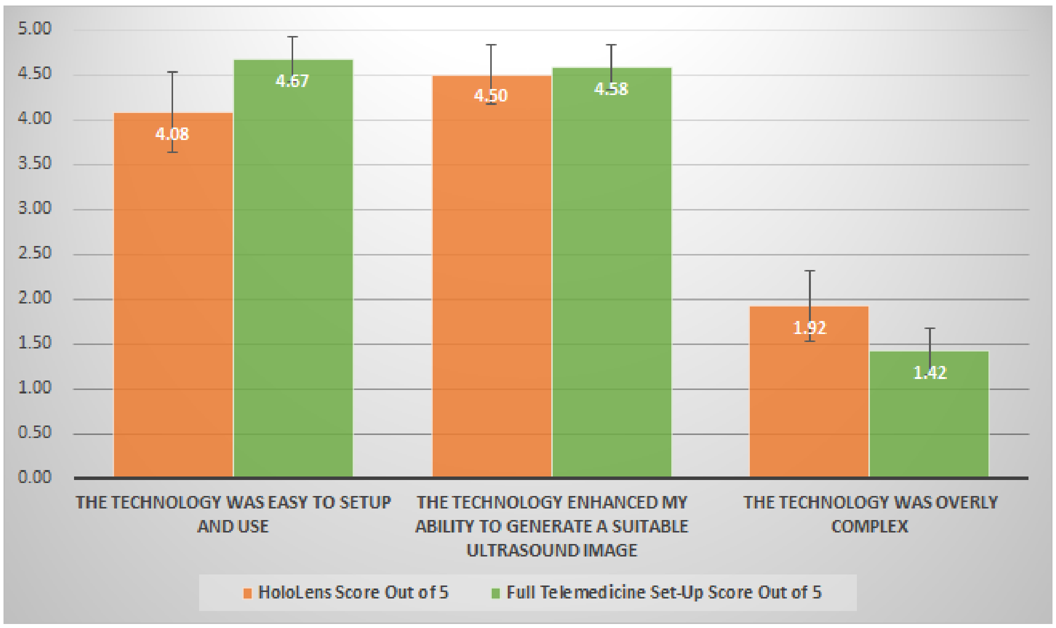

Figure 4.1: Trainee's opinions on the efficacy and difficulty of the HoloLens and Full Telemedicine Set-Up

4.2.2 Mentor

From the mentor's perspective, however, the technology did not reach expectations. For all categories from the mentor's perspective, the Full Telemedicine setup was significantly superior. A detailed comparison is shown in Figure 4.2, and the numeric

Table 4.1: Trainee's opinions on the efficacy and difficulty of the HoloLens and Full Telemedicine Set-Up

	HoloLens Score Out of 5 (Standard Deviation)	Full Telemedicine Set-Up Score Out of 5 (Standard Deviation)	P-Value
The technology was easy to setup and use	4.08 (0.90)	4.67 (0.49)	0.065
The technology enhanced my ability to generate a suitable ultrasound image	4.50 (0.67)	4.58 (0.51)	0.737
The technology was overly complex	1.92 (0.79)	1.42 (0.51)	0.081

values are shown in Table 4.2. It is important to note that there was only one mentor, so the results may have an inherent bias and cannot be generalized.

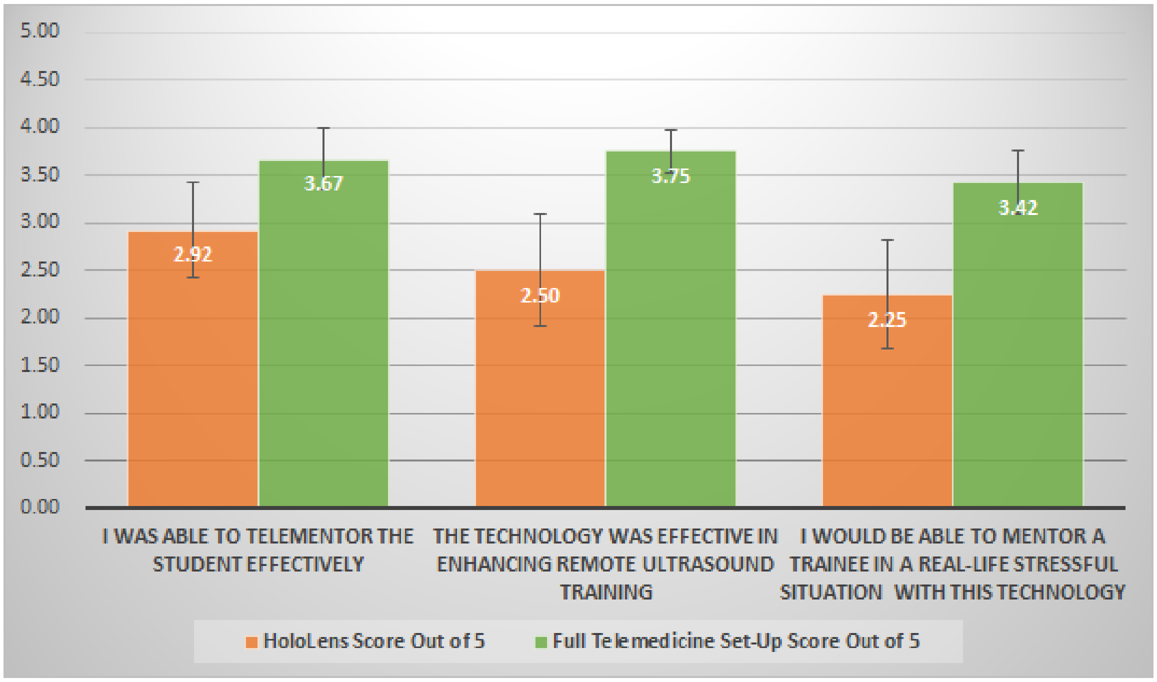

Figure 4.2: Mentor's opinions on the efficacy and real-life application of the HoloLens and Full Telemedicine Set-Up

4.2.3 Global Rating Scale (GRS)

From the expert evaluator's scores on the GRS for right upper quadrant exam, there was no significant statistical difference ($p = 0.534$) between the HoloLens application ($2.75, SD = 0.62$) and the full telemedicine setup ($2.91, SD = 0.67$) (Table 4.3).

