High-Resolution Stereoscopic Visualization of Pediatric Echocardiography Data on Microsoft HoloLens 2

by

Meruja Selvamanikkam

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Medical Sciences - Radiology and Diagnostic Imaging
University of Alberta

Abstract

Pediatric valve surgery typically requires cardiopulmonary bypass and cardioplegic arrest, which can have adverse effects on both post-operative recovery and long-term neurological function. However, non-bypass surgical techniques have limitations in pediatric patients, particularly due to challenges associated with the use of inserted devices. Balancing the necessity of bypass surgery with the unique considerations of pediatric cases remains a crucial aspect of pediatric valve surgery. Medical imaging, particularly Three-dimensional (3D) ultrasound (US), is crucial in providing volumetric images and detailed anatomical information for diagnosis and surgical planning. It is a key tool in the medical field to obtain a comprehensive view of the body. Ordinary two-dimensional displays do not provide depth perception and are not ideal for representing volumetric data, necessitating the use of more sophisticated visualization methods. Virtual and augmented reality (AR) displays can be used to improve the visualization of medical images, allowing for a more natural interaction with the environment. To address the issue of performing non-bypass surgery in children, a beating heart valve repair technique using augmented reality displays and three-dimensional echocardiography (3DE) is proposed to enhance visualization, offer natural interaction tools and voice commands, and facilitate the widespread use of 3D ultrasound in clinical settings. This study proposes custom software developed using the Unity3D platform to render high-resolution 3DE on the Microsoft HoloLens 2 using a phantom model with internal targets for papillary muscles and the atrioventricular leaflet-annular interface, providing an immersive AR experience for medical professionals. This research focuses on 3DE in children and uses a phantom

heart model to mimic a pulsating heart. The volume rendering algorithm utilizes the ray-marching technique, enabling direct volume rendering of high-quality volumetric models. To maintain a satisfactory frame rate, a Holographic Remoting approach is employed to reduce latency and enhance network transmission speed, utilizing the resources of a personal computer (PC). The custom software developed offers voice commands and an intuitive and interactive user interface that allows medical professionals to manipulate and explore 3DE images effectively. The interaction includes the ability to slice, modify the intensity range, and alter the voxel density. The experimental evaluations demonstrate that it is possible to produce a high-quality real-time display with HoloLens 2 and a PC-based remote rendering system, allowing intuitive control and exploration of 3DE. Overall, this research highlights the potential of AR rendering offered through Microsoft HoloLens 2 to advance pediatric 3DE rendering for medical professionals to enhance their decision-making and understanding of medical datasets.

Preface

This thesis was submitted as a partial fulfillment of the degree Master of Science (MSc) in Medical Sciences in Radiology and Diagnostic Imaging at the University of Alberta. The thesis is an original work by Meruja Selvamanikkam and the presented work was accomplished between January 2022 and December 2023. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board.

Material for this thesis is based on the following paper:

Meruja S, Noga M, Nee Scze K, and Punithakumar K. High-Resolution Stereoscopic Visualization of Pediatric Echocardiography Data on Microsoft HoloLens 2. IEEE Access. under review, 2023.

Acknowledgements

I would like to thank my supervisors, Dr. Kumaradevan Punithakumar and Dr. Michelle Noga, for giving me the opportunity to pursue a MSc in the Servier lab. From the first day, I felt extremely welcome working in their lab. They have provided me with so much technical and clinical guidance for the last couple of years. Moreover, I would like to thank Dr. Nee Khoo who was always readily available for meetings and discussions. It has been very enjoyable working in the lab over these past two years. The Servier lab would not be the same without its many members, and I would like to acknowledge my present colleagues who have made my MSc experience pleasant. To servier lab members, it was wonderful working with and getting to know each of them. I would like to especially thank Dr. Punithakumar and Dr. Nee Khoo who have graciously provided me with 3D echocardiography data and so much guidance over the last two years and many interesting lab-related discussions, especially about how the industry is focusing on medical images. My work would not be possible without their contributions. I will miss the lab lunches and parties all of us shared together and the many mandatory Monday Tim Hortons coffee breaks.

I extend my sincere thanks to my lab mates for supporting me throughout the journey. Lastly, I want to express my deep gratitude to my parents, siblings, and husband for their unwavering support during my MSc thesis. Their encouragement and assistance were invaluable, and I'm truly grateful for their contributions to my academic journey.

Table of Contents

List of Tables

List of Figures

Abbreviations

3D Three-dimensional.

3DE Three-dimensional echocardiography.

AR Augmented reality.

CT Computed tomography.

DICOM Digital imaging and communications in medicine.

DoF Degrees of freedom.

DVR Direct volume rendering.

FoV Field of view.

FPS Frame rate per second.

HI Holographic interface.

HPU Holographic processing unit.

JPEG Joint photographic expert group.

lerp Linear interpolation.

MIP Maximum intensity projection.

MR Mixed/Merged reality.

MRI Magnetic resonance imaging.

MRTK Mixed reality tool kit.

MV Mitral valve.

OOBE Out-of-box experience.

PC Personal computer.

PNG Portable network graphics.

SRGS Speech recognition grammar specification.

TIF Tagged image file.

TV Tricuspid valve.

UI User interface.

US Ultrasound.

UWP Universal windows platform.

VR Virtual reality.

XR Extended reality.

Chapter 1

Aid in Surgery with Medical Image Visualization

1.1 Overview

1.1.1 Medical imaging modalities

Medical imaging plays a crucial role in modern healthcare, significantly impacting the fields of diagnosis and treatment planning. It encompasses a diverse range of imaging modalities, each offering unique insights into the human body. Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and 3D ultrasound (US) stand out as crucial tools among the existing medical image modalities that enable healthcare professionals to delve deep into the intricacies of human anatomy and pathology. CT scans use X-rays to create detailed cross-sectional images of the body, aiding in the detection and diagnosis of fractures, tumors, vascular issues, and lung diseases. Their speed and precision make them invaluable in emergency medicine for swiftly and accurately assessing trauma-related injuries. MRI uses magnetic fields and radio waves to create highly detailed images of the body's internal structures, particularly excelling in visualizing soft tissues such as the brain, muscles, and joints. This non-invasive imaging technique is crucial for diagnosing neurological disorders, musculoskeletal conditions, and abdominal abnormalities, and is valuable for the long-term monitoring of chronic diseases and studying brain functions.

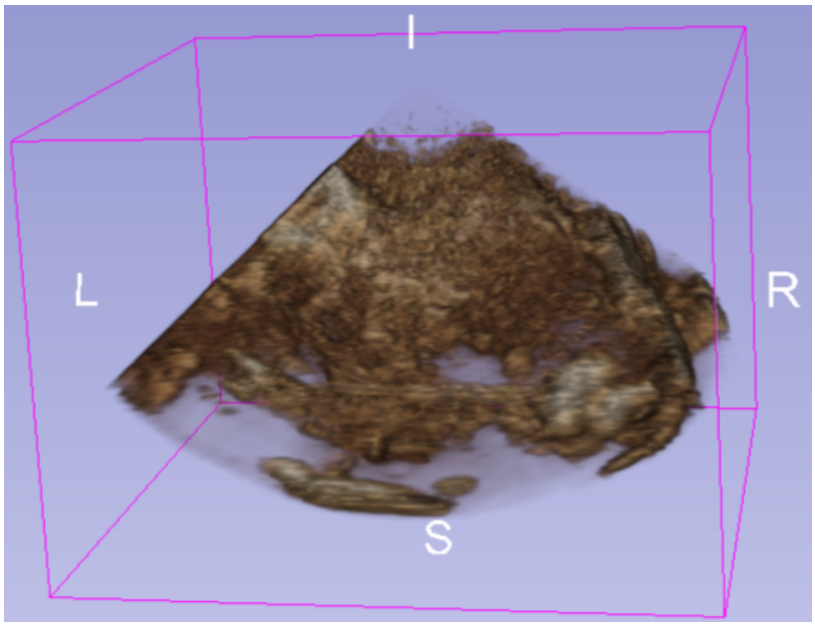

Figure 1.1: A volume-rendered 3D echocardiography scan of a heart phantom visualized in 3DSlicer.

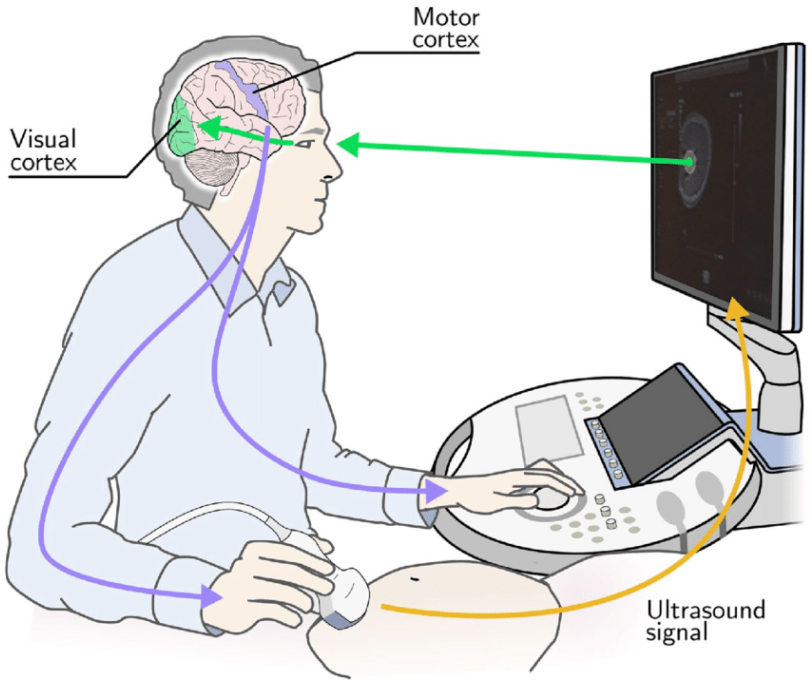

Figure 1.2: Sonographer loop [1]: gaze at the screen, adjust the transducer, alter the displayed image, and repeat while maintaining focus.

Using high-frequency sound waves (2 to 18 megahertz), US imaging produces images of interior organs and is affordable, non-invasive, portable, accessible, and real-

time capable. [2]. In pediatric medicine, where minimizing radiation exposure is of utmost importance, MRI and ultrasound are favored for imaging children due to their safety, making the US, which is inexpensive, a preferred choice for medical diagnosis and evaluation [3]. Among the various types, three-dimensional (3D) US, also known as volumetric US which is illustrated in Fig. 1.1, offers detailed anatomical information and visualization of volumes using orthogonal planes [2]. Nowadays, both 2D and 3D US are commonly employed in clinical settings. This project relies on echocardiography images, a specialized form of ultrasound used to create 3D US images of the heart and its structures that are used to evaluate the structure and function of the heart. It is crucial in diagnosing and monitoring heart conditions such as heart disease, valve disorders, and congenital heart defects. Echocardiography is a vital tool in cardiology, as it allows cardiologists to non-invasively assess the heart's condition and function. It is especially useful for imaging cardiac movement in real-time.

At the core of US devices lies the US transducer, responsible for transmitting sound waves and receiving echoes, which are then converted into image data [4]. Within the realm of US imaging, achieving precise diagnoses and effective treatments relies on a high level of proficiency and remarkable hand-eye coordination, ensuring the successful acquisition of quality ultrasound images [5]. Fig. 1.2 depicts the sonographer dynamically adjusting the transducer position to locate a standard plane, receiving real-time visual feedback on the display screen, and consequently changing the displayed ultrasound imaging.

1.1.2 Pediatric valve surgery

Fig. 1.3 and Fig. 1.4 depict the Mitral (MV) and tricuspid valve (TV) regurgitation, respectively. MV or TV occurs with acquired or congenital heart disease in children, and many require surgery.

Pediatric valve surgery traditionally involves cardiopulmonary bypass and stopping the heart, which can lead to negative effects on post-operative recovery, myocardial

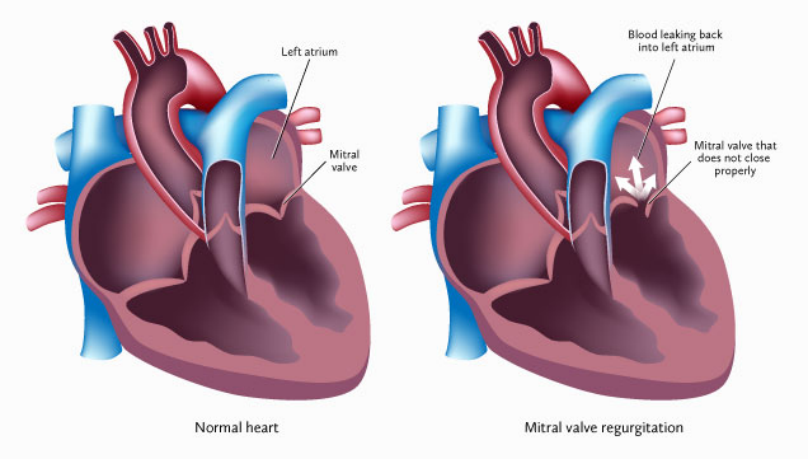

Figure 1.3: Mitral valve regurgitation is a valvular disease in which the mitral leaflets, which are comprised of two tissue flaps, do not seal properly [6].

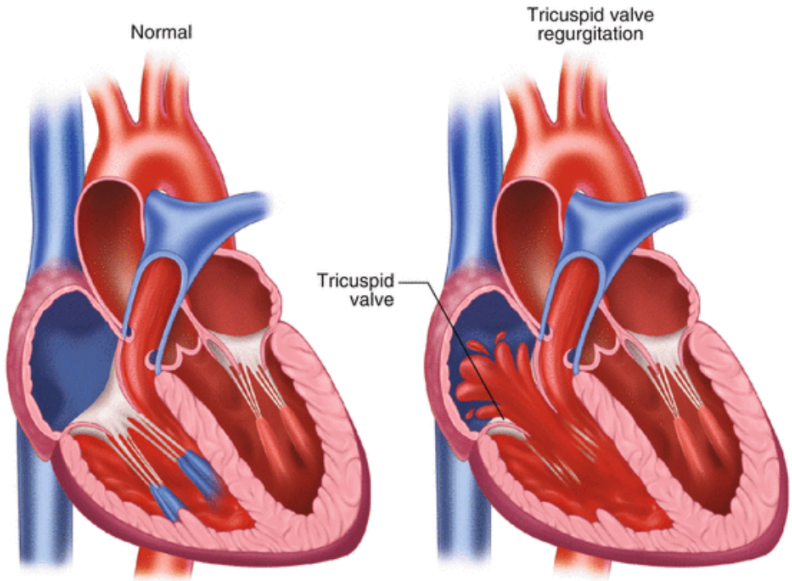

Figure 1.4: Tricuspid valve regurgitation is leakage of blood backward through the tricuspid valve each time the right ventricle contracts [7].

depression, and lasting neurological impact, especially in young patients. While newer techniques have reduced these impacts in adult valve surgery, they are not easily applicable to children due to the size of current instruments and limitations in long-term durability. This often leads to patient-prosthesis mismatch and multiple future surgeries to upsize the prosthesis.

Currently, there are no non-bypass surgical or percutaneous valve repair options for pediatric patients. However, the concept of non-bypass extracardiac surgery has been explored in adult procedures guided by echocardiography and manipulation of papillary muscles in a beating heart. These approaches avoid cardiopulmonary bypass's harmful effects and provide real-time feedback through live imaging of valve function, offering potential benefits for pediatric patients.

1.1.3 Volume visualization

Volume visualization, an essential technique in interactive graphics and imaging, enables the extraction of meaningful information from volumetric data, such as data representation, modeling, manipulation, and rendering. Volume rendering encompasses a set of techniques used to create a 2D projection of a 3D discretely sampled data set. This process enables the visualization of volumetric data, providing valuable insights and representations of complex structures and relationships in a three-dimensional space [8]. The 3DE voxel data obtained from a US scan is composed of individual data points that signify the density at a particular location. The information is read from a file and stored in a 3D texture. Although current 2D screen displays have limitations in conveying the true 3D nature of data, recent advancements in VR and AR displays offer better depth perception and natural interaction tools. In AR, virtual objects are overlayed on the real-world environment [9], unlike in VR where a purely artificial environment is generated with or without full user immersion.

A device is needed to display virtual objects and receive information about the real world to experience an AR application. Ray marching is a fundamental technique for

direct volume rendering, allowing the rendering of high-quality volumetric models on HoloLens 2 [10].

1.2 Thesis objectives

Create custom software for visualizing animated 3D echocardiography data in HoloLens 2, with the goal of establishing its potential utility for conducting surgical procedures in pediatric cases. The overarching objective of this thesis is to contribute to the advancement of medical imaging and intervention by developing innovative software for rendering stereoscopic 3DE images. Additionally, the research encompasses the design and testing of materials for creating accurate phantom heart models, achieving temporal synchronization with a dynamic *beating* heart phantom, conducting quantitative analyses of interventions using 3DE, and exploring the efficacy of AR-assisted extracardiac interventions in a simulated cardiac environment.

