AUGMENTED REALITY FOR MULTIMODAL ULTRASOUND-BASED BREAST BIOPSY

by

YASMEEN HASSAN

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

MASTERS OF SCIENCE IN MECHATRONICS AND ROBOTICS ENGINEERING

2025

Oakland University
Rochester, Michigan

Thesis Advisory Committee:

Alycen Wiacek, Ph.D., Chair
Khalid Mirza, Ph.D.
Jia Li, Ph.D.

ACKNOWLEDGMENTS

The author wishes to express her appreciation for extensive mentorship provided by Alycen Wiacek Ph.D, my adviser. Gratitude is also extended to committee members Khalid Mirza, Ph.D. and Jia Li, Ph.D. for their contributions to this dissertation. Special thanks are due to Michael Pasamba for constructing the custom water phantom used for system validation, Rehnuma Hasnat for developing the QUS processing code, and Andrew for for assistance with the AR system setup. The author also gratefully acknowledges the support of Oakland University's Augmented Reality Center (ARC), particularly for their help in configuring and debugging the HoloLens environment. This work was financially supported by the ORAU Ralph E. Powe Junior Faculty Enhancement Award. Additional funding was provided by The Augmented Reality for Enterprise Alliance (The AREA).

Yasmeen Hassan

ABSTRACT

AUGMENTED REALITY FOR MULTIMODAL ULTRASOUND-BASED BREAST BIOPSY

by

Yasmeen Hassan

Adviser: Alycen Wiacek, Ph.D.

Augmented Reality (AR) technologies have been demonstrated to enhance image-guided medical procedures by improving real-time visualization and spatial understanding. Ultrasound, in particular, is often integrated with AR due to its real-time functionality and portability. However, most ultrasound-based AR systems are limited to traditional B-mode ultrasound alone, visualizing the tissue morphology, but lacking mechanical or microsctructural properties of tissue. Newer modalities such as elastography and Quantitative Ultrasound (QUS) can provide these additional properties to improve diagnosis. This thesis presents an AR platform that integrates multimodal ultrasound imaging, including elastography and QUS, into a spatially registered system for biopsy guidance. The system combines a Clarius handheld ultrasound probe with a Microsoft HoloLens 2 using Unity, OpenCV, and marker-based tracking. A Qt and MATLAB-based data pipeline supports streaming and spatial alignment of multiple imaging modes. System evaluation on tissue-mimicking phantoms showed registration errors of 2.8mm for B-mode at shallow depths, 10.8mm for B-mode in deeper tissues, and 17.0mm for elastography. Latency ranged from 159ms using B-mode to 167s for QUS imaging. These results demonstrate the feasibility of the proposed AR platform for ultrasound-guided interventions driven by both morphology and diagnostic information,

while highlighting areas for future improvement in latency and registration accuracy and provides a foundation for future innovations in AR-based multimodal breast biopsies.

TABLE OF CONTENTS

TABLE OF CONTENTS—Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	Application Programming Interface
AR	Augmented Reality
HMD	Head Mounted Device
IQ	In-phase and Quadrature
LAN	Local Area Network
OST	optical see-through
QUS	Quantitative Ultrasound
RF	Radio Frequency
SDK	Software Development Kit
SoS	Speed of Sound
TCP	Transmission Control Protocol
UI	User Interface
US	Ultrasound
VST	video see-through
YAML	Yet Another Markup Language

CHAPTER ONE

INTRODUCTION

AR has become increasingly popular across various disciplines in recent years, particularly in healthcare, where technological advances are being integrated into clinical practice [2]. Among these advancements, the use of AR with medical imaging techniques is revolutionizing the way clinicians interact with and interpret diagnostic images. In radiology, AR systems have been explored to enhance spatial understanding, improve workflow, and provide real-time visualization support during procedures [3].

One key modality within radiology is ultrasound, which is valued for its real-time feedback, portability, and non-invasive nature. Ultrasound-guided procedures, in particular breast biopsies, offer a minimally invasive approach to obtain valuable information on the patient's diagnosis and plan subsequent treatment. The procedure consists of inserting a biopsy needle into tissue to obtain a sample from a specific area of concern within the breast. When guided by ultrasound, the physician first finds the biopsy target in the ultrasound image. Next, the biopsy needle is inserted and navigated toward the target; however, the ultrasound image used for this navigation is on an external monitor, requiring the physician to look up from the biopsy needle, to the monitor, and back again [4].

AR-assisted systems have been proposed to remove the reliance on this external monitor and assist in visualizing the target area in real-time within the physician's field of view [5]. These systems can overlay the ultrasound image directly onto the physical space, enhancing the surgeon's spatial awareness and improving accuracy during the procedure. By integrating AR with ultrasound guidance, these systems may offer several advantages, including reduced procedure time, minimized risk of error, and enhanced comfort for the patient and healthcare provider [6]. Additionally, AR can allow the physician to focus on

the procedure while maintaining situational awareness of both the target and surrounding anatomy, potentially reducing cognitive load and the need for constant adjustments [5].