Table 4.2: Mentor's opinions on the efficacy and real-life application of the HoloLens and Full Telemedicine Set-Up

	HoloLens Score Out of 5 (Standard Deviation)	Full Telemedicine Set-Up Score Out of 5 (Standard Deviation)	P-Value
I was able to telementor the student effectively	2.92 (1.00)	3.67 (0.65)	0.04
The technology was effective in enhancing remote ultrasound training	2.50 (1.17)	3.75 (0.45)	0.004
I would be able to mentor a trainee in a real-life stressful situation with this technology	2.25 (1.14)	3.42 (0.67)	0.007

Table 4.3: Global Rating Scale for right upper quadrant exam of the HoloLens and Full Telemedicine Set-Up

	HoloLens Score Out of 5 (Standard Deviation)	Full Telemedicine Set-Up Score Out of 5 (Standard Deviation)	P-Value
Preparation for Procedure	2.92(0.79)	3.00(0.60)	0.775
Patient Interaction	3.00(0.43)	3.08(0.51)	0.670
Image Optimization	3.00(0.60)	3.08(0.51)	0.719
Probe Technique	2.83(0.58)	2.83(0.72)	1.000
Overall Performance	2.75(0.62)	2.91(0.67)	0.534

4.2.4 Response time, mental effort and task difficulty ratings

We noticed that participants using the HoloLens application took much longer to finish the procedure (mean difference of 2.5 minutes) than participants completing the standard telemedicine setup (Figure 4.3). The time difference between the two was statistically significant. ($p = 0.01$). However, trends appeared to suggest that participants felt it was easier to use the HoloLens application to perform an ultrasound scan as the mental effort rating and task difficulty rating were lower than the full setup, though there was no significant difference between the groups (see Figure 4.4 and Table 4.4).

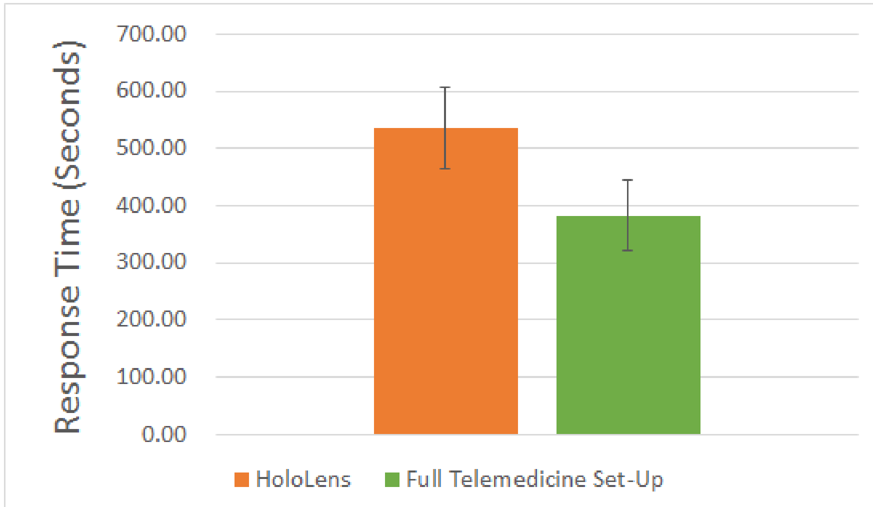

Figure 4.3: Trainee's Response time for the HoloLens and Full Telemedicine Set-Up

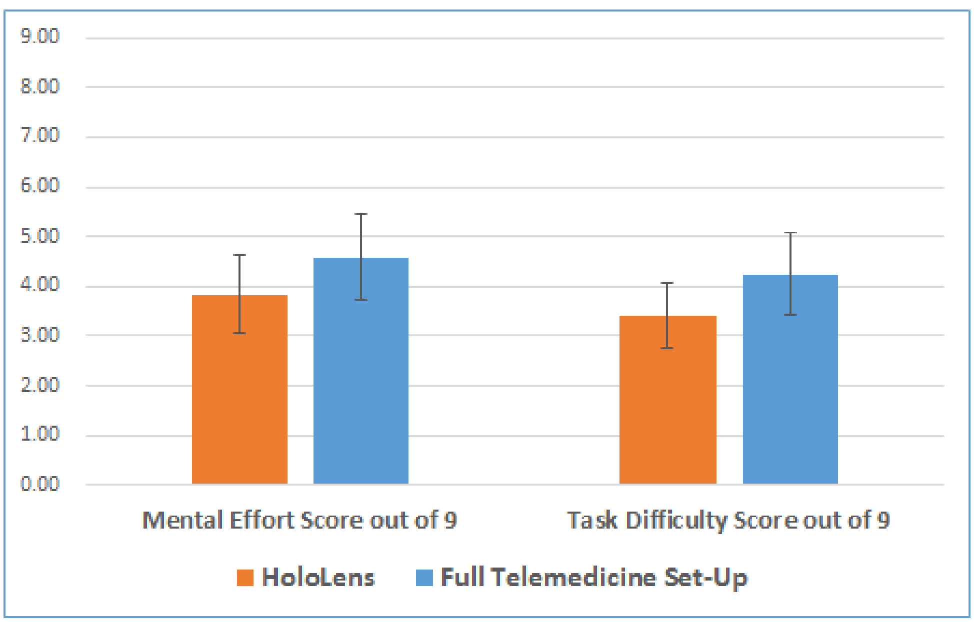

Figure 4.4: Trainee's perception of mental effort and task difficulty for the HoloLens and Full Telemedicine Set-Up (Mental effort and task difficulty were scored out of 9)

Table 4.4: Trainee's perception of response time, mental effort and task difficulty for the HoloLens and Full Telemedicine Set-Up

	HoloLens Score (Standard Deviation)	Full Telemedicine Set-Up Score (Standard Deviation)	P-Value
Response Time (Seconds)	536.00 (142.11)	382.25 (124.09)	0.010
Mental Effort Score out of 9	3.83 (1.59)	4.58 (1.73)	0.280
Task Difficulty Score out of 9	3.42 (1.31)	4.25 (1.66)	0.186

Chapter 5

Discussion

5.1 The Performance of the System

As described earlier in the Results section, there was no significant difference in overall trainee performance according to the expert evaluator. Also, the trainee rated mental effort and task difficulty slightly lower for the HoloLens, which suggested that the HoloLens application could potentially make the task easier, though there was not a statistically significant difference. However, the effectiveness of the system was rated low by the mentor. This suggests that the mentor felt it was harder to provide guidance with this setup. Furthermore, the HoloLens group took an average of 2.5 minutes longer to complete the ultrasound exploration compared to the full telemedicine group. This may be caused by frequent malfunction and bad connection quality of the HoloLens. During the study, the system did not perform as well as expected.