1.3 Thesis contributions

- **Developed software to render 3DE stereoscopic images:** The custom software was developed using the Unity3D platform to provide an immersive AR experience for medical professionals.
- **User interface/Voice input components:** User interface and voice command components for cardiologists to adjust the color transfer functions, such as intensity value change and the visibility window for intensity.
- **Enable spatial orientation to register 3DE in physical space:** Using object manipulation and manual spatial orientation, using hand interaction to support translation, rotation, and scaling of rendered objects
- **Quantitative analysis of rendering in HoloLens 2:** Due to the large 3DE image, there is a limitation to using holographic rendering. Therefore, holo-

graphic remoting was used. However, it is necessary to explore the limitations of holographic rendering in different aspects, which are unaddressed in many studies.

- **Temporal synchronization with *beating* heart phantom:** The visualization of the beating heart phantom model was achieved by importing all the raw files in the sequence. Then, a dataset is periodically selected from the array with a wait time based on the frame rate.

- **Design and testing of materials for phantom heart models with internal targets:** Fabricate a phantom heart model with internal targets for papillary muscles and the atrioventricular leaflet-annular interface using 3D printing. At the pre-defined position of the printed phantom, HoloLens 2, and PC, the volume is rendered in real-time to align properly with the physical model.

1.4 Thesis outline

This thesis includes six chapters:

- Chapter 2 reviews previous works in AR for surgery: It provides the literature on pediatric valve surgery, the basics of volume rendering, statistical analysis of usage of HoloLens 2, ultrasound volume rendering, and cardiac 3DE + t volume rendering.

- Chapter 3 presents the hardware options available for medical image visualization, explores the virtues and limitations of VR and AR, and highlights HoloLens 2.

- Chapter 4 proposes direct volume rendering and optimizing volumetric rendering. It also provides the details of data processing, algorithms, rendering approaches, environmental setup, and implementation.

- Chapter 5 describes the data that the study used and the results for the quantitative analysis of rendering in HoloLens 2 for different scenarios.
- Chapter 6 concludes the thesis with a summary and presents a direction for future work.

Chapter 2

AR for Surgery

2.1 Clinical review

Pediatric valve surgery using cardiopulmonary bypass can have negative effects on post-operative recovery, myocardial depression, and neurological impact [11–13]. Percutaneous techniques are not easily translated to children [14], highlighting the need for innovative tools for AR-assisted non-bypass heart valve surgery in children [15]. Valve replacement is avoided in children due to patient-prosthesis mismatch and the need for multiple future surgeries to upsize. There are no non-bypass surgical or percutaneous valve repair interventions for pediatric patients [16]. AR provides a stereoscopic display of 3DE images of intracardiac structures superimposed on the *beating* heart, improving accuracy [11]. Wearable stereoscopic displays enable AR exploration in surgical fields [17, 18]. AR's advantage of stereoscopic display allows a better appreciation of cardiac anatomy, even among experts [19]. The importance of adequate visualization for instrument guidance is reported in extracardiac interventions [20]. 3DE provides 3D spatial orientation of instruments but requires viewing on a non-stereoscopic screen [21]. The feasibility of extracardiac valve repair in adults has been reported [22–24]. Exploration in children is worthwhile due to the limitations of adult techniques and devices. Valve replacement is avoided in children due to future surgeries to upsize and patient-prosthesis mismatches. No non-bypass interventions exist for pediatric patients.

2.2 HoloLens 2

Augmented reality is increasingly being applied to applications in the medical field, particularly for diagnostic purposes, as numerous studies explore its potential. Prior research in the domain of augmented reality has paved the way for our current exploration. Rezende et al. [7] provided the basic idea to start the procedure as it analyzed volume rendering in HoloLens 1, which led us to work with HoloLens 2, a stand-alone device that does not require additional equipment, which is shown in Fig. 2.1. The HoloLens 2 is a head-mounted display where the holograms are generated by projecting images onto its lenses, creating an augmented reality environment with a display resolution of 2048×1080 pixels per eye. It boasts specific hardware specifications, including 64GB of onboard memory and 4GB of RAM, along with the 2nd Generation Custom-built Holographic Processing Unit (HPU).

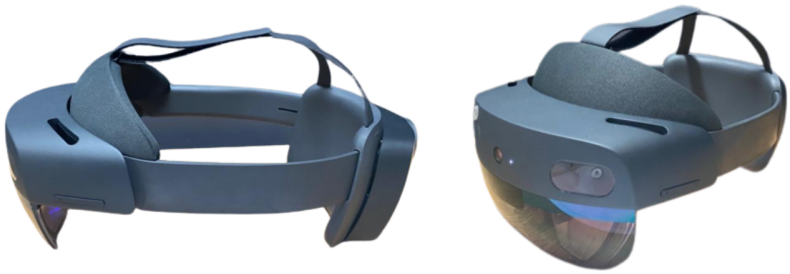

Figure 2.1: An up-close view of Microsoft's advanced mixed reality headset, the HoloLens 2.

2.3 HoloLens 2 usage statistics

Palumbo et al. [25] provide a state-of-the-art overview of the applications of Microsoft HoloLens 2 in the medical and healthcare context. The review thoroughly examines various studies conducted since 2019, focusing on HoloLens 2 sub-field applications, device functionality, and software used. Park et al. [26] summarize the results of 44 papers on the applications of HoloLens in several industries, including medical and

surgical aids and systems, medical education and simulation, industrial engineering, architecture, civil engineering, and other engineering fields. The review concluded that among the current applications, HoloLens was the most commonly applied in medical and surgical aids and systems, accounting for 43% of the existing applications [26]. Furthermore, [26] highlights that the utilization of Microsoft HoloLens 2 applications for medical auxiliary devices and systems remains relatively low as of 2021.

2.4 Medical volume rendering

Mojica et al. [27] introduced a holographic interface (HI) capable of visualizing 3D MRI data for planning neurosurgical procedures. The HI immersed users in a comprehensive MR scene, integrating actual MRI data and virtual renderings, enabling interactive interactions with objects within the MR environment. Currently, there is a substantial amount of ongoing research focused on implementing visualization tools for medical purposes. However, most articles [27, 28] rely on CT or MRI scans due to their clarity and ease of identifying targets. These modalities require fewer image processing techniques such as filtering, edge smoothing, and segmentation [29]. Conversely, in the case of the US, its greater computational complexity has limited its usage for rendering on devices such as the HoloLens 2. Although some researchers have used the US [5, 28–33], it comes with certain limitations. Costa et al. [30] explained that their application involves slicing a 3D breast volume template [29] with the input orientation. The resulting 2D US image is then transmitted back to the HoloLens 2 at a rate of 35 frames per second. This implementation builds on the HoloUS [31] application, which transmitted 2D images to HoloLens 1 at 25 frames per second. Another study [32] used 3D ultrasound rendered on a custom-made hybrid optical/video see-through head-mounted display.

2.5 Cardiac 3DE + t volume rendering

Background analysis reveals a scarcity of research in the US with HoloLens 2 for cardiac images [33] or HoloLens 2 combined with cardiac MRI [32]. This situation prompts the undertaking of a study on framerates during the rendering of 3D and 3D animated cardiac US volumes using HoloLens 2. Recent work by Maddali et al. [33] indicates that the study has been applied to 3DE images using HoloLens 2. However, specific minor issues remain unaddressed or unexplored. For instance, the performance statistics of the holographic rendering approach for different resolutions were not thoroughly examined. The application incorporates ray tracing; however, studies show that ray marching [34] enables efficient rendering of 3D fractals and faster calculation of surface normals compared to traditional ray tracing. The authors have created an application that is specifically designed to work with the data from GE Vivid E95 scanners and utilizes the same transfer functions or color maps as the scanner. The approach proposed in this study offers the option to edit the color and opacity transfer functions through a user interface, which allows for customizing the rendering by the end user prior to displaying the images through HoloLens. This study also explores the performance statistics of the holographic rendering using the ray-marching technique and animates the heart phantom to mimic the actual heart function to give a more realistic 3DE rendering.

Chapter 3

Medical Image Visualization Hardware

3.1 Stereoscopic displays

Stereoscopic vision is the ability of the visual brain to see an object in three dimensions. This is accomplished by using two eyes to view an object. The eyes work together to create a three-dimensional image. Real-world objects are 3D objects, thus having depth. Nevertheless, a human's eye produces only a 2D image. The 2D image from the left eye and the 2D image from the right eye are slightly different when looking at the same object. This is due to the different positions of the left eye and right eye on the human's face. The human brain creates a sensation of 3D or depth by means of processing these two different 2D images. This process is called stereoscopic vision. Fig. 3.1 illustrates the slightly different views of the cameras for the left eye and the right eye. The HoloLens 2 utilizes advanced optical and display technologies to provide a stereoscopic view. There are two distinct waveguide displays on the device, one for each eye. By projecting light directly on the user's retina, these displays create the illusion of 3D objects.

Key elements that enable HoloLens 2's stereoscopic view include:

Dual Displays: HoloLens 2 employs two separate displays, one for each eye. This binocular setup mimics the way human vision works, with each eye receiving a slightly

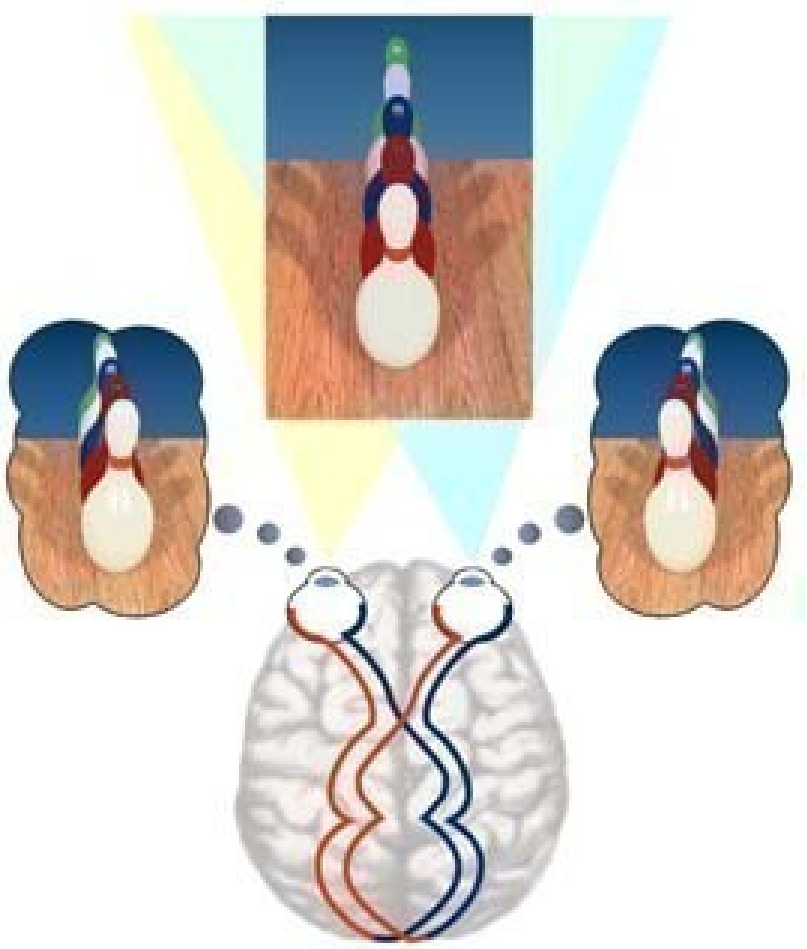

Figure 3.1: Illustrates the slightly different views of the cameras for the left eye and the right eye [35].

different image.

Optics: The waveguide optics in HoloLens 2 are responsible for directing the projected images from the displays into the user's eyes. These optics help in creating the perception of depth.

Eye Tracking: While HoloLens 2 primarily relies on the binocular display, it is also equipped with sensors for eye tracking. This tracking technology can detect where the user is looking, allowing for dynamic adjustments in real-time to maintain the illusion of depth.

By combining these components, HoloLens 2 offers a compelling 3D stereoscopic view, allowing users to interact with holograms and virtual objects seamlessly integrated into their environment. This technology enhances the overall mixed reality experience and makes it a powerful tool for applications such as remote collaboration, training, and visualization.

3.2 Virtuality continuum

The virtuality continuum represents the full spectrum of technological possibilities between the entirely physical world or real environment and the fully digital world or virtual environment [36]. The virtuality continuum is a theoretical framework shown in Fig. 3.2, which is derived from researchers Paul Milgram and Fumio Kishino's introduction on the virtuality continuum or reality-virtuality continuum concept in 1994 [37]. It helps to visualize and understand the differences between the various technologies that exist today and those that are yet to be invented. One end of the spectrum is low immersion, whereas the other end is high immersion. The technology is categorized based on the degree of immersion it provides. Fig. 3.3 represents the reality-virtuality continuum for the current technologies: AR, MR, and VR.

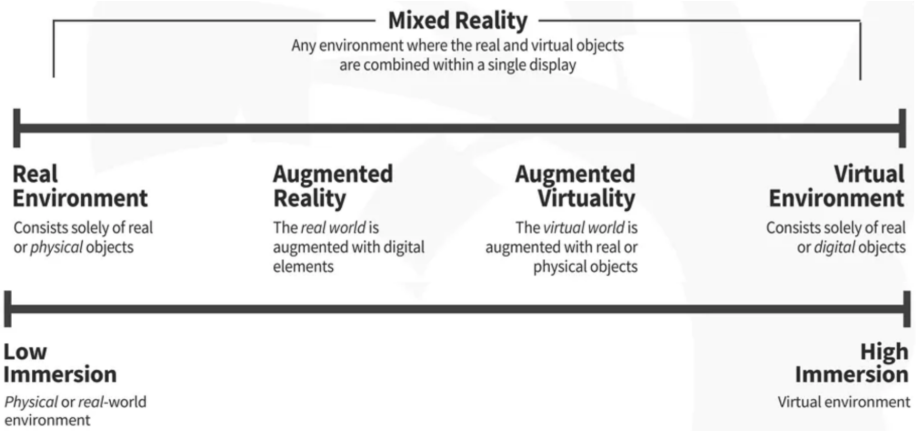

Figure 3.2: Representation of the virtuality continuum according to the spectrum of immersion [36, 37].

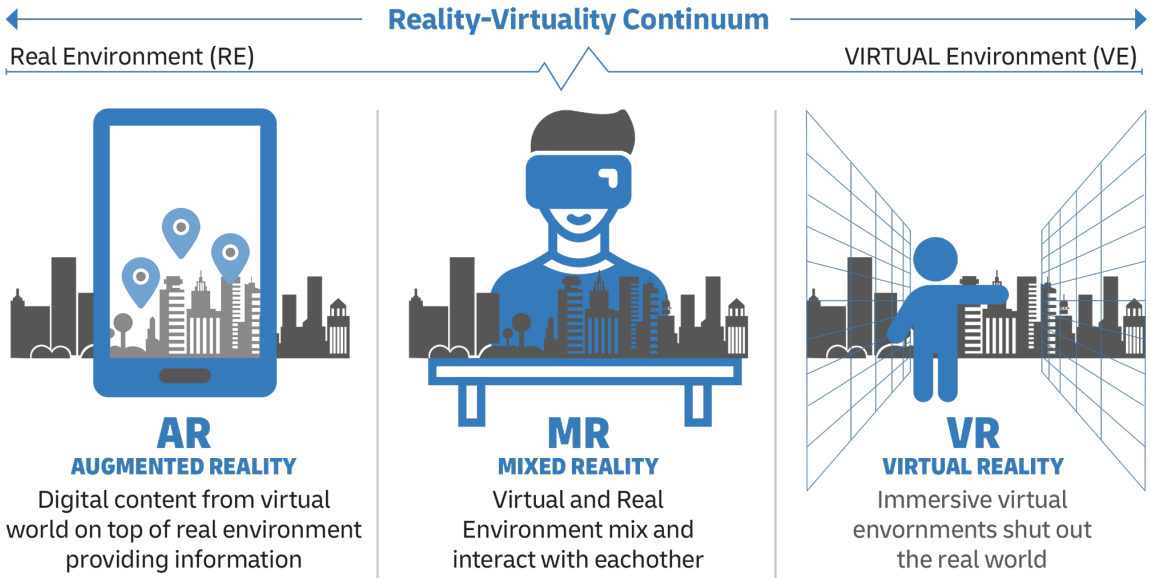

Figure 3.3: Representation of current XR technologies according to the reality-virtuality continuum [38, 39].