Despite these advancements, several challenges remain. Current AR systems for ultrasound-guided procedures typically rely on conventional grayscale B-mode ultrasound images [7]. While useful to visualize the morphology of the underlying tissue, these systems often overlook the diagnostic information available through multimodal ultrasound imaging, such as Doppler flow, elastography, QUS, or contrast-enhanced modes [8, 9]. Elastography is most commonly used to measure tissue stiffness to differentiate between benign and malignant lesions. Elastography paired with B-mode ultrasound has been shown to reduce false positive diagnoses [10]. Furthermore, QUS techniques such as Speed of Sound (SoS) analysis, attenuation and backscatter coefficient estimation, envelope statistics, and spectral parameters further enhance diagnostic power by evaluating acoustic and microstructural properties of tissue [11], adding to the diagnostic value of ultrasound. Together, these additional ultrasound-based modalities provide complementary views of tissue characteristics, vascular structures, and stiffness which are crucial for accurate targeting in biopsies, providing both morphological and diagnostic information [10, 12].

Despite their value, these advanced imaging modalities remain underutilized in AR environments due to differences in data formats such as grayscale or color imaging, spatial alignment, and real-time rendering requirements. In addition to visualization challenges, aligning and synchronizing multimodal ultrasound data within an AR environment requires precise calibration and efficient data handling to preserve diagnostic integrity. These limitations have led most AR solutions to prioritize spatial guidance over rich diagnostic support [4, 13]. The lack of robust frameworks to support these advanced imaging modes in AR applications represents a critical gap in current research and development.

This gap presents an opportunity to design an AR system that supports not only spatially aware guidance but also rich diagnostic imaging integration for biopsy procedures. Therefore, the overall goal of this thesis is to demonstrate the feasibility of integrating multimodal ultrasound imaging into an AR-guided biopsy system, specifically for breast cancer applications. By combining spatial AR guidance with functional data-overlays, the system aims to improve targeting precision, reduce diagnostic ambiguity, and support better clinical outcomes. The overall system workflow and contributions of this thesis are illustrated in Fig. 1.1, highlighting how multimodality information is integrated and displayed wirelessly in AR to enhance biopsy guidance.

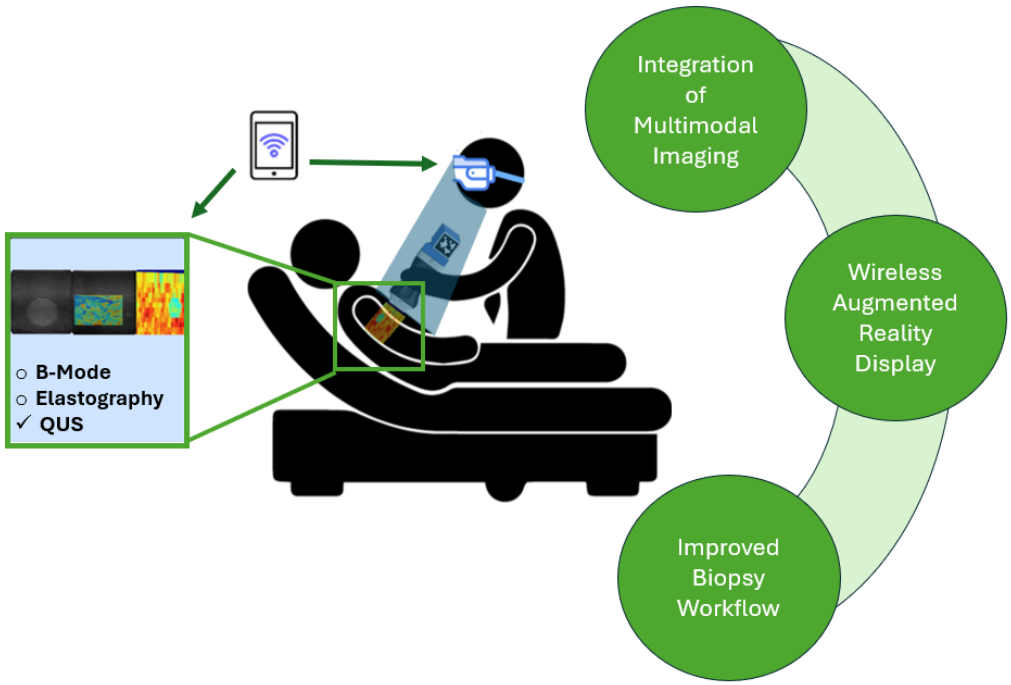

Figure 1.1: Overall workflow of the system showing multimodal data integration, wireless augmented reality display, and improved biopsy workflow.