There were several problems with the HoloLens that impacted the user experience.

For example, some trainees felt that the HoloLens was too heavy and found it painful to wear. Most participants felt uncomfortable with the nose pad in particular. The nose pad should actually never be place where the weight of the device relies, because the device is too heavy. Instead, the HoloLens should be worn as a headband, so that the skull carries the weight of the device. Furthermore, some participants could not find a suitable fit to their head, as they had a smaller skull than the smallest fit available in the device. Even though the HoloLens has a decent field of view of 120 degrees horizontally, for many users this is still too narrow. This is particularly relevant if we consider that the entire human field of view is 210 degrees [76]. This greatly influenced the user experience for all the participants.

In the HoloLens, a stereoscopic image pair is projected to the user [77]. However, the mentor's display is just a 2D screen without stereoscopic vision. This drawback affects the performance for remote pointing, as the mentor may lose the sense of depth. Another limitation was that the HoloLens could last for only approximately 4 participants or about 100 minutes before having to be charged again. One participant even had to finish the study with a connected charging cable. Another issue experienced was that the HoloLens would sometimes quit the current running application when the user was looking towards a dark area. The application would also quit when the user's hand movements were accidently recognized as the "bloom" gesture, which would quit the current application and open the Start menu. On the other hand, some participants enjoyed using the HoloLens. In particular, they liked how the open and non-occluding system allowed them to perform other activities while wearing the HoloLens. They were able to finish survey forms, answer their phone and talk to others without removing the device. Some participants with VR experience also mentioned that wearing the HoloLens would not cause them to get dizzy like other VR devices.

Further details about the limitations of the system are discussed later in this chapter (see Section 5.4);

5.2 General Insights

Though we chose to perform the telementoring for a specific area of telemedicine (ultrasound training), most of our results have the potential to inform other applications across various disciplines and areas. We learned that, for building a communication application, the quality of connection (latency) would be the first problem noticed by an operator [78, 79]. During the experiment, we noticed that traditional user interfaces such as buttons and keyboards were more reliable compared to new ones such as gesture and speech. For inexperienced users, if the new user interfaces worked improperly only one or two times, they may abandon them. The HoloLens still has some limitations and is not yet ready for practical application. However, the idea of presenting 3D objects in addition to one's vision may still be beneficial in various scenarios such as virtual fitting dressing room, remote presenting and remote teaching. We also learned that the performance was not always improved with new technology, as this AR setup did not show statistical difference when compared to a low cost setup. On the other hand, these type of systems have the potential to become a helpful tool in telemedicine, just like the full telemedicine set-up, if we can make them more robust and lightweight.

5.3 Privacy

Most patients are willing to have their doctor use face-mounted wearable computers, even when unfamiliar with the technology [80]. However, some patients have expressed concerns about privacy, which can certainly be a concern when a camera is pointing directly at them [80]. In this research, we serialized hand and audio data prior to network transmission. Compression and encryption can also be added into the serialization process. Furthermore, MRC is protected by a username and password combination. This setup provides a basic protection to the patient's privacy. However, all the data is transmitted through the Internet, which may make it vulnerable to hackers. The privacy regarding recording is also another concern when a videoconference is established.

5.4 Limitations

There were many limitations to this pilot study. First of all, the experiment was not under entirely real circumstances, as the connection was established in a local area network. The reason for using a local network was to provide a dedicated bandwidth and not rely on the variability of the university local area network, which was important to support the high bandwidth requirements of the application. Another limitation would be the technical problems that happened in the testing environment. Next, every participant reported different levels of discomfort with the HoloLens, which negatively impacted the experience. Also, only one mentor was involved in the study, so the mentor gradually familiarized himself with the whole setup, which may have caused an increasing trend in performance across trials due to the learning effect. Finally, the results could be biased due to a low sample size. Time and budget limitations forced us to have a small study. Future studies could measure performance using only one assessment which might save a substantial amount of time.

During the study, the mentor was able to indicate desired hand position through the Leap Motion sensing technology. After 5 participants used the system, however, Leap Motion appeared to be working improperly. It was unable to recognize and provide the orientation of the hand correctly. It is still unknown why this occurred. However, when we unplugged the Leap Motion sensor for a while, the problem can be solved. The study was then paused until the Leap Motion was working correctly again. For the next study, we plan to have multiple Leap Motion sensors to avoid this issue.

Most participants also found it difficult to locate the holograms (3D models). We put hand models at a fixed position and enabled the mentors to reset it to the trainee's current field of vision remotely. The trainee could also reset it by voice command. However, when a participant could not find the model, often times the participant would move their head rapidly in order to locate the model. This behaviour made the reset task even more difficult for the mentors. Additionally, the audio data was not streamed from the mentor to the HoloLens. Normally, a network connection will be created between two sides of network users, and network data will be sent byte by byte quickly. This is network streaming, which is considered a good way to transfer data. However, in our system, the audio was sent progressively after a short period. This may have required more bandwidth and lead to a higher latency. We believe that the latency should be considerably improved if we create network streaming with better protocol and hardware environment. Microsoft just released a project called Mixed Remote View Compositor, which provides the ability to incorporate near real-time viewing of mixed reality view. It is achieved through low level Media Foundation components, which tends to resolve the MRC latency problem with Dynamic Adaptive Streaming over HTTP (DASH), as discussed in the Appendix C.

5.5 Future Work

It is clear that the mentors' assessment of the technology was worse. The reason why this would happen is vital for making improvements in the future work. The study was paused frequently due to technical issues, which can be considered as the first negative factor to the mentor. Some students were not familiar with the HoloLens and the mentor should sometimes acted as a technician to provide them with guidance. This additional task would also be another key to the problem. Finally, the mentor was accustomed to traditional user interfaces such as buttons and keyboards rather than new ones such as gesture and speech. All these factors should be evaluated and covered in the future version.

In the user study, we noticed that the quality of the connection, in particular, the latency, was the key reason for poor performance. The latency came from two sides.