3.3 Extended reality (XR)

Extended reality (XR) is a universal term that combines any sort of technology that alters reality by adding digital elements to the physical or real-world environment to simulate the real world to any extent, blurring the line between the physical and the digital world. This includes virtual reality (VR), augmented reality (AR), mixed/merged reality (MR), and any technology—even those that have yet to be developed—situated at any point of the virtuality continuum. Any new technology that blends the physical and virtual worlds will also be categorized as XR. The "X" in XR stands for any variable—any letter of the alphabet—that may be used in the future for such technologies. Fig. 3.4, an illustration, vividly demonstrates the concept of XR, showcasing its immersive blend of VR, AR, and MR. Integrating XR into the project provides an immersive experience for the end-user in a multisensory environment that is more interactive, engaging, and effective long-term. XR training offers

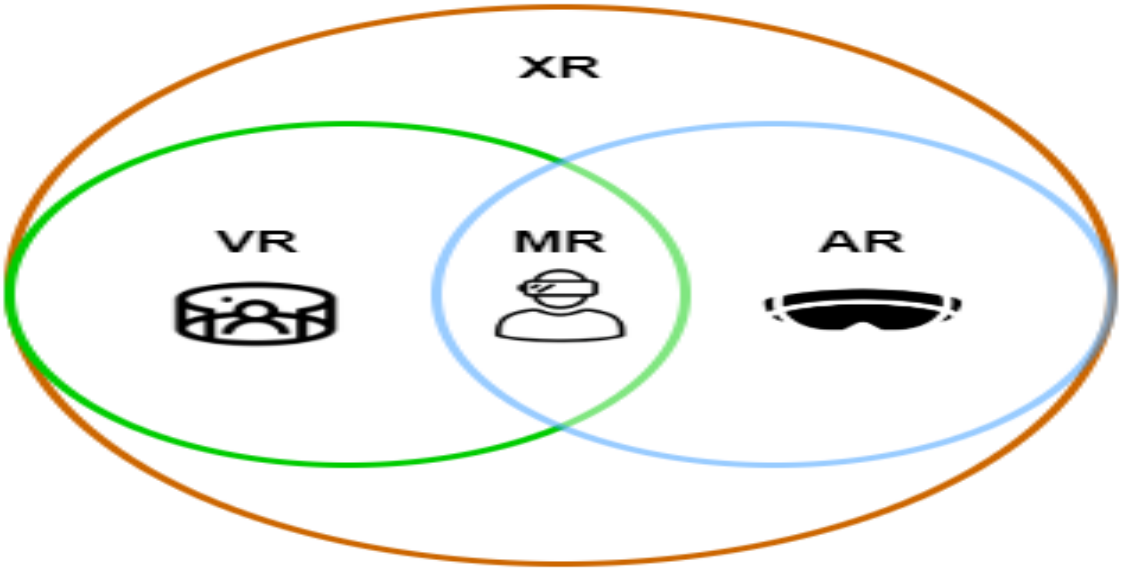

Figure 3.4: An illustration demonstrates the concept of XR, showcasing its immersive blend of VR, AR, and MR.

a secure environment for experiential learning, encouraging learners to take risks and learn from errors while also enabling the realistic development of both hard and soft

skills. Moreover, its scalability, coupled with the ability to collect essential training metrics through an XRS, ensures convenient deployment across organizations. On the other hand, XR training can involve high initial development and equipment costs, potential health risks, and the need for technology updates, which may pose a barrier for some organizations.

3.3.1 Augmented reality (AR)

Augmented reality is a technology that allows digital elements to be superimposed on real environments. In an AR experience, users can see a composite view of physical or real elements and digital elements. While some augmented reality experiences can provide some level of interaction between physical and virtual elements, direct interaction between elements of the physical and digital worlds is often overlooked and limited to non-existence. AR experiences are close to the physical world end of the virtuality continuum. The ability to overlay digital objects onto the physical world is revolutionizing many industries, such as video games, education, healthcare, construction, automobiles, and manufacturing.

3.3.2 Virtual reality (VR)

VR is a technology that enables the creation of a fully immersive digital environment. In a VR experience, the physical or realistic environment is completely blocked. VR experiences are at the all-virtual end of the virtual experience continuum. Many people have difficulty accepting the fact that virtual reality experiences produce real emotional responses, even though we know they are "fake." Humans are designed to construct reality from the input they receive from their senses. This leads to a sense of reacting to the complete virtual reality experience. In general, VR takes advantage of the visual and auditory systems. The end user will have a greater sense of presence and immersion if other senses are also added to the virtual environment.

3.3.3 Mixed reality (MR)

MR is also known as mixed, merged, or hybrid reality. It is a combination of the real world and the virtual world to create new environments. Physical and digital objects exist and interact with each other. MR is a technology that not only allows the superposition of digital elements in real environments but also allows interaction between them. In an MR experience, users can view and interact with both physical and digital elements. Therefore, MR experiments receive input from the environment and will change according to it. The MR experience spans the center of the virtual continuum.

3.4 Exploring the virtues and limitations of VR and AR

The integration of AR and VR in the healthcare sector brings forth a multitude of advantages. It serves as a valuable training tool for the next generation of medical professionals, enhancing their skills and knowledge base. Surgeons, too, reap the benefits of AR or VR, using cutting-edge technologies such as Microsoft Hololens to streamline surgical procedures and access patient data swiftly, crucial in time-sensitive situations. AR/VR also addresses common challenges in vein detection, making blood extraction less traumatic and easier for both patients and practitioners. Patients themselves benefit from AR/VR by gaining a deeper understanding of their medical conditions through immersive technologies, improving the quality of patient consultations. Additionally, they extend to enhancing the diagnostic process through 3D body scanning, augmenting conventional methods such as X-rays and CT scans. By generating precise 3D images that can be examined through AR/VR hardware, medical practitioners can expedite diagnosis and elevate patient care, ultimately revolutionizing the healthcare landscape.

The integration of AR and VR in healthcare presents a revolutionary frontier,

offering numerous advantages along with a few challenges. While the combined potential of AR and VR technologies is remarkable, there are hurdles to overcome in their widespread adoption. One major challenge is cost and accessibility, as AR and VR devices can be expensive, making their implementation at scale difficult, especially for smaller healthcare providers. Additionally, the security of patient data is a significant concern, demanding robust data protection measures to mitigate risks. Moreover, regulatory compliance and ethical considerations, such as ensuring patient autonomy and informed consent, are vital aspects. Furthermore, managing AR and VR devices within healthcare organizations requires tailored solutions due to specific charging, maintenance, and integration needs into existing healthcare systems and workflows. Overcoming these challenges is crucial to unlocking the full potential of AR and VR in healthcare.

3.5 Augmented reality hardware and software

The Microsoft HoloLens 2 was chosen as it is a stand-alone device that does not require additional equipment. HoloLens 2 utilizes a Microsoft HPU 2 with 64 GB of onboard memory and 4GB of random access memory. For the development environment, Unity3D v2020.3.33f1 was selected as the platform, along with Visual Studio 2019 and Mixed Reality Toolkit Foundation 2.3.0.

3.6 Highlights of HoloLens 2

Several AR devices are available on the market, with HoloLens 2 and Magic Leap 1 standing out as innovative augmented reality devices that provide cutting-edge experiences at the forefront of immersive technology. This thesis focuses on the development of the proposed method using HoloLens 2 due to the public development support available for the devices.

The HoloLens 2, developed by Microsoft, boasts an impressive set of features and

specifications. It provides a wider field of view (FoV), with a 52° diagonal rendered FoV compared to Magic Leap 1's 50°. Additionally, the HoloLens 2 offers a higher resolution of 1440x936 per eye, resulting in sharper and more detailed holographic displays. Its 6 Degrees of Freedom (6 DoF) tracking is achieved through four integrated cameras, ensuring precise and responsive interaction with holograms. Furthermore, the HoloLens 2 excels in terms of processing power with the Qualcomm Snapdragon 850 chipset, providing seamless performance. While both devices offer a similar battery life of around 3 hours, the HoloLens 2's superior specifications, wider field of view, and impressive tracking capabilities make it a more advanced and versatile choice for augmented reality applications, establishing it as a superior option in the standalone AR headset category.

3.7 Improve visual quality and comfort

3.7.1 Calibrating HoloLens 2

The use of eye-tracking technology, which improves the user's perception and engagement with the virtual environment, is a major component of HoloLens 2. This stage guarantees precise eye tracking for the gadget, which leads to improved comfort, precise hologram alignment, and precise hand tracking.

There are a few circumstances where the calibration process is required:

- **First-Time Use:** When unpacking the HoloLens 2 and initializing its use, the system prompts the user to calibrate the device to establish accurate eye-tracking parameters during the initial setup.
- **Opting Out:** If the user chooses to ignore the previous calibration prompt, the system will persistently request recalibration to ensure accurate performance.
- **Previous Calibration Failure:** Recalibration is recommended in the event of a calibration failure during the last usage of the device to rectify any tracking inaccuracies.

- **Deleted Calibration Profiles:** Recalibration becomes necessary if the user has deleted their calibration profiles or switched to a different user, ensuring precise tracking of eye movements.

- **Device Removal and Reapplication:** Following the removal and reapplication of the device, recalibration may be prompted if any of the aforementioned circumstances apply to maintain optimal performance.

A set of targets, or gems, is used during the calibration process to guide users. While blinking is acceptable, maintaining focus on the gems, rather than other objects in the room, is essential. This focus allows HoloLens to learn about the user's eye position, enabling the accurate rendering of holographic content in the virtual world. The outcome of successful calibration ensures that holograms appear correctly, even when the visor shifts on the user's head. This calibration process is pivotal in optimizing the HoloLens 2 experience for users.

3.7.2 Eye tracking

Eye tracking serves two key features: eye position tracking to optimize the viewing experience and eye gaze tracking, providing developers with gaze vectors for user input and interactions. Calibration can occur during the out-of-box experience (OOBE) or automatically when an uncalibrated user engages with an eye-tracking application.HoloLens 2 can store up to 50 calibration profiles, which facilitates device sharing among multiple users without requiring individual logins. Based on the user's previously calibrated visuals, the display automatically adjusts for quality and comfort upon reapplication.

Eye position tracking involves storing calibration information locally, correlated with iris bit codes from the calibration experience. This data is separate from login iris bit codes, ensuring automatic calibration retrieval for returning users without the need for recalibration. Calibration is not tied to a specific login account, ensuring

account agnosticism, and all calibration data can be deleted through the device's settings app.

3.7.3 Map physical spaces with HoloLens

HoloLens seamlessly integrates holograms into the user's physical environment by learning and memorizing the surroundings. Over time, the device constructs a spatial map of the environment, continually updating it as changes occur. The process occurs automatically as long as the user is logged in and the device is powered on. A device held or worn with the cameras pointing at an area enhances the mapping process. HoloLens automatically learns a space; however, users can optimize and expedite the mapping process.

For optimal results, it is recommended to choose a room with adequate light and space, avoiding dark areas and spaces with dark, shiny, or translucent surfaces. During mapping, a mesh graphic spreads across the space, visible in the mixed reality home. Spending more time in the areas where HoloLens will be used, ensuring smooth movements, and looking around to provide additional data all contribute to more efficient and accurate mapping. Additionally, traversing an area multiple times enhances the mapping process, helping to capture features that may have been missed during the initial walkthrough.

Chapter 4

Direct Volume Rendering & Optimization

4.1 Medical data processing

4.1.1 File formats

In the context of digital imaging, file formats serve as standards for storing image data. General applications often utilize Portable Network Graphics (PNG) and Joint Photographic Expert Group (JPEG) formats. HoloLens 2 supports a variety of image formats, including PNG, JPG, JPEG, Bitmap Image (BMP), and Tagged Image File (TIF), along with 3D models in formats such as FBX, GLB, glTF, STL, and PLY. Similarly, Unity is capable of reading standard 3D file formats such as .fbx, .dae, .dxf, and .obj. It is crucial to verify the compatibility of the model format with the development platform. Another critical consideration is the face count. Even if the 3D model type is supported on both Unity and HoloLens 2 development platforms and for rendering on head-mounted displays, it is essential to take into account the face count. Fig. 4.1 illustrates the FBX image of the original FBX file, decimated blend object, and optimized FBX file, showcasing their respective face counts. To ensure optimal performance on HoloLens 2, it is advisable to optimize 3D models as much as possible when focusing on surface rendering.

In the field of medical imaging, datasets typically encompass several image types: those representing anatomical volume projections onto an image plane, images de-

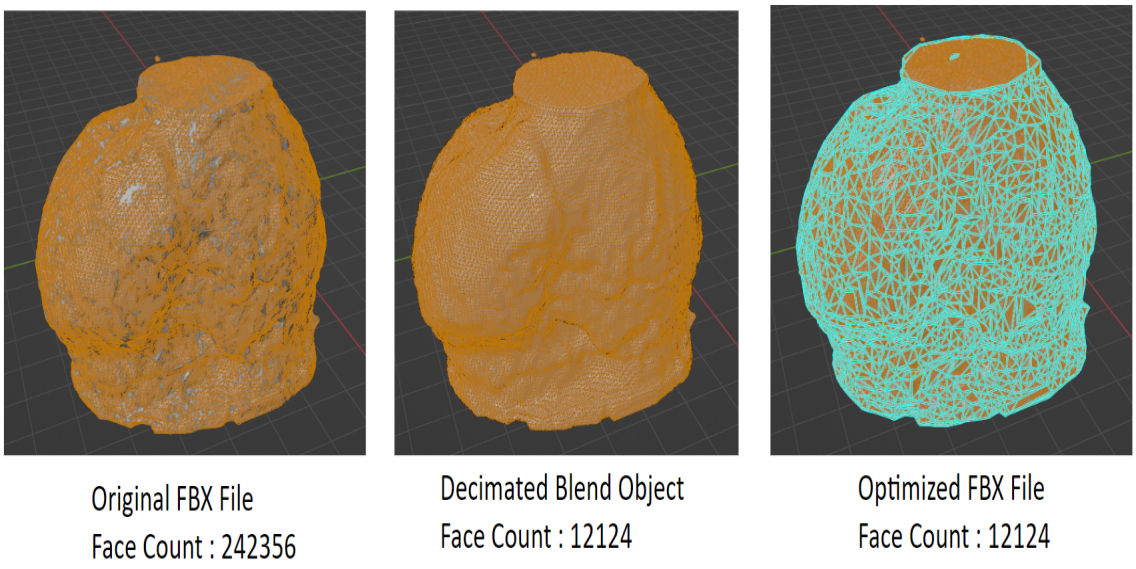

Figure 4.1: Visual representation comparing face counts in FBX images—depicting the original FBX file, decimated blend object, and optimized FBX file The varying face counts underscore the influence of optimization on 3D model complexity.

picting thin slices through a volume, datasets containing volume data (volume or 3D imaging), and acquisitions of the same tomographic or volumetric image over time, producing dynamic acquisition series (four-dimensional imaging). These file formats dictate how image data is structured within the file and guide software in correctly loading and interpreting pixel data for visualization. Medical image file formats can be categorized into two main groups. For this study, Digital Imaging and Communications in Medicine (DICOM), was used. DICOM stands as the standard for storing and transmitting medical images and has played a pivotal role in the development of modern radiological imaging.

In this study, datasets that contain volume data, often referred to as volume or three-dimensional imaging, are denoted as *static phantoms.* On the other hand, acquisitions of the same tomographic or volumetric image over time, leading to the creation of dynamic acquisition series commonly known as *four-dimensional imaging*, are termed *animated phantoms.*

4.1.2 Medical image processing

Image processing is a pivotal technique focused on improving and gleaning valuable insights from images. Image processing algorithms typically fall into three categories: Low-level methods, which address tasks such as noise reduction and color enhancement; medium-level techniques, including processes such as binarization and compression; and higher-level techniques, which encompass more advanced functionalities such as detection, segmentation, and recognition algorithms, are aimed at extracting semantic information from the captured data.

In US images, the signal-to-noise ratio is relatively low due to acoustic and electrical noise sources. When it comes to 2D ultrasound, noise can reduce image quality and make interpretation more difficult, while when it comes to 3D ultrasound, noise may occlude important structures and make them less visible. It is particularly difficult to delineate the boundary between heart tissue and the blood pool in 3D due to the noise generated by the cardiac ultrasound signals from the blood [33].

Fig. 4.2 illustrates the application of a recursive Gaussian image filter in this study to smooth, effectively removing noise and consequently eliminating the internal target. Therefore, the specific objectives of our study show that it is not necessary to apply any filters to enhance the visualization. However, to reduce the presence of noise, we employ color enhancement and support an interface or voice command to fine-tune visible intensity values.