The remainder of this thesis is organized as follows: **Chapter 2** reviews the current literature, **Chapter 3** discusses the methodology, system design, and comprehensive evaluation setup, **Chapter 4** describes the results, **Chapter 5** discusses the implications of the developed system including clinical implementation opportunities, and **Chapter 6** concludes with a discussion of future work.

CHAPTER TWO

LITERATURE REVIEW

This chapter reviews recent advancements in AR-assisted ultrasound-guided biopsies, examining their implementation strategies, clinical validation, and inherent limitations in order to identify open challenges and opportunities in AR-guided ultrasound interventions. A summary of representative systems that will be discussed in more detail below is provided in Table 2.1.

Underlying these visualization systems are two principal tracking approaches that tackle the challenges of simultaneous image registration and depth perception in AR environments. The outside-in method [14] uses external tracking systems such as optical cameras or infrared sensors to monitor the position and orientation of both the AR headset and the ultrasound probe. In contrast, the inside-out method [15] relies on integrated sensors and onboard cameras within the AR headset and probe to estimate their relative position from the user's point of view. These two paradigms offer distinct advantages in terms of setup complexity, environmental constraints, and tracking fidelity.

2.1 Inside-Out Method

The major advantage of the inside-out method is its portability and reduced setup complexity, with each implementation offering unique features to enhance procedural guidance. The general setup of a system that uses the inside-out method is seen in Figure 2.1. For example, Costa *et. al* [16] aimed to improve ultrasound breast imaging through real-time tracking of the biopsy needle in the field view using QR codes affixed to the ultrasound probe and biopsy needle. After gathering the size of the necessary coordinates and the size of the QR code in the real-world, a virtual square hologram with the exact dimensions and orientation can be seen via the AR glasses to confirm to the user that the ultrasound probe is being actively tracked. Costa *et. al* also included active needle

Table 2.1: Summary of methods for AR in ultrasound

Reference	Application	Unique Implementation	Performance
Inside-Out Tracking Methods			
Costa et al. [16]	Biopsy guidance	QR code tracking of probe and needle	Improved needle alignment
Haxthausen et al. [15]	Biopsy guidance	Reflective spheres on probe, HoloLens tracking	Low latency, image-probe alignment
Nguyen et al. [6]	Needle training	Hands-free (voice) + AR marker tracking	Novices completed 17% faster
Sato et al. [17]	Tumor model overlay	3D tumor reconstruction + rigid body and LED tracking	1.5 FPS and demonstrated clinically
Li et al. [18]	Biopsy guidance	Infrared mechanical trackers	122 ms latency and 95-98% punctuation accuracy
Outside-In Tracking Methods			
Kanbara et al. [13]	Ultrasound overlay	Stereo cameras + external markers	10 FPS with low registration error
State et al. [4]	Biopsy guidance	Hybrid magnetic, fluorescent, and mechanical tracking	Demonstrated on phantom
Cattari et al. [10]	Novice training	Hybrid vision + mechanical tracking	All participants accuracy <5mm
Ruger et al. [7]	Biopsy guidance	Reflective markers + 3D camera system	Decreased placement error by 2.5 mm

guidance for the clinician using the system by placing a red circle around the lesion center in the hologram along with a green line tracking from the needle tip during alignment or a red line during misalignment. This assisted the user in aligning the biopsy needle with the ultrasound image plane based on the measured angle between the needle tip and the lesion position [16]. The time latency between the ultrasound image being generated in the probe to that image being displayed in the AR system was tested by marking 2 distinct positions on a breast phantom and comparing the image's timestamps. The low latency of 16 ms between the probe's image generation and the AR system's display supports the effectiveness of this approach. A usability assessment was also performed among various