First, the audio data was progressively transferred together with the hand data from the mentor to the HoloLens instead of streaming. We believe that the latency should be considerably improved if we create a network streaming with better protocol and hardware environment. Microsoft released a Sharing server in their HoloToolkit project on Github.com. It allows applications to span multiple devices, and enables collaboration between remote users. The server runs on various platform, and can work with any programming language.

Second, the built-in Mixed Reality Capture (MRC) function is achieved by HTTP progressive download. The mixed reality view is continuously being recorded for a short period of time into a series of video files, and then exposed on the built-in web server (also known as the Device Portal) on the HoloLens. After that, other applications can then access the web server, download and play the recorded serial video files progressively. This method is suitable for live broadcast applications, but inappropriate for an application with instant communication requirements.

With the help of these projects, we improved our system, redesigned the whole networking connections, and reduced the latency from 2-3 seconds to less than 500 millisecond. The bandwidth requirement for this design is also reduced to 4 Mbps, which suggests the possibility to run this system under the LTE network.

The way to present the hand model is also changed. The hologram with a hand model will now be presenting right in the middle of the users' vision. Together with the latency, this improved version greatly changes the user experience. Figure 5.1 shows the pipeline of the improved system. To evaluate the effect of these improvements, a new user study will be performed to evaluate the performance of users under the improved system.

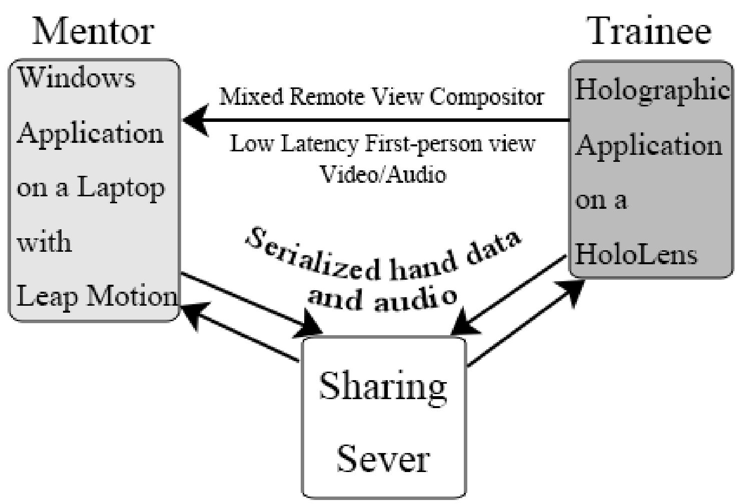

Figure 5.1: Overview of the improved system pipeline

Chapter 6

Conclusions

We have presented the design and implementation of an ultrasound telementoring application using the Microsoft HoloLens. Compared to available telementoring applications that mostly include visual and auditory instructions, the system introduced here is more immersive as it presents a controlled hand model with an attached ultrasound transducer. Compared to other gesture based AR system, our system is easier to setup and run. The trainee is wearing an AR headset and following voice instructions together with the mentor's transported hand gestures. The pilot user study with 12 inexperienced sonographers (medical school students) demonstrated that this could become an alternative system to perform ultrasound training. However, the HoloLens still needs to be improved, as every participant reported different levels of physical discomfort during the study, and an assistant must ensure the device is properly worn. In addition, the global rating scale used by an expert evaluator suggests that the trainees' performance is slightly worse using the HoloLens compared to the standard telemedicine setup with no significant statistical difference. Furthermore, the response time for the HoloLens application is longer than the other setup. Finally, the single mentor reported that the task became harder when using the HoloLens. A new system with significant improvements has the potential to be a feasible telemedicine tool, and we plan to evaluate this with a full user study in the near future. Part of this work has been submitted to journal Sensors (`http://www.mdpi.com/journal/sensors`). Other applications which could be studied in future research include other training systems and exploratory adventures in uncharted territories, such as creating an interactive social network application on the HoloLens.

Main Contributions of this Research

There are several components involved in this research, exploring the possibilities in different directions. The main contributions of this research are shown below:

- We have developed one of the first telemedicine mentoring systems using the Microsoft Hololens. We then demonstrated its viability and evaluated its suitability in practical use through a user study.
- We have tested various techniques and ported together inside the HoloLens, including: overlaying the holograms; controlling the hologram using a smart phone; implementing a videoconference with minimal latency; projecting Leap Motion recognized gesture inside the HoloLens. All these attempts are meaningful and useful for HoloLens-related developers due to its novelty.
- We have found that the performance of a telecommunication application is not

always improved with new technology, as the AR setup using the Hololens and Leap Motion did not show significant statistical difference when compared to a low cost setup.

- Until August 2017, the documentation about HoloLens development is still scarce. When you plan to develop a new application under the HoloLens, lack of support is a primary problem right now. We have provided a large amount of support material to follow up on this work, which should be considered as a valuable asset for researchers.

Above all, the most difficult part of this research would be the hand-shape hologram control part. We should gather the recognized hand data from the Leap Motion controller, serialize and transfer it to the HoloLens side, and then interpret the received serialized data into a hand-shape hologram. All of these work were done without enough documentation. After that, emerging this part together with Microsoft's github projects was also important for finally completing this work, as those github repositories were still under development.

Figure 6.1: A novel augmented reality telemedicine platform involving real-time remote pointing and gesture capture (Professor Andrew Smith and Professor Michael Parsons). For video example and further information, please see Appendix A.

Bibliography

[1] S. Laurent. *Rural Canada: Access to health care*. Parliamentary Research Branch Ottawa, 2002.

[2] G. Riva and K. Wiederhold. The new dawn of virtual reality in health care: medical simulation and experiential interface. *Annual Review of Cybertherapy and Telemedicine*, page 3, 2015.

[3] M. Rouse. Telepresence room. `http://whatis.techtarget.com/definition/telepresence-room`. Accessed: 2017-08-29.

[4] I. Mendez, M. Jong, D. Keays-White, and G. Turner. The use of remote presence for health care delivery in a northern Inuit community: a feasibility study. *International journal of circumpolar health*, 72, 2013.

[5] D. Fishman. Telemedicine: Bringing the specialist to the patient. *Nursing Management*, 28(7):30, 1997.

[6] M. Zundel. Telemedicine: history, applications, and impact on librarianship. *Bulletin of the Medical Library Association*, 84(1):71, 1996.

[7] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, 1992.

[8] T. Azuma. A survey of augmented reality. *Presence: Teleoperators and virtual environments*, 6(4):355–385, 1997.