4.2 Volume rendering

Volume rendering is a set of techniques used to visually represent a 2D projection of a 3D dataset composed of discrete samples. This approach finds applications in various fields, and one notable example is the visualization of medical images.
Key points regarding the data used in volume rendering:

- **3D Voxel Data:** The technique employs 3D voxel data, which is commonly

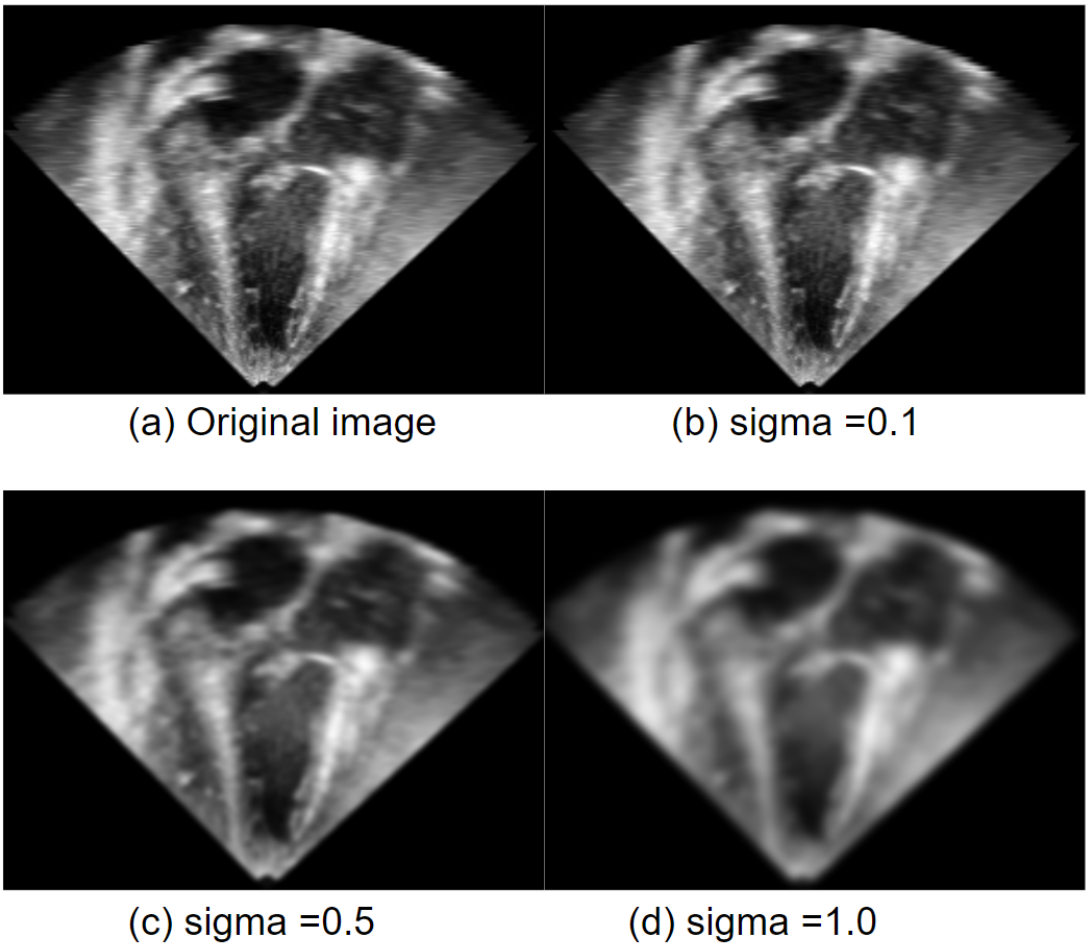

Figure 4.2: An illustration depicting the original images alongside the output images resulting from the application of a recursive Gaussian filter to smooth the 3DE at sigma values of 0.1, 0.5, and 1.0, respectively. The images showcase the coronal view of the 3DE and were viewed in 3D Slicer.

sourced from medical scans. Each data point within this 3D dataset represents the density at the corresponding location.

- **Data Acquisition:** The data is obtained through the scan and is saved in a 3D texture format. This texture contains the density information for each point.
- **Fitting into a Box:** Typically, the 3D data fits neatly within a bounding box. For rendering purposes, this box is visualized, and rendering techniques are employed to retrieve the densities of the visible voxels within the box.

In computer graphics, ray casting involves projecting rays or lines into a 3-dimensional environment to perform renderings, collision detection, and intersection calculations. It is commonly used for figuring out what objects or surfaces are visible from a certain point or for generating realistic images by simulating the path of light. Based on the application specification, a variation of the border concept of ray casting is used. By extending the concept of ray casting, ray tracing and ray marching provide realism and precision beyond traditional ray casting techniques [34].

4.2.1 Ray marching

The basic idea behind raymarching is to shoot a ray from the camera's viewpoint through each pixel of the screen and calculate the color of the object. The asset that the study used was implemented based on the ray-marching algorithm that supports back-to-front and front-to-back to get the ray for a specific fragment or vector.

Back face of a box by casting rays from each vertex toward the viewer's eye, dividing the rays into steps, fetching density information from a 3D texture at each step, and using this density data to decide the color of the resulting fragment or pixel. The exact color calculation logic is not provided and would depend on the specific rendering effect or material properties desired in the final visualization.

This set of instructions outlines a process for rendering the back faces of a box using a technique similar to ray marching. A breakdown of the steps is shown in

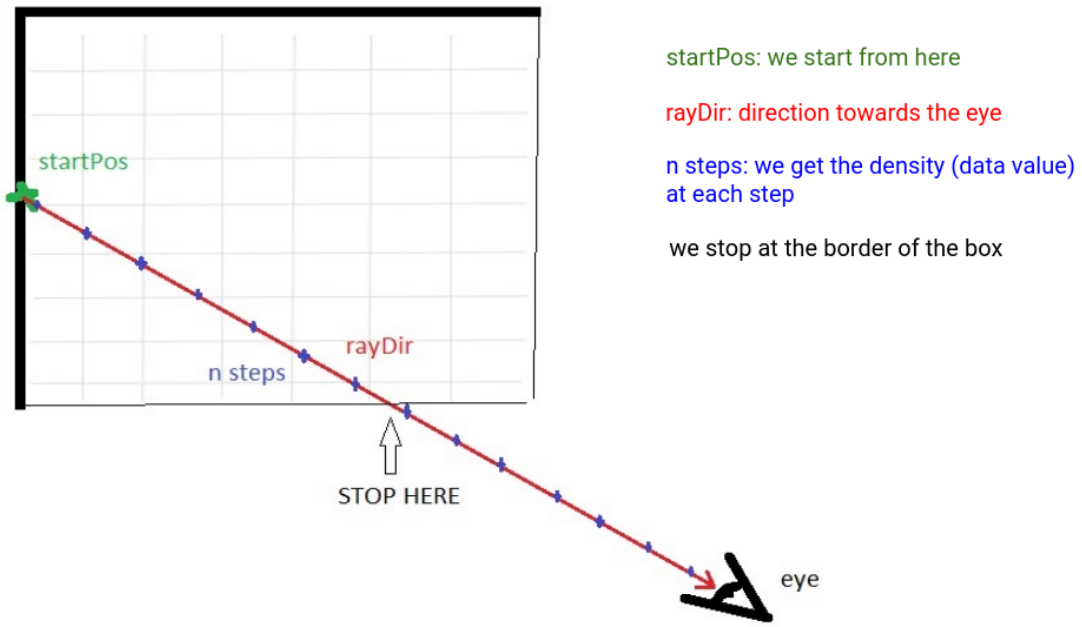

Figure 4.3: Rendering Back-Faces of a Box: Ray Marching and Density Mapping [40].

Fig. 4.3:

1. **Set start position (startPos):** For each vertex of the box, the local position of the vertex is considered the starting position. This local position represents a point in the 3D model's space.

2. **Set ray direction (rayDir):** The ray direction (rayDir) is determined from the starting position towards the eye or the camera. In the context of model space, this direction points from the vertex towards the viewer's eye.

3. **Divide the ray into N steps or samples:** The ray from the vertex to the eye is divided into N steps or samples. This division allows for more precise calculations along the ray's path, providing a higher-quality rendering result.

4. **Iterate through ray steps (Loop):** For each step along the ray (starting from the startPos and moving along the rayDir), the algorithm iterates through a loop.

5. **Get density of current position:** At each step, the current position's density

is calculated along the ray's path. This position is often used as a texture coordinate in a 3D texture. The density of the object at this point is obtained by fetching the corresponding value from the 3D texture.

6. **Use density for color calculation (Output color):** There are a few techniques to determine the output color based on the requirement or experimental type.

4.2.2 Rendering techniques to calculate colour for volume visualization

- Maximum intensity projection (MIP): It is used to visualize high-intensity structures within volumetric data. The highest data value encountered along corresponding viewing rays is displayed at each pixel. In this case, white will be the output color, where the alpha value will be assigned to maximum density. It illustrates big variations in density, which are commonly seen in CT scans, in a simple but powerful way.

- Direct volume rendering (DVR): In this technique, at each step along a ray, the current voxel's color is combined with the accumulated color value from the previous steps. The idea is to blend the colors smoothly to create a coherent and visually appealing result. To achieve this, linear interpolation (lerp) is used for blending the RGB values. It is a powerful method for volume rendering, especially when there is a need to combine different densities and transparencies to depict complex structures within the dataset. Fig. 4.4, which is a detailed flowchart diagram provides a comprehensive overview of the internal component and shows the steps involved in the DVR process in the plugin that the study used.

- Isosurface rendering: In isosurface rendering, the initial step involves setting a density threshold, which is a value that determines which voxels (data points)

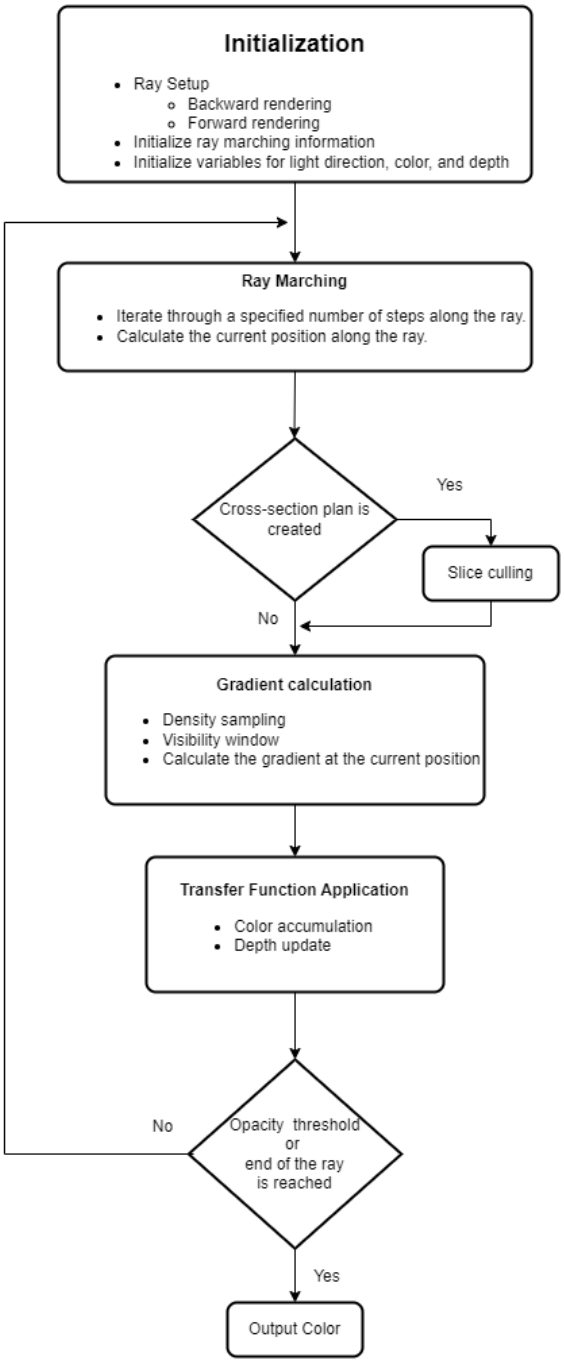

Figure 4.4: Flowchart diagram illustrating the process of how Direct Volume Rendering works.

will be considered. While raymarching away from the viewer, the moment the algorithm finds a voxel with a density exceeding the set threshold, the process halt and record that density value.

4.3 Rendering in HoloLens 2

4.3.1 Holographic rendering

In holographic rendering, holograms can be drawn at precise locations in the real world or in virtual realms created by the application. Mixed reality headsets continuously track the user's head's position and orientation in relation to their surroundings [41]. It predicts where the user's head will be when a frame appears on the display at the precise moment the application begins preparing its next frame. On the basis of this prediction, the system determines the view and projection transform for each frame. Once the holographic Unity project is ready for testing, the next step is to export and build a Unity Visual Studio solution. With that VS solution in hand, run the application in one of three ways on a real or simulated device.

To address the limitation of rendering the original animated phantom exceeding the maximum target size on HoloLens 2, the optimal solution involves modifying the size of the render volume or target buffers [42]. This can be achieved by invoking the RequestRenderTargetSize method to request a new render target size. However, it is essential to note that enabling a larger render target size will impose higher memory demands on the GPU. To effectively handle the rendering of larger volumes without overburdening the onboard resources of the HoloLens, it is advisable to leverage the capabilities of a PC. This is where holographic remoting comes into play.

4.3.2 Holographic remoting

Providing a user experience through the HoloLens but running the app on a PC to make use of its powerful resources is called holographic remoting [43]. This study is really helpful as it has high-resolution models and needs to maintain a good frame

rate. Holographic remoting is supported by a range of essential terms and core components. These include the player, remote server, and client, working together to create real-time experiences [43]. The Player, often found in the Windows Store, sends poses to the Remote, which performs rendering tasks and sends back video frames. Remote apps run on desktops or virtual machines, making them compatible with game engines.

A hostname and ports are essential for connectivity, with TCP and UDP protocols used. There are several challenges to overcome, such as cost, data privacy, and regulatory requirements. To achieve quality, data channels, bandwidth, and video encoding must be considered. The processing of video on hardware is accelerated to achieve low latency. Holographic Remote provides real-time streaming to XR devices and platforms using APIs such as Windows Mixed Reality and OpenXR. In fields such as healthcare and gaming, these technologies have the potential to revolutionize the industry.

4.4 Development of the application

4.4.1 Unity3D-based software rendering

Unity3D is widely used for creating 3D games and graphics applications and offers relative ease in building prototypes. It is also the most commonly used platform for HoloLens applications due to Microsoft's support. Unity 3D development is done using the C# programming language. Unity is a cross-platform environment for developing 2D, 3D, VR, and AR games on various platforms, supporting popular VR APIs such as Oculus and OpenVR, as well as other 3D display devices through manufacturer plugins. In the medical field, Unity is used for VR training environments [44, 45], and scientific/medical visualization [46] using surface rendering techniques. Directly rendering 3D medical images using volume rendering in Unity is desirable and has been implemented through fragment shaders [40], [22]. Unity's low-level na-

tive plugin interface enables multi-threaded rendering with external volume rendering libraries [40, 47, 48] for direct display in Unity.

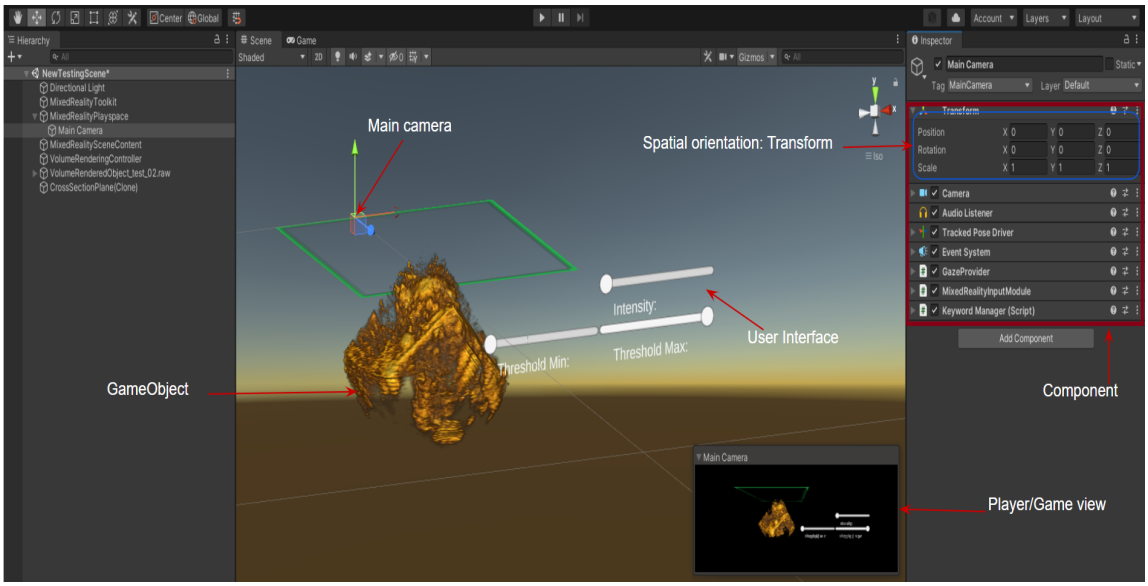

Figure 4.5: The setup of an implemented Scene with a GameObject, UI, a main camera, and other key terms in Unity development space.