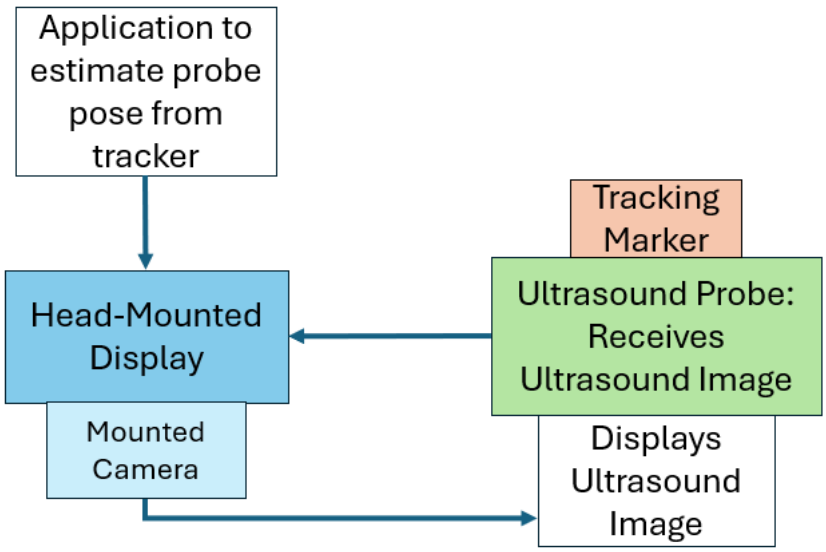

Figure 2.1: Inside-Out method employed using QR tracking

users and found that users reached the lesion more accurately while also finding that needle alignment was more precise with 92.67 degrees in the prototype and 89.99 degrees with the traditional system.

Similarily, Li *et. al* [18] used a retro-reflective tracking marker attached to the probe in order to simultaneously track the probe and the biopsy needle. Using a Unity rendering environment and the Microsoft HoloLens2, navigation information was provided for both in-plane and out-of-plane punctuation. Ultrasound images streamed to the HoloLens 2 had a latency of 122.5 ms with a user study resulting in 98% success rate for out-of-plane and 95% success rate for in-plane punctuation for novice users [18].

Another attempt at tracking the ultrasound probe was made by Haxthausen *et. al* [15] by using sensors and cameras that are embedded in the AR glasses (HoloLens 2) to track reflective spheres placed on the ultrasound probe. This work adopted a similar approach to Costa *et. al*[16] where instead of measuring the position of the QR code with

respect to the 3D coordinate system, the position of each sphere (using three in total) were measured via the depth camera.This resulted in the alignment of the 2D cross-sectional ultrasound images onto the 3D anatomy of the patient viewed on the HoloLens to the clinician. A drawback to this application, and many other applications that utilize the inside-out method, is the use of internal hardware such as depth cameras and inertial measuring units helps to maintain an accurate position of the probe.

Nguyen *et. al* [6] adopted a similar inside-out approach using physical markers for object tracking, utilizing the Microsoft HoloLens to virtually render and register the ultrasound image with a transducer. In the physical realm, the ultrasound transducer was tracked using two large AR markers for accuracy. Based on the 3D pose of these markers, the AR pattern was detected using the ARToolKit library. Although this resulted in a similar output to other works mentioned above, this work was unique in the way that the HoloLens voice command feature was utilized for hands-free procedure capabilities in order to maintain a sterile environment. The use of AR markers attached to the ultrasound transducer allowed for the limited additional hardware and gave this application the benefits of the inside-out method without the additional disadvantage of limited movement. Since the 3D pose was actively tracked, the ultrasound transducer did not have to constantly be in view of the HoloLens world-facing camera, as all variations in movement were processed by the ARToolKit. In order to test the efficiency of this system, a comparison of time to task completion was performed between novices and experts of ultrasound probe by using a phantom for the ultrasound-guided needle tasks. Novices completed 17% faster using the HoloLens compared to without it; however, experienced users were only 5% faster compared to the conventional approach.

Sato *et. al* [17] made a significant departure by using pre-surgical 3D tumor models instead of real-time ultrasound images. This approach uses rigid bodies attached to the ultrasound probe as its tracking system to precisely locate the position and

orientation of each cross section of the ultrasonic image. With the combination of a 3-D optical sensor, rigid body tracker and a video camera system the AR system takes the cross-section view of each 2-D live ultrasound image and performing volume composition to reconstruct a 3-D tumor model onto the field of view. The tracking in this system is done by attaching physical LED markers to the video camera (used as a reference frame) and probe to measure their exact orientation in the world coordinate system. Unlike previous studies, a 3-D model is displayed in the augmented reality space; therefore, the acquired image had to undergo further filtering to extract the volume composition of the tumor. Several experiments were performed to validate the use of pre-acquired images in the augmented scene including clinical validation showing 1.5 FPS for volume visualization.

2.2 Outside-In Method

The outside-in method allows for creativity in hardware options, marker tracking, and AR displays as shown in Figure 2.2. One early study sought to exploit this creative freedom by using an external stereovision sensor placed in the operating room and a video see-through (VST) Head Mounted Device (HMD) to display a superimposed image on the user's view of real environment [13]. Although dated, this work remains notable for demonstrating one of the first integrations of stereovision sensing with AR visualization for medical applications. The combination of these two technologies allow for the synchronization of display-timing between virtual and real worlds to reduce alignment error, which is significant in the case of maintaining correct registration of both worlds in real-time. To accomplish this precise registration, the stereoscopic cameras are mounted onto the HMD and the depth of the real world is estimated by using the camera's parameters. The camera parameters are found using marker-based tracking; by placing three blue markers in the real-world, feature points of the real environment are captured in a sequence of images. The resulting system is a video see-through AR system that