[9] I. Sutherland. The ultimate display. *Multimedia: From Wagner to virtual reality*, 1965.

[10] F. Jolesz. Image-guided procedures and the operating room of the future, 1996.

[11] M. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *ACM SIGGRAPH Computer Graphics*, volume 26, pages 203–210. ACM, 1992.

[12] R. Ohbuchi, M. Bajura, and H. Fuchs. Case study: Observing a volume rendered fetus within a pregnant patient. In *Visualization: Proceedings of the IEEE Conference on Visualization*, volume 5. IEEE Computer Society Press, 1998.

[13] G. Dussault and M. Franceschini. Not enough there, too many here: understanding geographical imbalances in the distribution of the health workforce. *Human resources for health*, 4(1):12, 2006.

[14] C. Kinsley. *Rural Health in Rural Hands: Strategic Directions for Rural, Remote, Northern and Aboriginal Communities.* Health Canada, 2002.

[15] Statistics Canada. 2011 Community Profiles. In: 2011 census. *Statistics Canada Catalogue*, (98-316-XWE), 2011.

[16] P. Aarnio, H. Rudenberg, M. Ellonen, and P. Jaatinen. User satisfaction with teleconsultations for surgery. *Journal of Telemedicine and Telecare*, 6(4):237–241, 2000.

[17] J. Moehr, J. Schaafsma, C. Anglin, S. Pantazi, N. Grimm, and S. Anglin. Success factors for telehealth: a case study. *International journal of medical informatics*, 75(10):755–763, 2006.

[18] S. Doze, J. Simpson, D. Hailey, and P. Jacobs. Evaluation of a telepsychiatry pilot project. *Journal of Telemedicine and Telecare*, 5(1):38–46, 1999.

[19] A. Shapiro and G. McDonald. I'm not a real doctor, but I play one in virtual reality: implications of virtual reality for judgments about reality. *Journal of Communication*, 42(4):94–114, 1992.

[20] K. Kaltenborn and O. Rienhoff. Virtual reality in medicine. *Methods of Information in Medicine*, 32(5):407–417, 1993.

[21] N. Statt. Facebook has Oculus, Google has cardboard, 2014.

[22] A. Taylor. What is the Microsoft HoloLens? In *Develop Microsoft HoloLens Apps Now*, pages 3–7. Springer, 2016.

[23] Ricoh Imagine Change. New ricoh theta model, capturing 360-degree images in one shot, is on sale soonspherical video function, api and sdk (beta version). *News Release*, 2014.

[24] Microsoft. Satya nadella, terry myerson, joe belfiore and phil spencer: Windows 10 briefing. `https://news.microsoft.com/speeches/satya-nadella-ter`

`ry-myerson-joe-belfiore-and-phil-spencer-windows-10-briefing`, 2015. Accessed: 2017-05-13.

[25] A.. Desai, L. Peña-Castillo, and O. Meruvia-Pastor. A window to your smartphone: Exploring interaction and communication in immersive vr with augmented virtuality. In *Computer and Robot Vision (CRV), 2017 Conference on*. IEEE, 2017.

[26] Y. Zhang and O. Meruvia-Pastor. Operating virtual panels with hand gestures in immersive vr games - experiences with the leap motion controller. In *Salento AVR 2017, Augmented and Virtual Reality, 4th International Conference on*. Springer, 2017.

[27] R. Elford. Telemedicine activities at Memorial University of Newfoundland: A historical review, 1975–1997. *Telemedicine Journal*, 4(3):207–224, 1998.

[28] R. Hoppmann, V. Rao, M. Poston, D. Howe, P. Hunt, S. Fowler, L. Paulman, J. Wells, N. Richeson, P. Catalana, et al. An integrated ultrasound curriculum (iusc) for medical students: 4-year experience. *Critical Ultrasound Journal*, 3(1):1, 2011.

[29] A. Smith and T. Renouf. Reflections on the development of a multidisciplinary point of care ultrasound program: Idealism vs reality. *Internal Medicine*, 8:15, 2013.

[30] V. Hurst, S. Peterson, K. Garcia, D. Ebert, D. Ham, D. Amponsah, and S. Dulchavsky. Concept of operations evaluation for using remote-guidance ultra-

sound for exploration spaceflight. *Aerospace medicine and human performance*, 86(12):1034–1038, 2015.

[31] N. Kolbe, K. Killu, V. Coba, L. Neri, K. M Garcia, M. McCulloch, A. Spreafico, and S. Dulchavsky. Point of care ultrasound (pocus) telemedicine project in rural nicaragua and its impact on patient management. *Journal of ultrasound*, 18(2):179–185, 2015.

[32] D. Elford and A. House. Telemedicine experience in Canada (1956–1996). In *Medicine 2001 Conference, Montreal*, 1996.

[33] J. Picot. Telemedicine and Telehealth in Canada: Forty years of change in the use of information and communications technologies in a publicly administered health care system. *Telemedicine Journal*, 4(3):199–205, 1998.

[34] R. ter Haar. Virtual reality in the military: present and future. In *3rd Twente student conference on IT*, 2005.

[35] Facebook News Room. Facebook to acquire Oculus. `http://newsroom.fb.com/news/2014/03/facebook-to-acquire-oculus/`, 2014. Accessed: 2016-09-27.

[36] E. Ackerman. Innovega delivers the wearable displays that science fiction promised. *IEEE Spectrum, January*, 9, 2014.

[37] M. Satava. Emerging medical applications of virtual reality: A surgeon's perspective. *Artificial Intelligence in Medicine*, 6(4):281–288, 1994.

[38] G. Riva. Medical clinical uses of virtual worlds, 2014.

[39] G. Riva, A. Dakanalis, and F. Mantovani. Leveraging psychology of virtual body for health and wellness. *The Handbook of the Psychology of Communication Technology, SS Sundar, ed., John Wiley & Sons, Ltd, Chichester, UK*, pages 528–547, 2015.

[40] G. Riva, S. Raspelli, D. Algeri, F. Pallavicini, A. Gorini, B. Wiederhold, and A. Gaggioli. Interreality in practice: Bridging virtual and real worlds in the treatment of posttraumatic stress disorders. *Cyberpsychology, Behavior, and Social Networking*, 13, 2010.

[41] G. Riva, S. Raspelli, D. Algeri, F. Pallavicini, A. Grassi, B. Wiederhold, and A. Gaggioli. Interreality in the management of psychological stress: A clinical scenario. *Studies in health technology and informatics*, 154, 2010.