In the world of Unity development, various key terms and components play pivotal roles:

- **Game Engine:** A game engine is the core software that powers video games, managing everything from graphics and physics to player input and artificial intelligence.
- **GameObject:** GameObjects are basic entities in a game world that can represent characters, objects, or elements. They are the building blocks of the game.
- **Component:** Components are functional parts that can be added to GameObjects to give them specific behaviors or attributes, such as movement or sound.
- **Gestures:** Gestures are specific actions or motions used by users to interact with software, such as swiping a touchscreen or making hand movements in an AR environment.

- **User Interface (UI):** The UI is the interface through which users interact with software or games, including menus, buttons, and displays.

- **Mesh Renderer:** Mesh renderers are responsible for displaying 3D objects in a game world, ensuring they look as intended.

- **Colliders:** Colliders define the physical boundaries and interactions of objects in the game, enabling collision detection and response.

- **Spatial Orientation:** Spatial orientation involves determining the position and orientation of objects in a 3D space, which is crucial for creating immersive virtual environments and responsive interactions.

The setup of a scene is shown in Fig. 4.5 with a GameObject, UI, a main camera, and three directional arrows representing the 3D coordinate system. In Unity 3D, there are several essential Unity 3D components—materials, shaders, and textures—that contribute to the development and rendering of 3D objects in the program. These elements are perfectly blended to create realistic and visually appealing 3D graphics. It becomes necessary to manipulate materials, shaders, and textures in order to accomplish particular visual effects and improve the overall aesthetics and guiding principles of the program.

- **Materials:** Materials provide texture references, tiling specifications, color tints, and other characteristics that specify how a surface should be created. The shader that the material uses determines the possibilities that are available to it.

- **Shaders:** Shaders are brief scripts that contain mathematical formulas and techniques for determining the color of every pixel that is displayed in relation to the lighting input and material settings.

- **Textures:** Textures are made from bitmap images. Material may have texture references, which enable the shader of the Material to use the textures to determine the surface color of an object. In addition to an object's primary color, textures can reveal a variety of surface characteristics, including roughness and reflection.

4.4.2 Mixed reality in HoloLens 2

HoloLens 2 is designed as an AR device, but most of the documentation sometimes uses the term *mixed reality* to describe its approach, which is to blend digital and physical content. There are a few key structural elements that affect and contribute to the quality of the mixed reality experience of the app.

Coordinate systems: Spatial coordinate systems within the Windows environment encompass various real-world coordinate systems used for delineating spatial geometry. These systems facilitate the calculation of hologram attributes such as position, orientation, gaze ray, and hand locations. These spatial coordinate systems are metrically measured. Consequently, objects positioned at a distance of two units along the X, Y, or Z axes will appear two meters apart in the mixed reality setting. A comprehensive understanding of spatial coordinate systems empowers the seamless rendering of objects and environments on a true real-world scale. Unity utilizes a left-handed coordinate system, while spatial coordinate systems on Windows, including Windows Mixed Reality, consistently adhere to a right-handed orientation. Fig. 4.6 visually illustrates the left and right coordinate systems for enhanced comprehension.

Spatial mapping: Spatial mapping serves as a blueprint for precisely placing holograms within a designated area. It enhances the mixed reality experience by enabling holographic objects to interact seamlessly with the physical surfaces it defines.

Holographic frame: Through the headset, users can experience the world in mixed reality via a rectangular viewport. The holographic frame, a designated area on

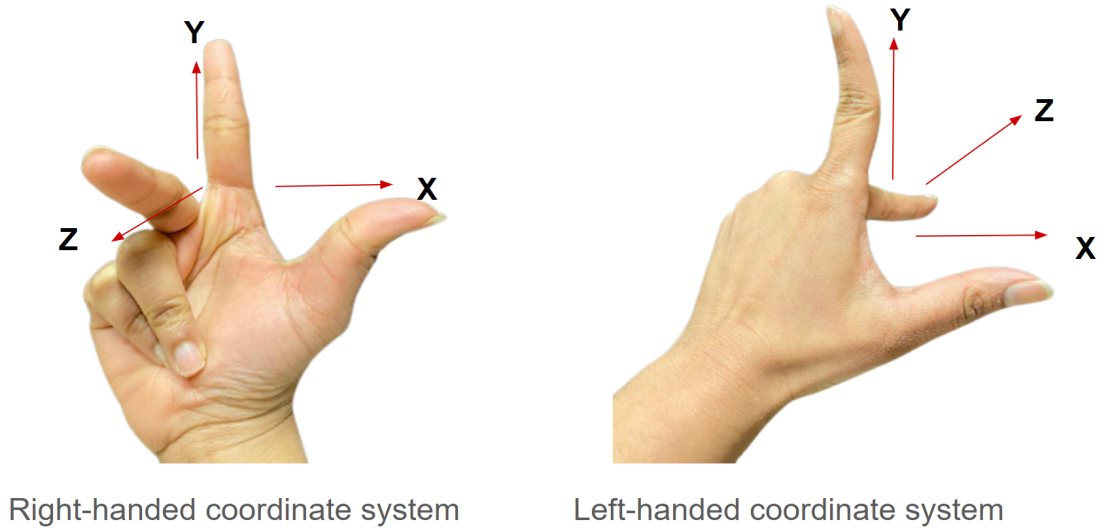

Figure 4.6: Visual representation depicting the left and right coordinate systems for clarity and understanding.

the HoloLens, allows users to observe digital content seamlessly integrated with the surrounding real world.

Mixed reality development tool

Microsoft provides the Mixed Reality Toolkit as an open-source project to assist developers in creating mixed reality (MR) applications. The Mixed Reality Toolkit supports both AR and VR development across various platforms, including Microsoft HoloLens, Windows Mixed Reality headsets, and other Windows Mixed Reality immersive devices. The Mixed Reality Toolkit for Unity comprises components and scripts aimed at expediting the development of applications for HoloLens and Windows Mixed Reality headsets. This toolkit incorporates UX controls, fundamental user inputs, and spatial mapping APIs to facilitate a swift start to the development process. For HoloLens 2, user input can be accepted in various forms, including speech, gesture, or gaze. In this study, two plugin components associated with the MRTK have been used. These plugin components, integral to the research methodology, contribute to the development and enhancement of mixed-reality experiences.

Mixed Reality Toolkit Foundation

The toolkit simplifies the development of applications that leverage features such as spatial mapping, hand tracking, gesture recognition, and interaction with holographic content.

Mixed Reality OpenXR

OpenXR is an open standard for AR and VR platforms. The Mixed Reality OpenXR plugin is an extension or module designed to integrate OpenXR support into mixed-reality applications, ensuring compatibility across different platforms.

4.4.3 System architecture

The system developed in this thesis is an application module that acts as an interface between the input DICOM US images and the output HoloLens 2 device, which would be used by a surgeon. The developed system is depicted using a simplified architecture diagram in Fig. 4.7. The DICOM images are converted to individual raw images and provided as input to the system. This information is processed to provide a mesh using Unity 3D, which can then be visualized on the output device, a Microsoft HoloLens 2. In addition to providing the user with the ability to view the 3D surface, a plane is developed to view the cropped version of the phantom and see internal structures. Moreover, hand tracking and object interaction are enabled to perform spatial orientation to register 3DE in physical space. Additionally, to support the precision of object manipulation and noise removal, a UI interface is developed, and voice commands are integrated.

4.4.4 3D+t rendering

It is challenging to generate an accurate surface representation of the data where it contains high-noise components, as in the case of echocardiography. Instead, a volumetric rendering approach that does not require the surface to be extracted can

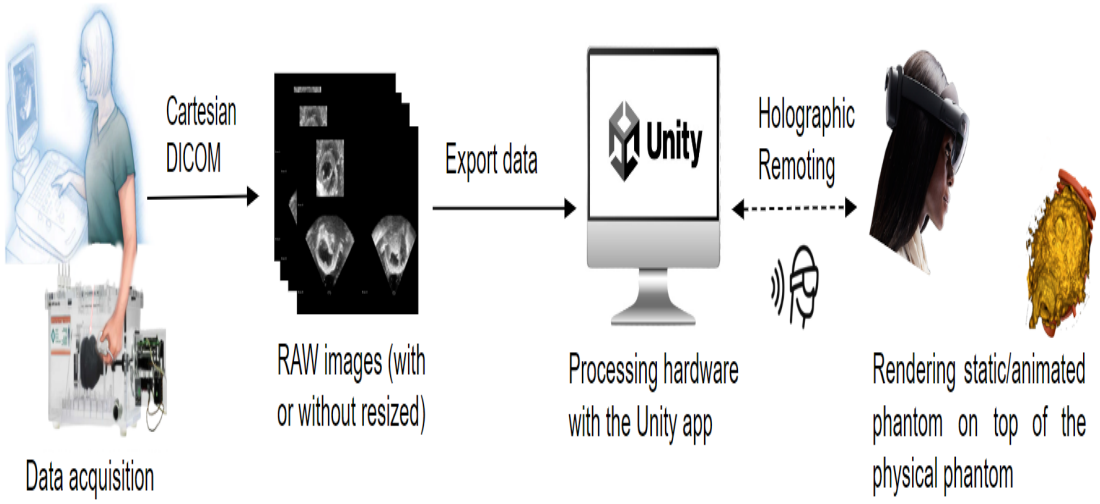

Figure 4.7: Overall components showing the acquisition, processing, and display of the 3D echocardiography data used in this study. The proposed approach uses Unity3D for rendering the volumetric data on a Microsoft HoloLens 2 augmented reality device.

be utilized for the problem. Thus, a ray marching-based DVR approach was adopted as one of the newer approaches to the ray model. It attempts to render faster than ray tracing by jumping (or marching) in fixed steps along the ray, making the time until an intersection occurs shorter.

The volume rendering algorithm used in this study is based on ray marching, which is available as a Unity plugin [40] to visualize static medical images. The image processing and rendering algorithms were implemented using C# and utilized the free asset of the plugin [40] from the Unity Asset Store. The plugin offers an option to import raw data, and the header file is read to determine how the data is organized (e.g., content format, dimension, endianness, number of bytes to skip, etc.) for static volume rendering. Once imported, the datasets are converted into game objects to be compatible with the Unity platform, and a script is attached to them to render them as volumes in Unity. In this study, to implement the beating

heart phantom model, a new script was added to support time series rendering of the echocardiography datasets by storing them sequentially in an array of data textures. The static rendering plugin [40] is then used to individually render the textures at each time frame. The visualization of the beating heart phantom model was achieved by importing all the raw files in the sequence. Then, a dataset is periodically selected from the array with a wait time based on the frame rate, allowing it to support temporal synchronization with a *beating* heart phantom. The proposed development does not support the live 3DE image rendering in HoloLens 2, and the heart rate or R—R interval value could not be located in the header information of the file used for this project. In that case, this study assumes the R—R interval to be 800 ms. Based on the predefined value of the wait time to visualize the 3DE volumes in the array list, it is defined to support the *beating* heart phantom.

4.4.5 Remote rendering

Initially, the static volume data was transferred from a PC to HoloLens 2 using a wireless network. However, due to the large image size, the frame rate dropped significantly, leading to a loss in performance. To maintain an acceptable frame rate, the data had to be resized. The frame rate dropped to an unacceptable 16 fps when the rendering of the beating heart phantom was enabled, which was not suitable for the application. One solution to address this issue is to increase the network transmission bandwidth and perform remote rendering to transfer the results to the display device. To achieve this, a higher computing power device, such as a PC, was utilized to calculate the view and then transfer the final view to the HoloLens 2. This approach helped minimize latency and optimize network transmission speed. To implement this solution, the holographic remoting option was chosen to utilize the PC's resources for faster calculations in the app. The PC was connected to the network via Ethernet, while the HoloLens 2 was connected through a mobile hotspot to support wireless network transmission and ensure faster data transfer.

Fig. 4.8 illustrates the comprehensive architecture of the rendering pipeline. This visual representation provides a detailed overview of the entire system, showcasing the interconnected components and the flow of data through each stage.

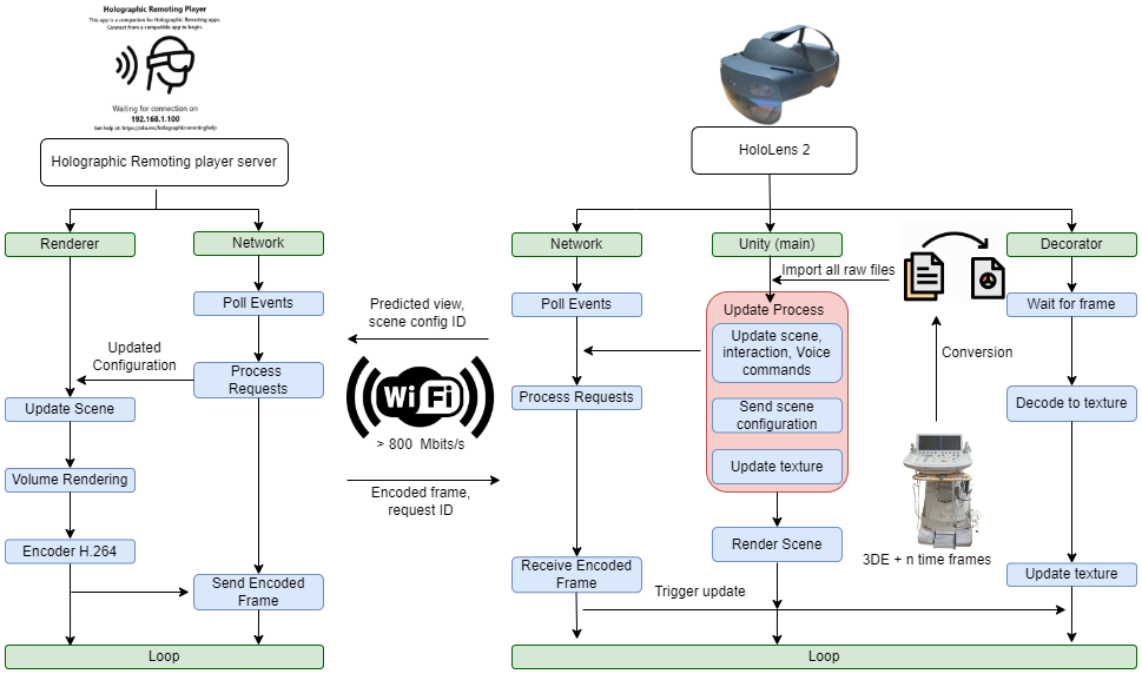

Figure 4.8: Architecture diagram of the remote rendering pipeline.

4.4.6 User interface

One of the key strengths of HoloLens 2 in the medical field is its intuitive and interactive user interface. This software allows medical professionals to manipulate and explore medical images in a novel way. For example, the user interface provides options to cut off the data in all three dimensions, adjust the visible intensity ranges, and change the density of the voxel. These interactive tools facilitate a deeper understanding of complex medical datasets and enhance the precision of diagnosis and treatment planning. Fig. 4.9 shows the interface developed to interact with the rendered phantom to change its transfer functions.

The cross-section plane is a versatile tool that can be moved and rotated to clip the volume at any arbitrary position and angle, providing a comprehensive view of the

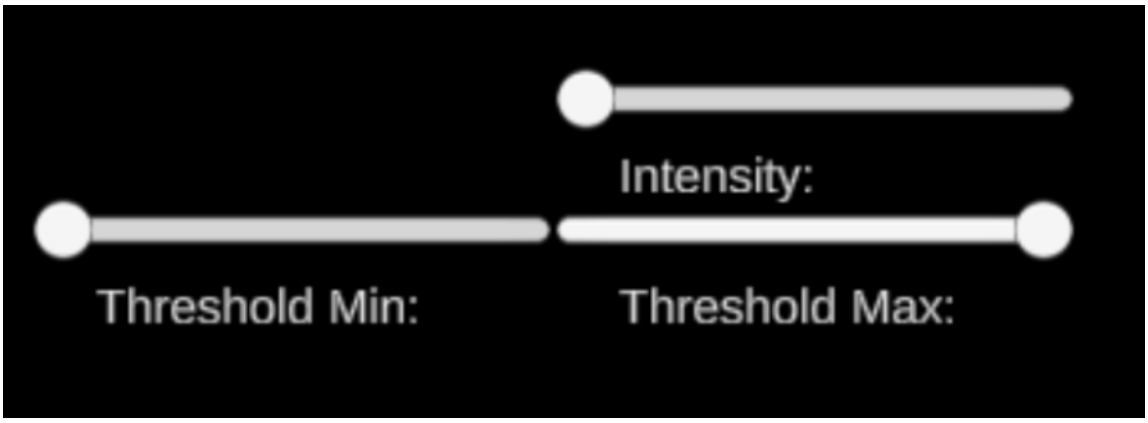

Figure 4.9: Implemented interface for HoloLens 2 to interact with the phantom and adjust its transfer functions.