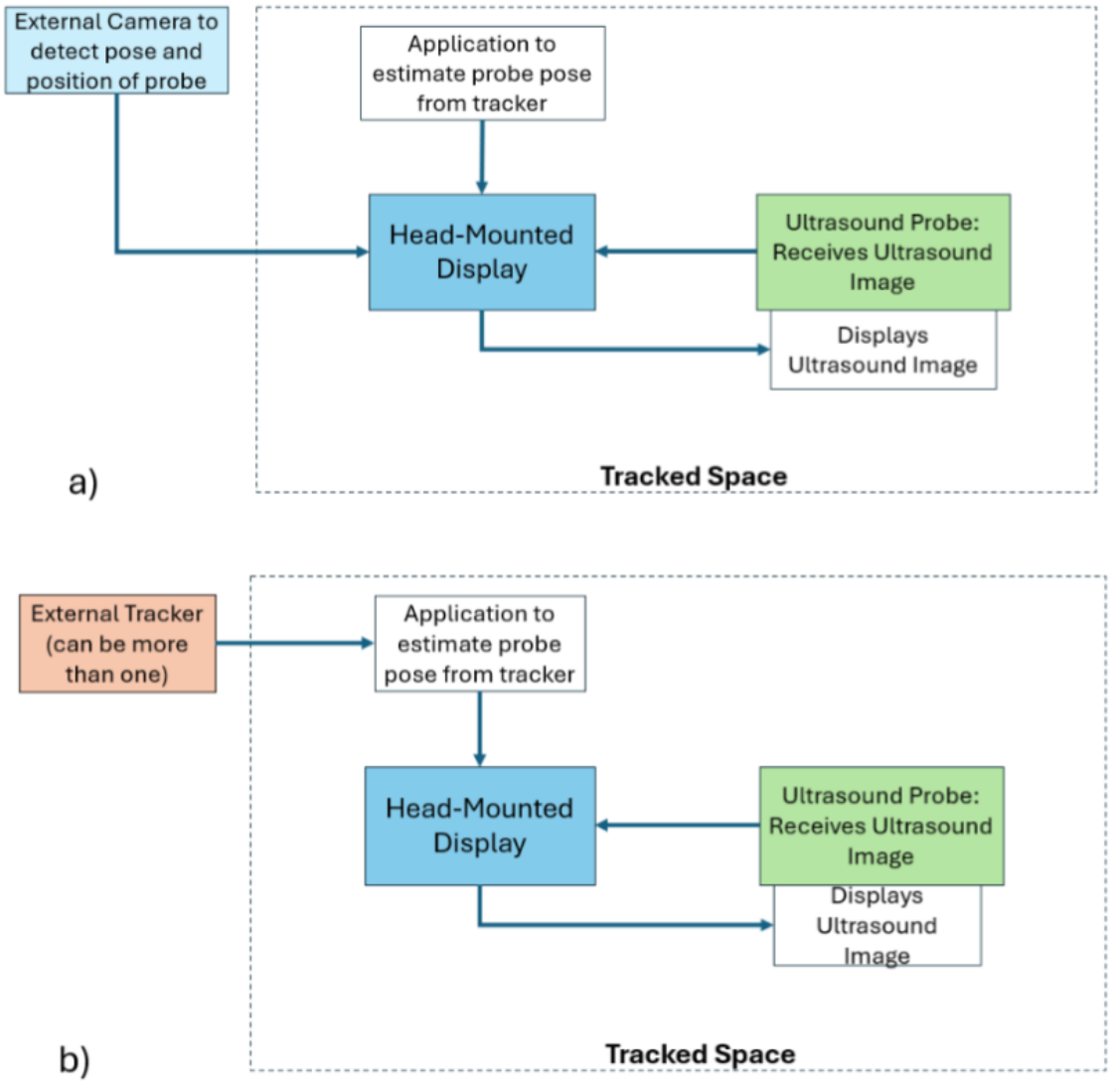

Figure 2.2: Outside-In setup using either (a) an external camera to track the position and orientation of the ultrasound probe to communicate with the display or (b) external marker(s) placed in various locations to estimate the probe's pose.

consists of a pair of stereo cameras mounted on the HMD that registers an average frame rate of 10 frames per seconds or 100 milliseconds per frame. This quick image registration allows for an accurate real-time tracking of the object in the AR system. While this registration speed was impressive for its time, automatic detection and tracking

of natural feature points in a real scene cannot be employed without the predefined markers that the inside-out method provides.