[42] S. Müller, L. Maier-Hein, A. Mehrabi, F. Pianka, U. Rietdorf, I. Wolf, L. Grenacher, G. Richter, C. Gutt, J. Schmidt, et al. Creation and establishment of a respiratory liver motion simulator for liver interventions. *Medical physics*, 34(12):4605–4608, 2007.

[43] T. Ogata, J. Onuki, K. Takahashi, and T. Fujimoto. The use of computer-assisted system in ear surgery. *Oto-Rhino-Laryngology Tokyo*, 48(6):47–51, 2005.

[44] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari, and V. Ferrari. Augmented reality in neurosurgery: a systematic review. *Neurosurgical review*, pages 1–12, 2016.

[45] R. Bly, M. Chang, S.and Cudejkova, J. Liu, and K. Moe. Computer-guided orbital reconstruction to improve outcomes. *JAMA facial plastic surgery*, 15(2):113–120, 2013.

[46] M. Qu, Y. Hou, Y. Xu, C. Shen, M. Zhu, L. Xie, H. Wang, Y. Zhang, and G. Chai. Precise positioning of an intraoral distractor using augmented reality in patients with hemifacial microsomia. *Journal of Cranio-Maxillofacial Surgery*, 43(1):106–112, 2015.

[47] G. Badiali, V. Ferrari, F. Cutolo, C. Freschi, D. Caramella, A. Bianchi, and C. Marchetti. Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning. *Journal of cranio-Maxillofacial surgery*, 42(8):1970–1976, 2014.

[48] J. Herron. Augmented reality in medical education and training. *Journal of Electronic Resources in Medical Libraries*, 13(2):51–55, 2016.

[49] C. Foronda, C. Alfes, P. Dev, A. Kleinheksel, D. Nelson Jr, J. O'Donnell, and J. Samosky. Virtually nursing: Emerging technologies in nursing education. *Nurse Educator*, 42(1):14–17, 2017.

[50] L. Alem, F. Tecchia, and W. Huang. Handsonvideo: Towards a gesture based mobile ar system for remote collaboration. In *Recent trends of mobile collaborative augmented reality systems*, pages 135–148. Springer, 2011.

[51] F. Tecchia, L. Alem, and W. Huang. 3d helping hands: a gesture based mr system for remote collaboration. In *Proceedings of the 11th ACM SIGGRAPH*

International Conference on Virtual-Reality Continuum and its Applications in Industry, pages 323–328. ACM, 2012.

[52] D. Anton, G. Kurillo, A. Yang, and R. Bajcsy. Augmented telemedicine platform for real-time remote medical consultation. In *International Conference on Multimedia Modeling*, pages 77–89. Springer, 2017.

[53] G. Kurillo, A. Yang, V. Shia, A. Bair, and R. Bajcsy. New emergency medicine paradigm via augmented telemedicine. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 502–511. Springer, 2016.

[54] M. Carbone, C. Freschi, S. Mascioli, V. Ferrari, and M. Ferrari. A wearable augmented reality platform for telemedicine. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 92–100. Springer, 2016.

[55] M. Carbone, C. Freschi, S. Mascioli, V. Ferrari, and M. Ferrari. A wearable augmented reality platform for telemedicine. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 92–100. Springer, 2016.

[56] M. Shenai, M. Dillavou, C. Shum, D. Ross, R. Tubbs, A. Shih, and B. Guthrie. Virtual interactive presence and augmented reality (vipar) for remote surgical assistance. *Operative Neurosurgery*, 68(suppl_1):ons200–ons207, 2011.

[57] S. Beck, A. Kunert, A. Kulik, and B. Froehlich. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):616–625, 2013.

[58] C. Zhang, Q. Cai, P. Chou, Z. Zhang, and R. Martin-Brualla. Viewport: A distributed, immersive teleconferencing system with infrared dot pattern. *IEEE Transactions on MultiMedia*, 20(1):17–27, 2013.

[59] Mark H. We found 7 critical HoloLens details that Microsoft hid inside its developer docs. `http://www.pcworld.com/article/3039822/consumer-electronics/we-found-7-critical-hololens-details-that-microsoft-hid-inside-its-developer-docs.html`, 2016. Accessed: 2017-06-04.

[60] Gabe M. Microsoft HoloLens may cause discomfort as it gets extremely hot. `https://infinityleap.com/microsoft-hololens-may-cause-discomfort-as-it-gets-extremely-hot/`, 2016. Accessed: 2017-06-04.

[61] R. Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[62] S. Izadi, D. Kim, O.r Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, et al. KinectFusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 559–568. ACM, 2011.

[63] R. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. In *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*, pages 127–136. IEEE, 2011.

[64] R. Rusu and S. Cousins. 3d is here: Point cloud library (pcl). In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 1–4. IEEE, 2011.

[65] J. Zhang. Microsoft HoloLens technology: the reduction of 3d scene and sdk. `http://prog3.com/article/1970-01-01/2823744`, 2015. Accessed: 2017-05-13.

[66] O. Muensterer, M. Lacher, C. Zoeller, M. Bronstein, and J. Kübler. Google Glass in pediatric surgery: an exploratory study. *International Journal of Surgery*, 12(4):281–289, 2014.

[67] M. Cicero, B. Walsh, Y. Solad, T. Whitfill, G. Paesano, K. Kim, C. Baum, and D. Cone. Do you see what i see? insights from using Google Glass for disaster telemedicine triage. *Prehospital and Disaster Medicine*, 30(01):4–8, 2015.

[68] H. Knight, P. Gajendragadkar, and A. Bokhari. Wearable technology: using Google Glass as a teaching tool. *BMJ case reports*, 2015, 2015.

[69] O. Moshtaghi, K. Kelley, W. Armstrong, Y. Ghavami, J. Gu, and H. Djalilian. Using Google Glass to solve communication and surgical education challenges in the operating room. *The Laryngoscope*, 125(10):2295–2297, 2015.

[70] A. Widmer and H. Müller. Using Google Glass to enhance pre-hospital care. *Swiss Medical Informatics*, 30, 2014.

[71] N. Cui, P. Kharel, and V. Gruev. Augmented reality with Microsoft HoloLens Holograms for near infrared fluorescence based image guided surgery. In *Proc.*

of SPIE (The International Society for Optical Engineering) Vol, volume 10049, pages 100490I–1, 2017.