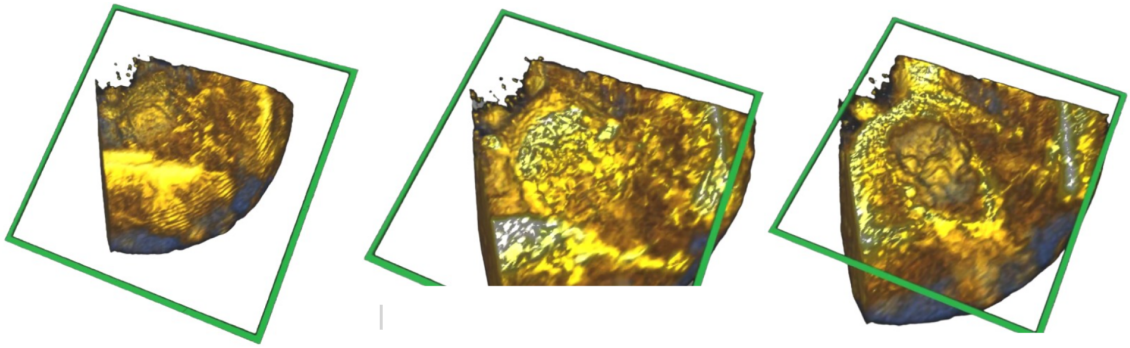

Figure 4.10: Illustration of the phantom's appearance during movement across the cross-sectional plane.

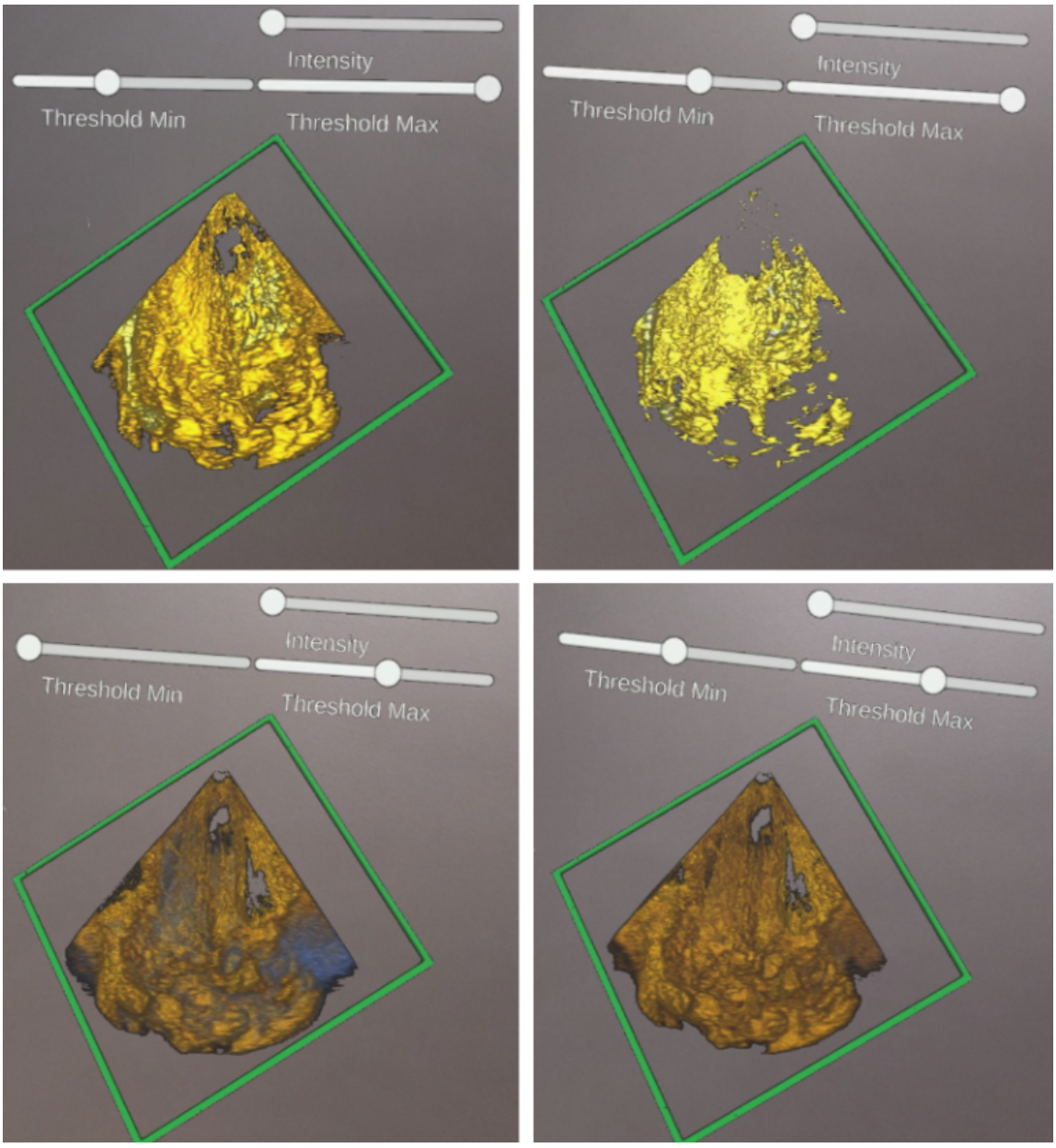

Figure 4.11: Example images of a static phantom of the size of $224 \times 208 \times 208$ rendered with various minimum and maximum intensity values set through a user interface for holographic remote rendering.

volumetric data from different perspectives. This approach allows for a more comprehensive exploration and analysis of the 3D data, enhancing the visualization and understanding of the medical images in the application. Fig. 4.10 shows a rendering of a 3DE of a heart phantom and a sequence of the beating heart phantom while the cross-section plane is moved towards the phantom.

The slider threshold control illustrated in Fig. 4.11 allows developers to adjust the appearance of the mesh in real-time, enhancing user interaction and exploration. This feature could be utilized to reveal internal structures gradually as the user interacts with the 3D model. When the threshold value is adjusted for the upper and lower bounds through the slider control, specific regions of the mesh can be made transparent or opaque, providing a detailed view of complex anatomical structures.

The functionality of the slider labeled *Intensity* is employed to adjust and elevate the intensity of each individual pixel value within the displayed image. As the slider is adjusted, it systematically increases the brightness or luminosity of the image pixels, effectively enhancing the overall intensity of the visual representation, as shown in Fig. 4.12. The utilization of the intensity slider serves as a dynamic and interactive means to modify the visual characteristics of the image and tailor its appearance to meet specific analytical or perceptual requirements. Fig. 4.12 illustrates how the slider intensity is used to increase the intensity of each pixel value.

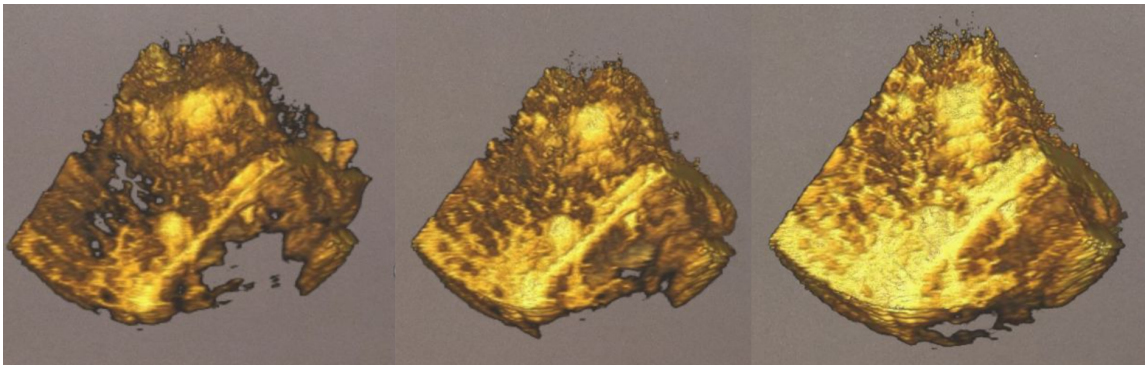

Figure 4.12: The *Intensity' slider* in action dynamically adjusts and elevates the intensity of each pixel value within the image.

4.4.7 Voice inputs

Voice input [49] is a fundamental method for users to interact with the HoloLens, offering the advantage of natural and efficient communication with holograms without relying on gestures. Speech recognition, an integral part of voice input, is based on the same engine that supports speech in Universal Windows Apps, ensuring users can communicate seamlessly in their device's configured Windows display language. To enhance user experiences, integrating voice commands into applications is highly recommended. These voice commands should be carefully designed to ensure they are clear and universally understood, given the diversity of accents and dialects. It is essential to keep commands concise, opt for simple vocabulary, and make sure they are non-destructive, allowing easy undoing of actions.

Maintaining consistency in voice commands across applications and avoiding similar-sounding commands are essential practices to ensure users can navigate mixed-reality environments without confusion. Moreover, there are challenges related to voice input, such as the difficulty of fine-grained control, especially when quantifying commands such as adjusting volume. Reliability remains a concern as voice systems can occasionally misinterpret user commands, necessitating informative feedback from users to address potential misunderstandings. Further, voice input may not always be appropriate in shared spaces or for sensitive information, as users may feel awkward or uncomfortable when speaking aloud. Unfamiliar or unique words pose challenges, as the system may not recognize them, and teaching users a significant set of voice commands without overwhelming them is an ongoing challenge [49].

Within the context of the MRTK, developers can effortlessly assign voice commands to objects, simplifying interactions with holograms. By using MRTK's Speech Input Profile and the SpeechInputHandler script, developers can make objects respond to specific keywords, improving user confidence and enhancing the overall mixed reality experience.

Speech input profile [50], offers a way to trigger events based on recognized keywords, bypassing the need for physical controllers. These keywords are configured in the Speech Commands Profile within the Input System Profile. For each keyword, there is flexibility to [50]:

- **Map to an Input Action:** Recognize keyword with an input action.
- **Specify a Key Code:** Define a key code that, when pressed, triggers the same speech event, providing an alternative input method for those who prefer key presses.
- **Include a Localization Key:** To ensure compatibility with Universal Windows Platform (UWP) apps, a localization key can be added. This key is used to fetch the localized keyword from the app's resources, making voice commands more user-friendly and adaptable to various languages.

The Speech Input Handler script can be incorporated into a GameObject to manage speech commands using UnityEvents. This script conveniently displays the list of defined keywords from the Speech Commands Profile, streamlining the process of handling speech commands within Unity.

Due to the larger image size and inability to render stable 3DE data in the UWP app, we rely on the holographic remoting approach to render in HoloLens 2. In that case, we cannot use the Speech Input Handler to set up voice commands. The alternative solution is to add voice input to the Unity application.

In Unity, there are three primary methods to incorporate voice input into your application, with the first two categorized under the PhraseRecognizer [41]:

- **KeywordRecognizer:** This approach equips the application with a predefined array of string commands to actively listen for.
- **GrammarRecognizer:** In this method, the application is provided with a Speech Recognition Grammar Specification (SRGS) file. This file defines a

specific grammar or set of rules for recognized speech. It enables the app to listen for speech patterns that adhere to the defined grammar, ensuring more precise recognition of user commands.

- **DictationRecognizer:** This recognizer allows the application to listen for any spoken words from the user. Unlike the previous two methods that rely on predefined commands or grammar, DictationRecognizer is more flexible. It can be used to transcribe speech for note-taking or generate a display of the transcribed speech.

Each of these recognition methods offers varying degrees of flexibility and specificity in handling voice input, allowing developers to choose the most suitable option based on the application's requirements.

Since this study focuses on recognizing keywords and not grammatical correctness or speech-to-text. The voice command relied on KeywordRecognizer. It allows users to interact with the application by speaking recognized voice commands, providing a hands-free and intuitive means of controlling the following functions:

- **Position:** To disable the object manipulation to avoid minimal hand tracking errors when using spacial orientation results in an adjustment in object position.
- **Remove Minimum:** Remove the minimum voxel density in the object to view the clear structure.
- **Remove Maximum:** Remove the maximum voxel density in the object to view the clear structure.
- **Add Minimum:** Add the minimum voxel density in the object to view the clear structure.
- **Add Maximum:** Add the maximum voxel density in the object to view the clear structure.

Chapter 5

Evaluation and Results

5.1 Data description

This study primarily focuses on the analysis of pediatric cardiology data. Notably, the research did not involve the utilization of actual patient data during its initial Phase I analysis. Instead, a sophisticated phantom model was employed, featuring multiple internal targets representing structures such as papillary muscles and the atrioventricular leaflet-annular interface. This 3D phantom heart model was created using silicone rubber and a molding process that incorporated heated plastisol, red dye, and chalk powder. This phantom model was carefully designed to include both inflow and outflow components, ensuring its compatibility with a dynamic heart phantom system at the Servier Virtual Cardiac Center. Fig. 5.1 provides a visual representation of this phantom model. Additionally, the dynamic heart phantom system, as depicted in Fig. 5.2, was set up for the study. This system incorporated a pneumatic regulator to simulate a *beating* heart, capable of reaching up to 110 bpm.

5.2 Data acquisition

During the research, echocardiography images of the heart phantom were acquired using the X3-1 probe on a Philips system. The acquired data was initially formatted in Cartesian DICOM, which was then further processed to create individual raw images for loading into the application, thus facilitating the analysis.

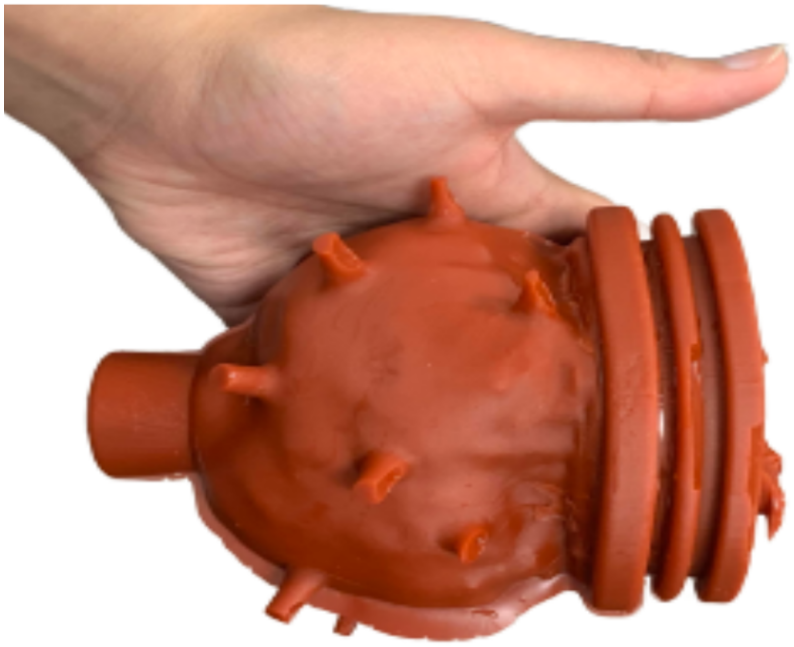

Figure 5.1: A 3D heart phantom made of silicone consisting of internal heart structures used for the experiment.

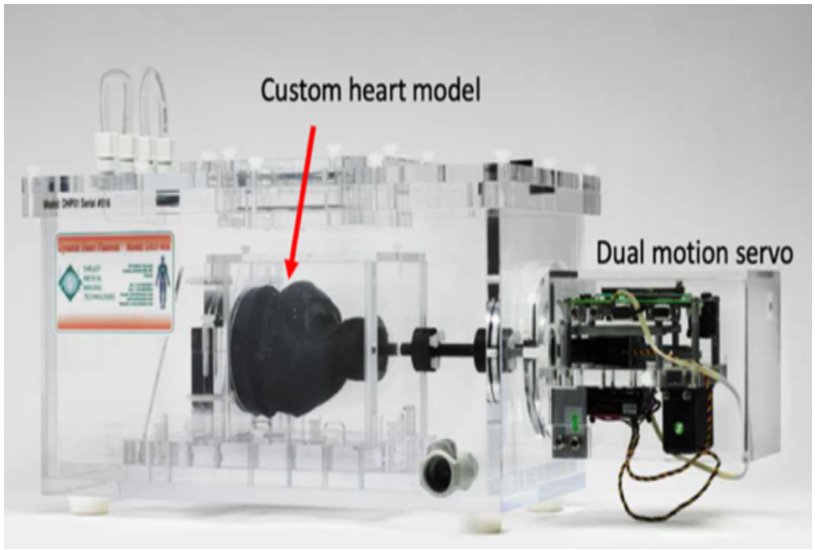

Figure 5.2: Dynamic heart phantom system [51] used in this study to simulate heart motion, which is captured using an echocardiography scanner and rendered through HoloLens 2.