State *et. al* [4], though now over two decades old, remains influential for its introduction of volumetric ultrasound visualization in AR with a physician-focused goal. To track the probe, a combination of vision and mechanical tracking is performed to identify landmarks through the video see-through system. A magnetic tracker is used for the physician's head using a HMD with vision-based landmark tracking of the work environment via fluorescent discs to reduce interference with magnetic elements in the room. A six degree-of-freedom mechanical tracker is placed on the probe for ease in tracking probe position; however, this introduces a limited range of probe motion as the probe is tethered to the ultrasound machine. The overall inputs to this system are an ultrasound scanner video, HMD stereo cameras, mechanical tracker on the probe, and a magnetic tracker on the head of the physician. These inputs result in the final output of a stereo AR image of the patient on the HMD with volume visualization. The combined use of vision and mechanical tracking is advantageous due to improved positional and orientation accuracy of the probe and patient. While hardware constraints limited clinical usability, the hybrid tracking approach laid important groundwork for today's multimodal tracking systems.

For the skilled physician, the needle-guided biopsy does not require a drawn out path of trajectory. However, in order to train the novice, a path of needle trajectory can significantly reduce errors, improve accuracy, and efficiently train the novice to skillfully locate the location of the biopsy. The aim of Cattari *et. al* was to use the AR headset in conjunction with a 3D probe to display the 3D model of the target superimposed on the patient with two cross hairs that will guide the user in the puncture and insertion of the needle [19]. Similar to the way [4] combined the use of a mechanical tracker on the probe, [19] tracked the probe's motion by using three fluorescent green markers attached to the

probe. The hardware, although an AR headset, was unique in the way that it exploits both optical see-through (OST) and VST methods. The software, an open source optical tracking software, provided the registered *in situ* visualization of both optical and video based augmentation. A user study was conducted to quantify the accuracy of the AR system by including ten participants, all inexperienced, to line up the viewpoint of the HMD so the circular cross-hair was inscribed, point the tip of the needle at the center of two cross-hairs and align the syringe along the trajectory by means of the cross drawn on the bottom of the syringe. Four different sessions were conducted and it resulted in all participants achieving accuracy within 5 mm with 40 percent achieving accuracy within 1.5 mm.

Rüger *et. al.* [7] focused on the novice or physician-in-training by attempting to create an AR-HMD system that results in "reduced workload, fewer complications, or shorter interventions" [7]. The system consisted of a surgical navigation system, 3D camera and ultrasound system directly communicating to the Microsoft HoloLens. Reflective markers were attached to the ultrasound transducer and the HoloLens to track the position and orientation of the ultrasound image relative to the tracker on the transducer and the projected image. The calibration of the tracker was done using proprietary software, and its data along with the real-time tracking from the 3D camera were used to generate a transformation matrix of the ultrasound image relative to the HMD tracker. Participants with varying skill levels used the AR-HMD system to perform needle placement on a training model. Placement error, task duration, and perceived workload were compared, demonstrating greater placement accuracy with a mean error decreasing from 7.4 mm to 4.9 mm, but no significant change in task duration. The qualitative study resulted in the participants reporting that they were unfamiliar with the system, but would like to test the technology in practice.

2.3 Challenges and Limitations

While each of the examples described in Sections 2.1 and 2.2 represent significant advancements in AR-based ultrasound-guided biopsy procedures either in novice training or in procedure accuracy, none of the existing applications introduce multimodality information into their systems, relying on only B-mode ultrasound. One key limitation of relying only on B-mode ultrasound is that it primarily provides morphological images by reflecting differences in acoustic impedance but does not provide direct quantitative information about tissue microstructure or mechanical properties. While B-mode remains highly effective for visualizing tumors, small lesions, and general anatomical features, it offers limited contrast when differentiating between subtle variations in soft tissue properties [20]. This lack of contrast may lead to difficulty in identifying clinically significant changes in tissue properties, potentially resulting in missed biopsy regions and suboptimal clinical outcomes.

To understand this limitation, it is useful to understand how B-mode images are generated. Ultrasound imaging begins by sending high-frequency sound waves from the probe into the body, with echoes returning from tissue boundaries. These echoes are first recorded as raw Radio Frequency (RF) signals which are the most detailed form of ultrasound backscatter [9]. The RF signal is then demodulated into in-phase (I) and quadrature (Q) components, collectively called IQ data, which represent the amplitude and phase of the echo at baseband. From the In-phase and Quadrature (IQ) data, the envelope (magnitude) is computed, followed by log compression and quantitization to produce the grayscale B-mode image [9]. Thus the familiar brightness-based B-mode image is ultimately a compressed representation of the original RF data.