[72] T. Pham and A. Tang. User-defined gestures for holographic medical analytics. In *Poster Proceedings of Graphics Interface*, 2017.

[73] A. De Lattre, J. Bilien, A. Daoud, C. Stenac, A. Cellerier, and J. Saman. Videolan streaming howto. *GNU General Public License*, 2005, 2002.

[74] A. Taylor. Develop microsoft hololens apps now. 2016.

[75] H. Black, G. Sheppard, B. Metcalfe, J. Stone-McLean, H. McCarthy, and A. Dubrowski. Expert facilitated development of an objective assessment tool for point-of-care ultrasound performance in undergraduate medical education. *Cureus*, 8(6), 2016.

[76] B. Pourazar and O. Meruvia-Pastor. A comprehensive framework for evaluation of stereo correspondence solutions in immersive augmented and virtual realities. *Journal of Virtual Reality and Broadcasting*, 13(2), 2016.

[77] CACM Staff. Address the consequences of AI in advance. *Communications of the ACM*, 60(3):10–11, 2017.

[78] M. Anvari, T. Broderick, H. Stein, T. Chapman, M. Ghodoussi, D. Birch, C. Mckinley, P. Trudeau, S. Dutta, and C. Goldsmith. The impact of latency on surgical precision and task completion during robotic-assisted remote telepresence surgery. *Computer Aided Surgery*, 10(2):93–99, 2005.

[79] E. Geelhoed, A. Parker, D. Williams, and M. Groen. Effects of latency on telepresence. *HP Labs Technical Report: HPL-2009-120*, 2009.

[80] M. Prochaska, V. Press, D. Meltzer, V. Arora, et al. Patient perceptions of wearable face-mounted computing technology and the effect on the doctor-patient relationship. *Applied Clinical Informatics*, 7(4):946–953, 2016.

[81] Microsoft. Mixed reality capture for developers. `https://developer.microsoft.com/en-us/windows/mixed-reality/mixed_reality_capture_for_developers`, 2015. Accessed: 2017-08-06.

[82] A. Aloman, A. Ispas, P. Ciotîrnae, R. Sanchez-Iborra, and M. Cano. Performance evaluation of video streaming using MPEG DASH, RTSP, and RTMP in mobile networks. In *IFIP Wireless and Mobile Networking Conference (WMNC), 2015 8th*, pages 144–151. IEEE, 2015.

Appendix A

Video and Development Related

- A video about this study is available online at `http://www.wsycarlos.com/teleholo_video.html`.

- To provide an overview of the lessons learned in this research, the advantages and disadvantages of the different prototypes attempted to reach our proposed solution are illustrated in Appendix B.

- Specific technical details about the video streaming solutions explored for the HoloLens are discussed in Appendix C.

- Source code of the whole project can be accessed via `https://bitbucket.org/wsycarlos/mrcleaphand`.

There are several projects and plugins involved in the implementation of our study. They played an important role in the implementation of this research:

- Mixed Remote View Compositor in HoloLensCompanionKit:

`https://github.com/Microsoft/HoloLensCompanionKit/tree/master/MixedRemoteViewCompositor`

- Sharing Sever in HoloToolkit:

  `https://github.com/Microsoft/HoloToolkit/tree/master/Sharing`

  `https://github.com/Microsoft/HoloToolkit-Unity`

- Leap Motion for Unity Development

  `https://developer.leapmotion.com/unity`

- AVPro Video plugin developed by RenderHeads

  `http://renderheads.com/product/avpro-video/`

If you still have any questions, please contact sw7164@mun.ca for detail information. Thank you.

Appendix B

Development of a Telemedicine Prototype Using the Hololens

B.1 Gyroscope-controlled Probe

In order to test the possibility that the HoloLens can be used in the field of remote ultrasound training, we developed an initial prototype simulating a virtual ultrasound transducer on the HoloLens with its orientation controlled by the gyroscope inside a mobile phone (Figure B.1). A hologram of an ultrasound transducer was projected within the trainee's field of view while the gyroscope was accessed in an Android phone. The orientation information of the phone was live-streamed to a local HoloLens. The orientation data enabled the hologram to be adjusted accordingly. The basic objective was to demonstrate that a mentor could represent a motion or gesture in the HoloLens AR space and provide user feedback.

This early-stage prototype was deployed on the HoloLens with 10 participants agreeing to do a pilot test of the application. The research protocol involving human subjects for this and other related trials was reviewed and approved by the Health Research Ethics Authority in St. John's, Newfoundland and Labrador. Each participant used the system for 5 minutes prior to providing general feedback. One participant indicated that "having virtual objects around the actual environment is so cool". Most people felt they were able to gain some additional information without extra effort. However one concern highlighted the challenges associated with how the trainees should actually hold the ultrasound probe. This resulted in the addition of a hand model to the virtual transducer. Other feedback highlighted the importance of two-way communications, ability to manipulate the probe in 3-D space (as opposed to simply roll, pitch, yaw), and the importance of capturing hand as well as probe motion.

Figure B.1: Virtual Probe controlled by the gyroscope located in the mobile phone. Remote drawing can also be achieved by drawing on the screen of the phone.

B.2 Video Conferencing

Feedback from the first prototype prompted us to consider a video conferencing application (Figure B.2) between the HoloLens and a desktop computer. Microsoft provides a built-in function called Mixed Reality Capture (MRC) for developers. The HoloLens can create an experience of mixing the real and digital worlds, with the MRC becoming a valuable tool to capture this experience from a first-person point of view. The lack of compatibility between the HoloLens and video streaming protocols is the chief obstacle of this video conferencing task. All immersive apps built for the HoloLens run on the Universal Windows Platform (UWP) and hence, are required to be built with the Unity Engine. Unity owns the “Asset Store” which contains many free and paid plugins. However, the HoloLens is a new product with limited access and no related video plugins available in the Asset Store yet. Finally, after several failed attempts (more detail in Appendix C), a plugin developed by RenderHeads called AVPro Video was located. AVPro Video provides powerful video playback solutions on various platforms including the HoloLens.