5.3 Test environment

The system was tested using a desktop personal computer (Intel® Core™ i7-12650H, 16 GB RAM, NVIDIA® GeForce RTX™ 4070 GPU, and Windows 11 64-bit operating system) as the computer that processes the 3DE images and sends the stereo images to HoloLens 2.

5.4 Quantitative comparisons

5.4.1 Performance metrics analysis across diverse holographic scenarios

The performance assessment of the proposed rendering system included different scenarios, each designed to thoroughly examine its capabilities. These scenarios included static and animated representations of heart phantoms derived from echocardiography data, which included image volumes in their original dimensions and scaled-down versions. A tabulated summary of the results of these diverse scenarios, along with their corresponding performance metrics for holographic rendering and remoting, is presented in Table 5.1 with the movement of HoloLens 2.

The findings from these evaluations reveal intriguing insights into the system's behavior. In particular, it was observed that resizing the phantom data and reducing its inherent complexity led to tangible improvements in the performance of the holographic rendering. However, it is imperative to acknowledge a specific warning encountered during the assessment of the animated phantom scenario. At a data transfer rate of 150 Mbits/s with device connectivity during holographic remoting, occasional issues were reported, including instances of *Device not found* and *Tracking lost*. It is essential to underscore that the WiFi setup in use boasts an available capacity of 832 Mbit/s on the HoloLens 2. This underscores the importance of maintaining a consistent WiFi connection with a stable bandwidth, especially within the sterile environment where these visualizations are deployed, to prevent any potential disrup-

tions in the rendered scenes. This is particularly critical, as user mobility during the visualization process can further compound the demands on network stability.

Table 5.1: The performance of holographic rendering and remoting in different scenarios is evaluated using metrics that measure the quality of the holographic experience, such as the frame rate, latency, and image quality.

Model	**Size**	**Range of Frame Rate Per Second (fps)**	
		Holographic Rendering	***Holographic Remoting***
Static Phantom - Original	$224 \times 208 \times 208$	$15 - 30$	$47 - 58$
Static Phantom - Resized	$64 \times 64 \times 64$	$18 - 35$	$55 - 60$
Animated Phantom - Original	$224 \times 208 \times 208 \times 12$	N/A	$38 - 52$
Animated Phantom - Resized	$64 \times 64 \times 64 \times 12$	$6 - 16$	$42 - 55$

5.4.2 Analysis of GPU resource utilization for static and dynamic phantom rendering

Table 5.2 provides a detailed breakdown of the maximum GPU memory usage, GPU computational unit usage, and video engine load during both static and animated phantom rendering. These critical performance metrics were precisely measured us-

ing the GPU-Z tool, affording a broad view of the system's resource consumption dynamics. Especially, the data disclosed that the animated phantom rendering approach consumed a maximum GPU memory of 1117 MB, around 200 MB higher than the maximum GPU memory consumption observed during the rendering of the static phantom. This divergence in resource usage underlines the additional demands imposed by animated phantom on the GPU, emphasizing the importance of managing memory resources efficiently for optimal performance.

Moreover, the original-sized animated phantom utilized a GPU computational unit usage rate averaging 74.2%, which is more than 2.3 times the utilization of the original-sized static phantom. This shows that rendering multiple frames will not have a significant difference in rendering using holographic remoting.

Table 5.2: The Peak GPU memory usage, GPU computational unit usage, and video engine load for the original size static and animated phantom rendering.

Data	**Size**	**Peak GPU Memory Usage (MB)**	**GPU Load (%)**	**Video Engine Load (%)**
Static Phantom	$224 \times 208 \times 208$	919	31.841 ± 12.656	23.853 ± 4.664
Animated Phantom	$224 \times 208 \times 208 \times 12$	1117	74.189 ± 33.068	25.861 ± 8.187

Fig. 5.3 shows the system memory usage in MB for the original static phantom size of $224 \times 208 \times 208$ and the original animated phantom size of $224 \times 208 \times 208 \times 12$.

5.4.3 Performance evaluation of direct volumetric rendering on HoloLens 2 with varied distances and measures

For a more comprehensive discussion of the performance analysis performed on the algorithm, the direct volumetric rendering on HoloLens 2 was rigorously examined

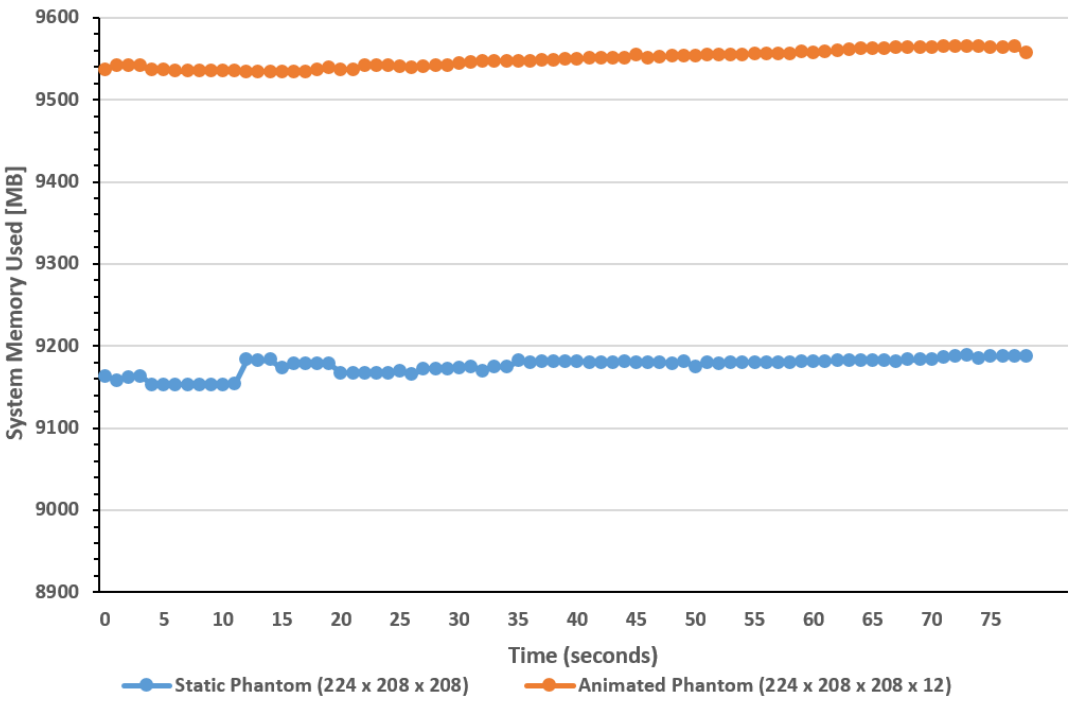

Figure 5.3: The amount of system memory used in megabytes for the static and animated phantom data for the original echocardiography data sizes.

in various scenarios involving different volumes and measures. A pivotal aspect of this evaluation was determining the highest achievable frame update rates for the HoloLens 2, taking into account the dynamic interplay between the rendered volume's proximity to the device and the utilization of holographic remoting and rendering techniques.

The data obtained with performance tests are shown in Fig. 5.4 and Fig. 5.5, which are, respectively, the tests for holographic remoting and holographic rendering with the movement of the HoloLens 2.

Based on the graphs in Fig. 5.4 and Fig. 5.5, it can be seen that the fps rate of the lowest volume is very accentuated, contrary to the larger volumes, which presented performance below 16 fps. The performance of the rendering is adversely affected by the increase in volumetric data size and any sudden movement of the user wearing the HoloLens 2.

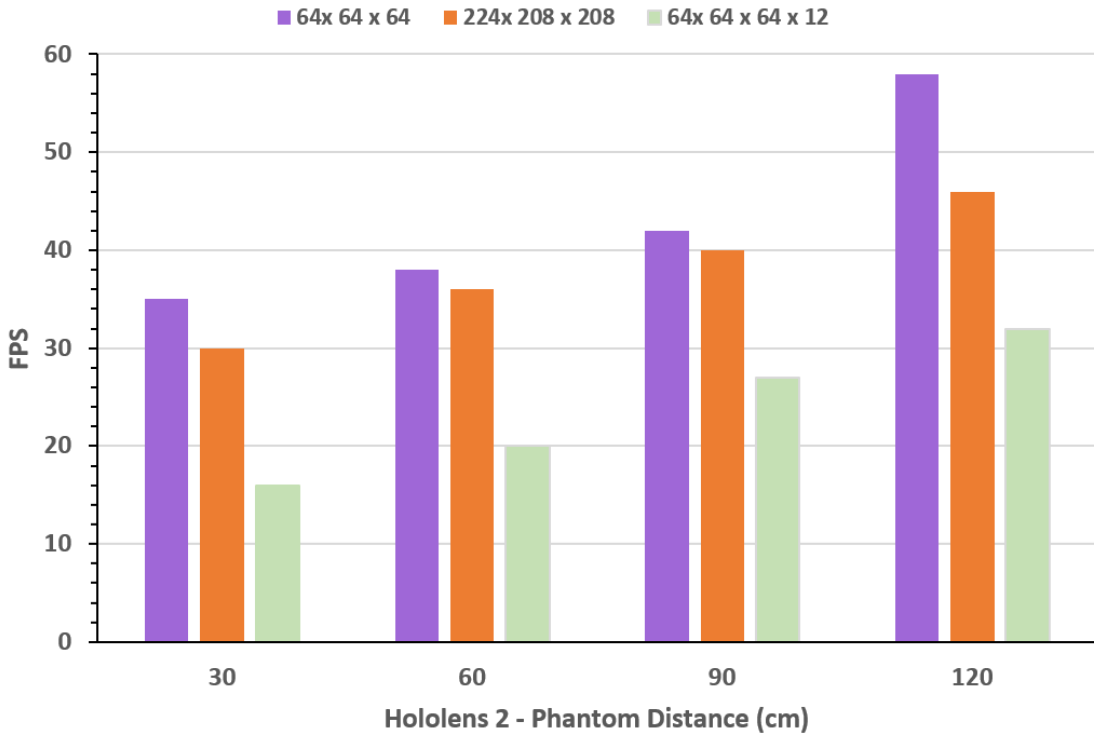

Figure 5.4: Maximum frame rate attained at different distances for holographic rendering.

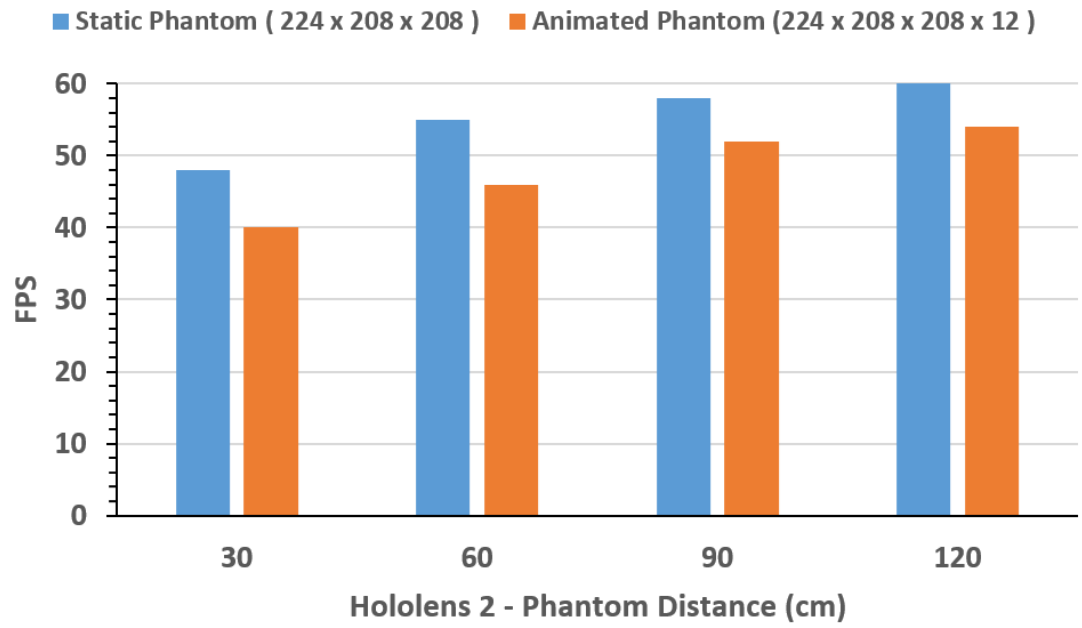

Figure 5.5: Maximum frame rate attained at different distances for holographic remoting.

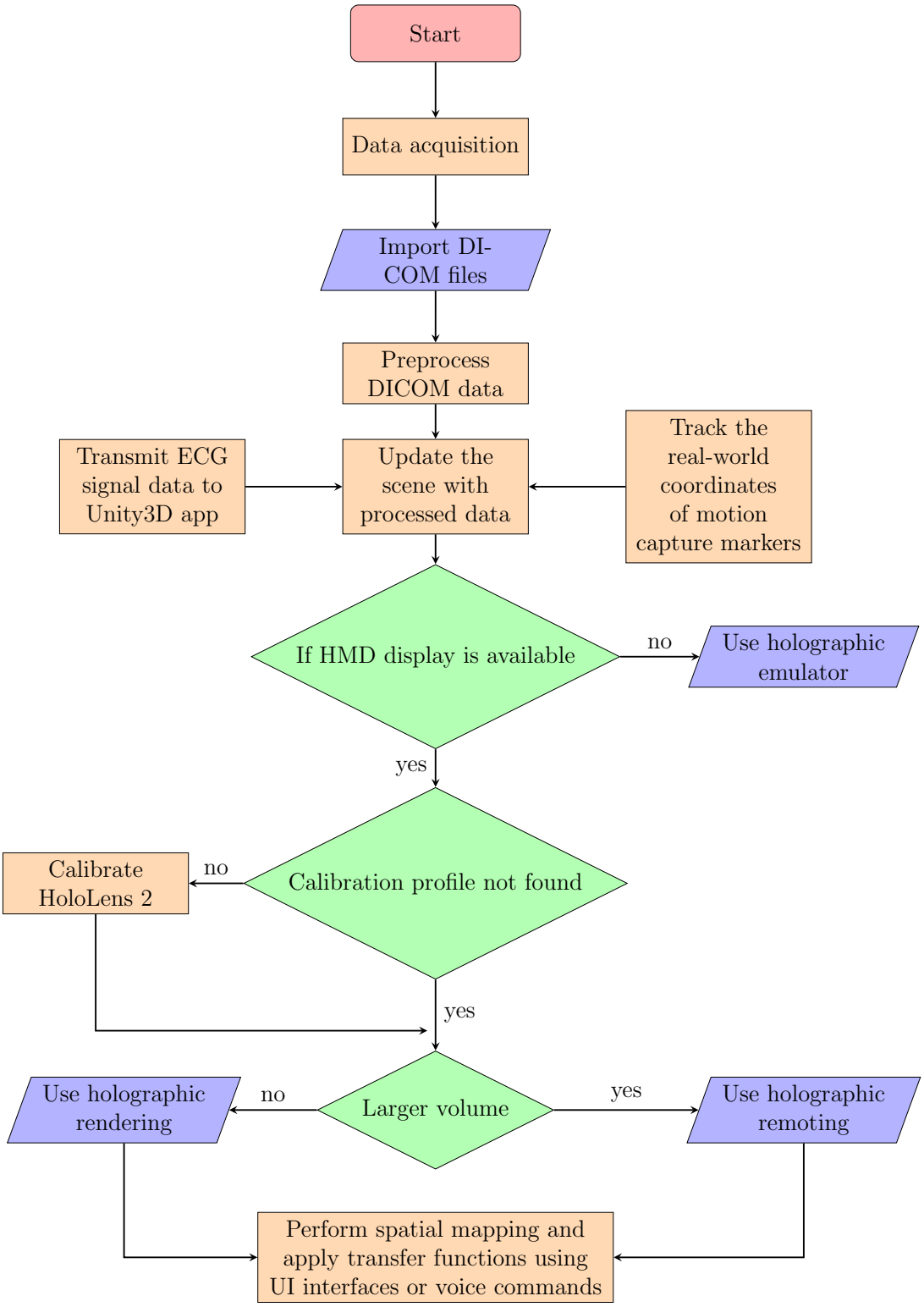

Figure 5.6: Flow chart of tasks required to achieve real-time rendering and guided surgical intervention

Fig. 5.6 outlines the schematic representation of tasks necessary to accomplish real-time rendering and guided surgical intervention. It serves as a visual guide, depicting the sequential steps involved in the process.