QUS, a category of techniques that quantitatively assesses tissue properties such as microstructure and texture, can improve B-mode imaging techniques by providing additional information on tissue properties extracted from the raw data [21]. Strain

elastography measures tissue stiffness and improves tumor detection and procedural guidance by tracking tissue movement as a result of a mechanical disturbance, visualizing differences in mechanical properties [10]. These modalities provide diagnostic information beyond what is visible in conventional B-mode imaging, allowing for improved tissue characterization and more accurate targeting in interventional procedures [11]. Despite their proven clinical benefits, QUS, elastography, and other advanced modalities remain largely absent from AR guidance systems. By combining multimodal ultrasound imaging into the AR environment alongside B-mode images, we predict that clinicians will gain a more comprehensive understanding of both morphology and microstructure, ultimately improving biopsy target localization and clinical outcomes. This approach has the potential to address current limitations in AR-guided procedures and pave the way for more effective interventions.

CHAPTER THREE

METHODS

3.1 Overview

The core objective of this system is to integrate multimodal ultrasound imaging into an AR interface for guiding biopsy procedures with improved spatial awareness and diagnostic feedback. The developed system superimposes real-time ultrasound data in the clinician's field of view using the Microsoft HoloLens 2, enabling hands-free, in-situ visualization. Figure 3.1 shows an overview of the system setup. Specifically, ultrasound data from a wireless Clarius probe is transmitted via wireless Local Area Network (LAN) to a companion PC, which processes and relays the image stream to the HoloLens application built in Unity. Spatial tracking of the probe is achieved through marker-based computer vision techniques using ArUco markers and OpenCV, and spatial alignment is maintained by applying a fixed offset transformation between the tracked marker and the ultrasound imaging plane. Two secondary diagnostic modalities are available: (1) elastography and (2) QUS. Elastography is activated from within the Clarius Mobile application and automatically streamed to the HoloLens. Parametric QUS images are requested through a button press in the HoloLens user interface, which triggers data to be downloaded, processed via MATLAB and streamed back to the HoloLens.

3.2 Hardware and Software Setup

3.2.1 Hardware

Table 3.1 provides an overview of the hardware components and their functions in our custom multimodality AR setup. The developed system runs directly on the Microsoft HoloLens 2, a self-contained mixed reality headset equipped with integrated depth sensors, inertial measurement units (IMUs), and visible light cameras. Research Mode

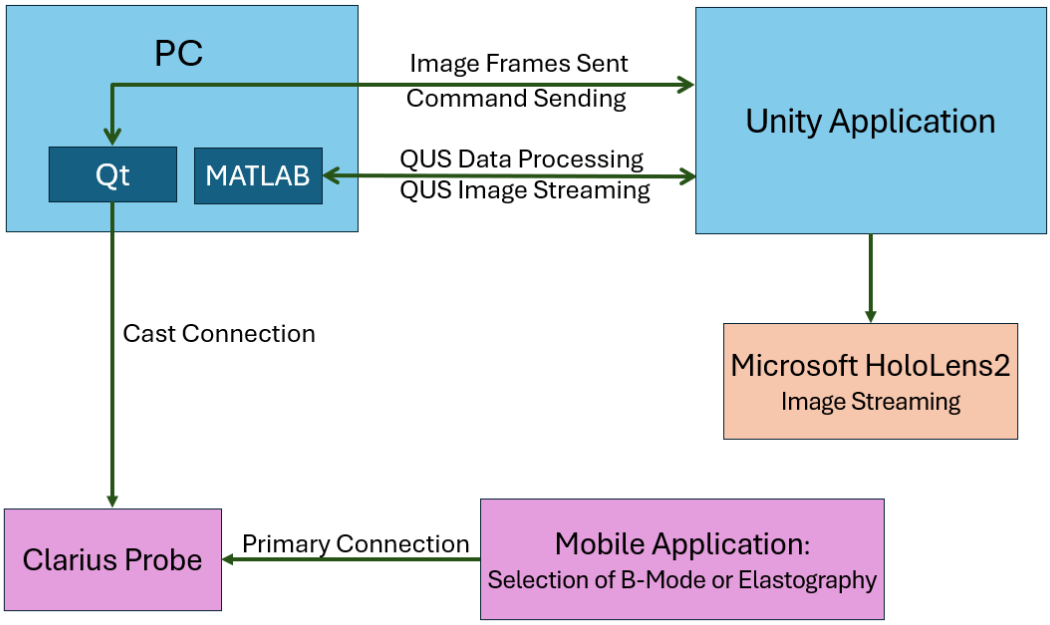

Figure 3.1: Overview of the AR-guided multimodal ultrasound system architecture. The Clarius wireless ultrasound probe streams B-mode, elastography, and raw data over Wi-Fi to a PC running a custom Qt application and MATLAB-based quantitative ultrasound (QUS) processing pipeline. Processed image frames and QUS results are then streamed via TCP to the HoloLens 2 running a Unity application. Unity handles real-time AR visualization, spatial tracking via OpenCV-based ArUco marker detection, and user interface interactions within the AR environment.