Our team created a video conference in the lab using a local area network and again sought user feedback. During this iteration, the participants had difficulty focusing on performing the ultrasound procedure with a video-feed streaming in their field of view. It was deemed uncomfortable to have both a video and a dynamic probe hologram simultaneously. Furthermore, the latency of the live-stream video, which could reach as high as 10 seconds, was unacceptable. On the other hand, attempting to use an MRC tool would cause the system to reduce the rendering to 30Hz, and would also cause the hologram content in the right eye of the device to “sparkle” [74, 81], which would be an undesirable artifact. The HoloLens was unable to maintain sufficient rendering quality with the MRC enabled, subsequently explaining the increased latency during a video decoding task. For this reason, the team concluded that video conferencing was not a suitable choice for the HoloLens and its Holographic Processing Unit (HPU) at the time.

Further details about video streaming on the HoloLens can be found in the Appendix C.

B.3 AR together with VR

Armed with new knowledge and experience learned from previous attempts, we hypothesized that the 3D appearance of the remote environment could be captured by the HoloLens and represented locally to an expert observer wearing an immersive VR HMD (the Oculus Rift). Such an immersive telepresence interface [57, 58] had the potential to enhance the mentors’ sense of presence. The MRC live video was broadcasted from the HoloLens’ first-person perspective through a built-in web server on the HoloLens. The Wowza Streaming Engine [82] was chosen to re-broadcast the MRC video using the Real Time Streaming Protocol (RTSP). MRC video was transferred and played to the mentor using the VR headset, recreating the first person view for the mentor in an attempt to provide a high sense of telepresence.

Figure B.3 shows a sample of the mixed reality view of a hologram (a dog) inside a VR headset.

Feedback from this system pertained mainly to the mentor’s experience. The expectation was that the mentor, wearing the VR headset, would be able to adjust

Figure B.2: Videoconferencing, with the HoloLens, was displayed on a floating mesh.

their view by moving their head. In this iteration, the VR headset simply acted as a monitor displaying the HoloLens' perspective. This resulted in the remote trainee being the individual controlling the expert's view while requiring additional guidance on where to look. Anecdotally, this appeared to increase the cognitive load on both the trainee and the mentor and occasionally triggered VR sickness symptoms due to the visually-induced perception of self-motion. Finally, the mentor indicated that he was significantly more comfortable with traditional input methods such as mouse, keyboard and touch rather than a new type of user interface such as voice and gesture.

Figure B.3: Displaying a mixed reality view of a hologram inside a VR headset.

Prior to the study, several streaming protocols had been used to implement prototypes. In the Appendix C, those protocols will be introduced.

Appendix C

Video Streaming on the HoloLens

Prior to the study, when we tried to implement a videoconferencing application on the HoloLens, several streaming protocols [82] were tested. All holographic applications built for the HoloLens run on the Universal Windows Platform (UWP). The Unity Engine was the default option for this task. The Unity Engine is mainly designed to develop games, and lacked adequate support related to video playback for the HoloLens. The "Asset Store" contains many plugins that can do this job. However, the HoloLens is a new product with limited access and most video plugins cannot work properly on the HoloLens. The compatibility of the HoloLens was becoming a problem for this videoconferencing task. We performed several experiments to find a best solution. Some details are listed below.

C.1 Web Real Time Communication (WebRTC)

Web Real Time Communication (WebRTC) is the latest protocol for real-time video streaming, but it is still an experimental project. It is fast, low latency and compatible with the newest browsers. The problem was how to integrate it into Unity. We tried different ways to make it work inside Unity, but failed. The first attempt was to use a WebGL build in Unity. WebGL is still a project in development, and is not compatible with the HoloLens. The performance of this implementation was not quite good. An empty scene with a cube required more than 20 seconds to load on a new iPhone, and at a frame rate no better than 15 frames per second (fps). The second thought was to embed a browser (web-view) inside a .Net application. However, it was hard to find a good framework. Several frameworks were written for web-view function, but either were not compatible with HTML5 elements or only worked on Unity Windows PC.

C.2 HTTP Live Streaming (HLS)

HTTP Live Streaming (HLS) was the first protocol that worked in our experiments. We could use either Open Broadcaster Software (OBS) or ffmpeg (video encoding/decoding library) to capture the stream for HLS. Nginx web server and nginx-rtmp-module with HLS support was used as the server. It is fast, smooth, high-resolution, compatible with almost every platform (works perfect on Windows, OSX, iOS, Android, HoloLens as well as browser). The problem was the prior delay for every stream, because it needs to slice the video into small parts. If the parts are too small, it would require more CPU usage, while if the parts are large, then the latency would become quite high.

C.3 Real-Time Messaging Protocol (RTMP) and Real Time Streaming Protocol (RTSP)

Real-Time Messaging Protocol (RTMP) and Real Time Streaming Protocol (RTSP) were the next two we tried. Open Broadcaster Software (OBS) or ffmpeg (video encoding/decoding library) were also acceptable for capturing the stream for RTMP and RTSP. Nginx web server and nginx-rtmp-module could also be used as the server. They were fast, smooth, high-resolution and low latency. The problem was the compatibility. For RTMP, it required a flash player on the client side, which made it inaccessible on many platforms, and hard to integrate into Unity. We had tried to use FluorineFx, a framework made RTMP compatible with Microsoft .Net. It was obsolete with a lot of errors, not very helpful. For RTSP, some plugins provided the support to play it under the Unity environment. However, most of them were not designed for the HoloLens.

We finally found a plugin developed by RenderHeads called AVPro Video, which provided powerful video playback solutions on various platforms including the HoloLens. HTTP Progessive Streaming, HLS and RTSP were all supported on the HoloLens. With RTSP protocol, the latency caused by the video streaming was eliminated. However, the latency caused by the Mixed Reality Capture was still significant in video conferencing.

C.4 Dynamic Streaming over HTTP (DASH)

Media Foundation is a video engine developed by Microsoft. It is designed to provide video playback support for the Universal Windows Platform (UWP). This is the native video support for the HoloLens. We could add playback of adaptive streaming multimedia content to UWP apps using Media Foundation. This feature currently supports playback of HTTP Live Streaming (HLS) and Dynamic Streaming over HTTP (DASH) content. For HLS, the latency is the problem. Therefore, DASH was chosen in the end. The Mixed Remote View Compositor project released by Microsoft is used in our upgraded version, which provides the ability to incorporate near real-time presenting of mixed reality view. It is achieved through low level Media Foundation components. The video feed is captured through the camera on the HoloLens, and then mixed with the virtual 3D objects. After that, the DASH protocol is used to stream the final mixed reality video. With all these efforts, a low latency mixed reality playback is finally achieved.