Chapter 6

Conclusions, & Future Work

6.1 Conclusions

This study presents an innovative approach that utilizes HoloLens 2 for augmented reality rendering of 3DE datasets, enabling volumetric visualization of the data. Our proposed method encompasses the rendering of both static and dynamic echocardiography images, facilitating the creation of animations for enhanced visualization. Leveraging advanced volumetric rendering capabilities, an interactive user interface, and stereoscopic image rendering, HoloLens 2 emerges as a compelling choice for visualizing echocardiography datasets within an augmented reality environment. Evaluations have demonstrated that employing the holographic remote rendering option yields improved framerates compared to rendering the volumes directly on the head-mounted hardware. This advancement opens up new possibilities for efficient and immersive echocardiography data exploration.

6.2 Limitations & future work

6.2.1 The hardware and software limitations

Utilizing HoloLens 2 for rendering 3D presents several hardware and software limitations. The device's processing power and display resolution may not suffice for intricate 3D rendering, leading to performance issues in augmented reality. Users of the HoloLens 2 face the challenge of a limited field of view, which requires them

to frequently adjust their viewing angle to explore the entire 3D environment. To enhance the overall experience of viewing and interacting with 3D content on the HoloLens 2, key challenges such as software optimization, efficient data transfer, user interface design, and memory management must be addressed.

6.2.2 Manual interaction

Hand gestures may not be accurate enough for precise alignment, which could cause slight discrepancies between physical and holographic objects. The HoloLens 2's limited tracking capabilities could make it difficult to maintain a stable alignment, particularly when users move or perform gestures. Additionally, the physical phantom may not always be visible, making real-time alignment a challenge. To overcome these limitations, advancements in hand-tracking accuracy, stability, and the ability to seamlessly interact with real-world objects would be necessary.

6.2.3 Application of spatial registration methods

The easiest option is to perform spatial registration using the planar image fiducial marker. However, in general, they lack robustness and accuracy. It has been shown that the tracking error using common libraries can range from several millimeters to even centimeters [52, 53] and is highly dependent on viewing angles, distance, lighting conditions, and movement patterns. Marker-less approaches such as the possibility of using the various onboard sensors of the HoloLens for inside-out, marker-less registration. In this study, the inside-out approach could not be used as it relies on holographic remoting. An alternative solution might be using an outside-in marker-based approach is recommended.

Bibliography

[1] L. Drukker *et al.*, "Transforming obstetric ultrasound into data science using eye tracking, voice recording, transducer motion and ultrasound video," *Scientific Reports*, vol. 11, no. 1, p. 14 109, 2021. DOI: 10.1038/s41598-021-92829-1. (visited on 11/07/2023).

[2] A. Carovac, F. Smajlovic, and D. Junuzovic, "Application of ultrasound in medicine," *Acta informatica medica: AIM: journal of the Society for Medical Informatics of Bosnia & Herzegovina: casopis Drustva za medicinsku informatiku BiH*, vol. 19, no. 3, pp. 168–171, 2011. DOI: 10.5455/aim.2011.19.168-171.

[3] R. S. o. N. A. R. a. A. C. o. Radiology (ACR), *General Ultrasound*. [Online]. Available: https://www.radiologyinfo.org/en/info/genus (visited on 06/16/2023).

[4] D. Neumann and E. Kollorz, "Ultrasound," in *Medical Imaging Systems: An Introductory Guide*, A. Maier, S. Steidl, V. Christlein, and J. Hornegger, Eds., Cham (CH): Springer, 2018, pp. 237–249.

[5] F. von Haxthausen, R. Moreta-Martinez, A. Pose Díez de la Lastra, J. Pascau, and F. Ernst, "UltrARsound: In situ visualization of live ultrasound images using HoloLens 2," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 11, pp. 2081–2091, 2022. DOI: 10.1007/s11548-022-02695-z.

[6] *Mitral Valve Regurgitation*. [Online]. Available: https://www.cardioshoals.com/cardiovascular-conditions/mitral-valve-regurgitation/mitral-valve-regurgitation-2/ (visited on 10/10/2023).

[7] U. F. O. Themes, *Tricuspid Valve: Tricuspid Regurgitation*, 2016. [Online]. Available: https://thoracickey.com/tricuspid-valve-tricuspid-regurgitation/ (visited on 10/10/2023).

[8] L. S. O. Rezende, P. H. M. Sá, M. C. F. Macedo, A. L. Apolinário, I. Winkler, and M. A. Moret S. G., "Volume Rendering: An Analysis based on the HoloLens Augmented Reality Device," in *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, 2020, pp. 35–38. DOI: 10.1109/SVR51698.2020.00021.

[9] B. Arnaldi, P. Guitton, and G. Moreau, *Virtual Reality and Augmented Reality: Myths and Realities*. John Wiley & Sons, 2018.

[10] B. Cetinsaya, C. Neumann, and D. Reiners, "Using Direct Volume Rendering for Augmented Reality in Resource-constrained Platforms," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2022, pp. 768–769. DOI: 10.1109/VRW55335.2022.00235.

[11] G. D. Lopaschuk *et al.*, "Plasma fatty acid levels in infants and adults after myocardial ischemia," *American Heart Journal*, vol. 128, no. 1, pp. 61–67, 1994. DOI: 10.1016/0002-8703(94)90010-8.

[12] P. F. Kantor, J. R. Dyck, and G. D. Lopaschuk, "Fatty acid oxidation in the reperfused ischemic heart," *The American Journal of the Medical Sciences*, vol. 318, no. 1, pp. 3–14, 1999. DOI: 10.1097/00000441-199907000-00002.

[13] M. Karimi *et al.*, "Neonatal vulnerability to ischemia and reperfusion: Cardioplegic arrest causes greater myocardial apoptosis in neonatal lambs than in mature lambs," *The Journal of Thoracic and Cardiovascular Surgery*, vol. 127, no. 2, pp. 490–497, 2004. DOI: 10.1016/j.jtcvs.2003.07.052.

[14] J. S. Gammie *et al.*, "Safety and performance of a novel transventricular beating heart mitral valve repair system: 1-year outcomes," *European Journal of Cardio-Thoracic Surgery: Official Journal of the European Association for Cardio-Thoracic Surgery*, vol. 59, no. 1, pp. 199–206, 2021. DOI: 10.1093/ejcts/ezaa256.

[15] B. M. Cohn, K. Bartus, G. H. L. Tang, and T. C. Nguyen, "Finding the Future: The 10 Commandments of Beating Heart Mitral Valve Repair," *Innovations (Philadelphia, Pa.)*, vol. 15, no. 1, pp. 17–21, 2020. DOI: 10.1177/1556984519899638.

[16] G. Gerosa *et al.*, "Transapical off-pump echo-guided mitral valve repair with neochordae implantation mid-term outcomes," *Annals of Cardiothoracic Surgery*, vol. 10, no. 1, pp. 131–140, 2021. DOI: 10.21037/acs-2020-mv-86.

[17] G. Burström *et al.*, "Frameless Patient Tracking With Adhesive Optical Skin Markers for Augmented Reality Surgical Navigation in Spine Surgery," *Spine*, vol. 45, no. 22, pp. 1598–1604, 2020. DOI: 10.1097/BRS.0000000000003628.

[18] L. Greuter *et al.*, "Augmented Reality Fluorescence Imaging in Cerebrovascular Surgery: A Single-Center Experience with Thirty-Nine Cases," *World Neurosurgery*, vol. 151, pp. 12–20, 2021. DOI: 10.1016/j.wneu.2021.03.157.

[19] D. Harake *et al.*, "Stereoscopic Display Is Superior to Conventional Display for Three-Dimensional Echocardiography of Congenital Heart Anatomy," *Journal of the American Society of Echocardiography: Official Publication of the American Society of Echocardiography*, vol. 33, no. 11, pp. 1297–1305, 2020. DOI: 10.1016/j.echo.2020.06.016.

[20] L. L. Gheorghe *et al.*, "Imaging for Native Mitral Valve Surgical and Transcatheter Interventions," *JACC: Cardiovascular Imaging*, vol. 14, no. 1, pp. 112–127, 2021. DOI: 10.1016/j.jcmg.2020.11.021.

[21] A. K. Lim, J. Ryu, H. M. Yoon, H. C. Yang, and S.-K. Kim, "Ergonomic effects of medical augmented reality glasses in video-assisted surgery," *Surgical Endoscopy*, vol. 36, no. 2, pp. 988–998, 2022. DOI: 10.1007/s00464-021-08363-8.

[22] J. Li and C. Gu, "Off-pump mitral valve repair: Primary result of treating moderate ischemic mitral regurgitation during off-pump coronary artery bypass grafting," *Journal of Thoracic Disease*, vol. 11, no. 7, pp. 3191–3194, 2019. DOI: 10.21037/jtd.2019.07.15.

[23] J. P. Dal-Bianco *et al.*, "Active adaptation of the tethered mitral valve: Insights into a compensatory mechanism for functional mitral regurgitation," *Circulation*, vol. 120, no. 4, pp. 334–342, 2009. DOI: 10.1161/CIRCULATIONAHA.108.846782.

[24] V. H. Thourani *et al.*, "First in human experience with an epicardial beating heart device for secondary mitral regurgitation," *The Journal of Thoracic and Cardiovascular Surgery*, vol. 161, no. 3, 949–958.e4, 2021. DOI: 10.1016/j.jtcvs.2020.11.169.

[25] *Sensors | Free Full-Text | Microsoft HoloLens 2 in Medical and Healthcare Context: State of the Art and Future Prospects*. [Online]. Available: https://www.mdpi.com/1424-8220/22/20/7709 (visited on 11/07/2023).

[26] S. Park, S. Bokijonov, and Y. Choi, "Review of Microsoft HoloLens Applications over the Past Five Years," *Applied Sciences*, vol. 11, no. 16, p. 7259, 2021. DOI: 10.3390/app11167259.

[27] C. M. Morales Mojica *et al.*, "Holographic Interface for three-dimensional Visualization of MRI on HoloLens: A Prototype Platform for MRI Guided Neurosurgeries," in *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2017, pp. 21–27. DOI: 10.1109/BIBE.2017.00-84.

[28] M. Doughty and N. R. Ghugre, "Head-Mounted Display-Based Augmented Reality for Image-Guided Media Delivery to the Heart: A Preliminary Investigation of Perceptual Accuracy," *Journal of Imaging*, vol. 8, no. 2, p. 33, 2022. DOI: 10.3390/jimaging8020033.

[29] N. Costa *et al.*, "Augmented Reality-Assisted Ultrasound Breast Biopsy," *Sensors*, vol. 23, no. 4, p. 1838, 2023. DOI: 10.3390/s23041838.

[30] J. N. Costa *et al.*, "Ultrasound training simulator using augmented reality glasses: An accuracy and precision assessment study," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2022, pp. 4461–4464, 2022. DOI: 10.1109/EMBC48229.2022.9871406.

[31] T. Nguyen, W. Plishker, A. Matisoff, K. Sharma, and R. Shekhar, "HoloUS: Augmented reality visualization of live ultrasound images using HoloLens for ultrasound-guided procedures," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 2, pp. 385–391, 2022. DOI: 10.1007/s11548-021-02526-7.

[32] N. Cattari, S. Condino, F. Cutolo, M. Ferrari, and V. Ferrari, "In Situ Visualization for 3D Ultrasound-Guided Interventions with Augmented Reality Headset," *Bioengineering (Basel, Switzerland)*, vol. 8, no. 10, p. 131, 2021. DOI: 10.3390/bioengineering8100131.

[33] D. D. Maddali, H. Brun, G. Kiss, J. M. Hjelmervik, and O. J. Elle, "Spatial Orientation in Cardiac Ultrasound Images Using Mixed Reality: Design and Evaluation," *12*, 2022, Accepted: 2023-01-19T09:22:11Z Publisher: Frontiers Media. DOI: 10.3389/frvir.2022.881338.

[34] A. Fajar, D. W. Santoso, Z. R. Bahen, R. Sarno, and C. Fatichah, "Color Mapping for Volume Rendering Using Digital Imaging and Communications in Medicine Images," in *2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, 2021, pp. 1–5. DOI: 10.1109/AIMS52415.2021.9466012.

[35] *What is Stereoscopic Vision? - Tech-FAQ*, 2019. [Online]. Available: https://www.tech-faq.com/stereoscopic-vision.html (visited on 10/15/2023).

[36] L. Tremosa, *Beyond AR vs. VR: What is the Difference between AR vs. MR vs. VR vs. XR?* 2023. [Online]. Available: https://www.interaction-design.org/literature/article/beyond-ar-vs-vr-what-is-the-difference-between-ar-vs-mr-vs-vr-vs-xr (visited on 09/10/2023).

[37] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE Trans. Information Systems*, vol. vol. E77-D, no. 12, pp. 1321–1329, 1994. (visited on 09/10/2023).

[38] *Extended Reality*. [Online]. Available: https://www.dhl.com/global-en/home/insights-and-innovation/thought-leadership/trend-reports/augmented-and-extended-reality.html (visited on 10/10/2023).

[39] M. D. Osorto Carrasco and P.-H. Chen, "Application of mixed reality for improving architectural design comprehension effectiveness," *Automation in Construction*, vol. 126, p. 103 677, 2021. DOI: 10.1016/j.autcon.2021.103677.

[40] M. Lavik, *UnityVolumeRendering*, 2023. [Online]. Available: https://github.com/mlavik1/UnityVolumeRendering (visited on 11/05/2022).

[41] thetuvix, *Voice input in Unity - Mixed Reality*, 2021. [Online]. Available: https://learn.microsoft.com/en-us/windows/mixed-reality/develop/unity/voice-input-in-unity (visited on 10/05/2022).

[42] thetuvix, *Rendering - Mixed Reality*, 2022. [Online]. Available: https://learn.microsoft.com/en-us/windows/mixed-reality/develop/advanced-concepts/rendering-overview (visited on 10/20/2022).

[43] vtieto, *Use PC resources to power your app with Holographic Remoting remote app - Mixed Reality*, 2022. [Online]. Available: https://learn.microsoft.com/en-us/windows/mixed-reality/develop/unity/use-pc-resources (visited on 03/08/2023).

[44] J. Cecil, P. Ramanathan, M. Pirela-Cruz, and M. B. R. Kumar, "A Virtual Reality Based Simulation Environment for Orthopedic Surgery," in *On the Move to Meaningful Internet Systems: OTM 2014 Workshops*, R. Meersman *et al.*, Eds., Springer, 2014, pp. 275–285. DOI: 10.1007/978-3-662-45550-0_28.

[45] D. Escobar-Castillejos, J. Noguez, L. Neri, A. Magana, and B. Benes, "A Review of Simulators with Haptic Devices for Medical Training," *Journal of Medical Systems*, vol. 40, no. 4, p. 104, 2016. DOI: 10.1007/s10916-016-0459-8.

[46] Q. Lin *et al.*, "Immersive Virtual Reality for Visualization of Abdominal CT," *Proceedings of SPIE*, vol. 8673, 2013. DOI: 10.1117/12.2008050.

[47] mattatz, *Unity-volume-rendering*, 2023. [Online]. Available: https://github.com/mattatz/unity-volume-rendering (visited on 06/07/2023).

[48] U. Technologies, *Unity - Manual: Low-level Native Plugin Interface*. [Online]. Available: https://docs.unity3d.com/2018.1/Documentation/Manual/NativePluginInterface.html (visited on 03/07/2023).

[49] Hak0n, *Voice input - Mixed Reality*, 2022. [Online]. Available: https://learn.microsoft.com/en-us/windows/mixed-reality/design/voice-input (visited on 10/05/2023).

[50] keveleigh, *Speech - MRTK 2*, 2022. [Online]. Available: https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/features/input/speech?view=mrtkunity-2022-05 (visited on 10/05/2023).

[51] A. Pashaei, G. Piella, N. Duchateau, L. Gabrielli, and O. Camara, "Image-Based Estimation of Myocardial Acceleration Using TDFFD: A Phantom Study," in *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, O. Camara, T. Mansi, M. Pop, K. Rhode, M. Sermesant, and A. Young, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2014, pp. 262–270. DOI: 10.1007/978-3-642-54268-8_31.

[52] M. Brand, L. A. Wulff, Y. Hamdani, and T. Schüppstuhl, "Accuracy of Marker Tracking on an Optical See-Through Head Mounted Display," in *Annals of Scientific Society for Assembly, Handling and Industrial Robotics*, T. Schüppstuhl, K. Tracht, and D. Henrich, Eds., 2020, pp. 21–31. DOI: 10.1007/978-3-662-61755-7_3.

[53] A. Cao, A. Dhanaliwala, J. Shi, T. Gade, and B. Park, *Image-based marker tracking and registration for intraoperative 3D image-guided interventions using augmented reality*, 2019. DOI: 10.48550/arXiv.1908.03237.