Table 3.1: Overview of hardware components and their purposes

Hardware	**Purpose**
Clarius Probe	Acquires ultrasound images and transmits them to the PC using Clarius's internal API over Wi-Fi.
PC	Executes a custom Qt application and MATLAB executable script to stream image frames over TCP to the Unity application on the HoloLens and process the QUS images.
HoloLens 2	Executes a Unity app to receive the image stream, detect ArUco markers using camera feeds, and render the ultrasound images as textured overlays aligned with the tracked marker pose.
Marker Cube	Affixed to the probe and provided a spatial reference for aligning ultrasound images in 3D.

was enabled to access low-level sensor streams including the IR-reflectivity feed and spatial environment tracking data, allowing for robust marker detection and spatial mapping [22]. The device used in this study was running Windows Holographic OS version 23H2.

Ultrasound data was acquired from a Clarius L7 HD3 wireless ultrasound probe on software version 12.0.2. The probe supports multiple imaging modes including B-mode (grayscale), color Doppler, and elastography, and is equipped with the research package, providing access to raw IQ or RF data. To establish communication, the probe must first connect to the Clarius mobile application, which initializes the device and activates the internal Application Programming Interface (API) services. After this initial connection, the probe communicates over a dedicated Wi-Fi network to a local PC, which acts as a proxy to stream ultrasound image frames and raw data to the HoloLens 2 for AR visualization.

For tracking, a rigid 3D-printed cube was mounted at the base of the ultrasound probe. The cube had 5 cm sides and featured printed 6×6 ArUco markers (ID 0–3) on its faces to enable stable detection regardless of probe orientation [23]. Marker visibility was optimized by aligning the cube's face normals with expected camera viewing angles during scanning [24].

The off-board processing unit was a Windows 11 laptop equipped with an Intel Core i9-13950HX CPU, 32 GB of RAM, and an NVIDIA RTX 4060 GPU. This computer ran the Clarius Caster application (compiled using Qt 6.5), managed all Transmission Control Protocol (TCP) server operations, and optionally launched MATLAB for offline processing of raw ultrasound data such as QUS analysis. The PC connected to the Clarius probe over Wi-Fi and forwarded ultrasound frames to the HoloLens in real-time.

3.2.2 Software

The software system was composed of four main components: (1) the Unity application on the HoloLens 2, (2) a custom Caster application using the Qt framework, (3) a MATLAB script for offline analysis and (4) supporting utilities for OpenCV-based tracking.

3.2.2.1 Unity Application

The user interface of the AR visualization was implemented in Unity 2022.3.62f, compiled for the Universal Windows Platform (UWP) and deployed to the HoloLens 2. The application used the OpenCV for Unity 2.6.6 plugin to perform marker-based tracking via the ArUco module [25]. The tracking relied on a calibrated pinhole camera model using intrinsic parameters derived from checkerboard-based calibration using the HoloLensCamCalib toolkit [26]. Real-time pose estimation was performed using the `solvePnP()` function to estimate the 6DOF transform from the marker to the camera.

In Unity, ultrasound images were received from a TCP stream and rendered onto a `RawImage` component placed on a World Space canvas. The pose of this canvas was

dynamically updated to align with the detected marker pose. To improve visualization clarity, the canvas orientation was adjusted based on the user's viewpoint (e.g., front vs. side perspective) and resolution scaling was additionally supported.

The Unity User Interface (UI) included several Canvas-based controls rendered in 3D world space, as shown in Figure 3.2. These controls allowed the user to:

- Toggle the ultrasound stream (freeze/unfreeze)
- Control imaging parameters (e.g., depth)
- Trigger raw image capture from Clarius
- Initiate MATLAB-based QUS analysis

These controls were implemented using Unity's `Button`, `TextMeshProUGUI`, and `Image` components and integrated with the real-time tracking and rendering pipeline. These controls allowed hands-free operation in a clinical or research setting, improving ergonomics and maintaining sterility during probe use.

The application displayed on a 2D window within the HoloLens 2 display, from there it can be aligned to the ultrasound probe and either enlarged or shrunk in size based on the user's preference.

3.2.2.2 Qt Caster Application To access the ultrasound data, the Clarius Caster Software Development Kit (SDK) (v12.0.2) was compiled into a custom application using Qt 6.5. This application established a local connection to the Clarius probe over Wi-Fi and streamed frames (in raw pixel buffer format) to the Unity application via a TCP socket. This application acted as a server with image metadata passed alongside each frame.

3.2.2.3 MATLAB Processing For generation of the parametric QUS images, MATLAB R2024b was used to process raw ultrasound data from the Clarius probe, including the generation of QUS parametric maps. A TCP server was implemented using MATLAB's `tcpserver()` interface, allowing Unity to issue commands (e.g., `RUN`, `EXIT`) for

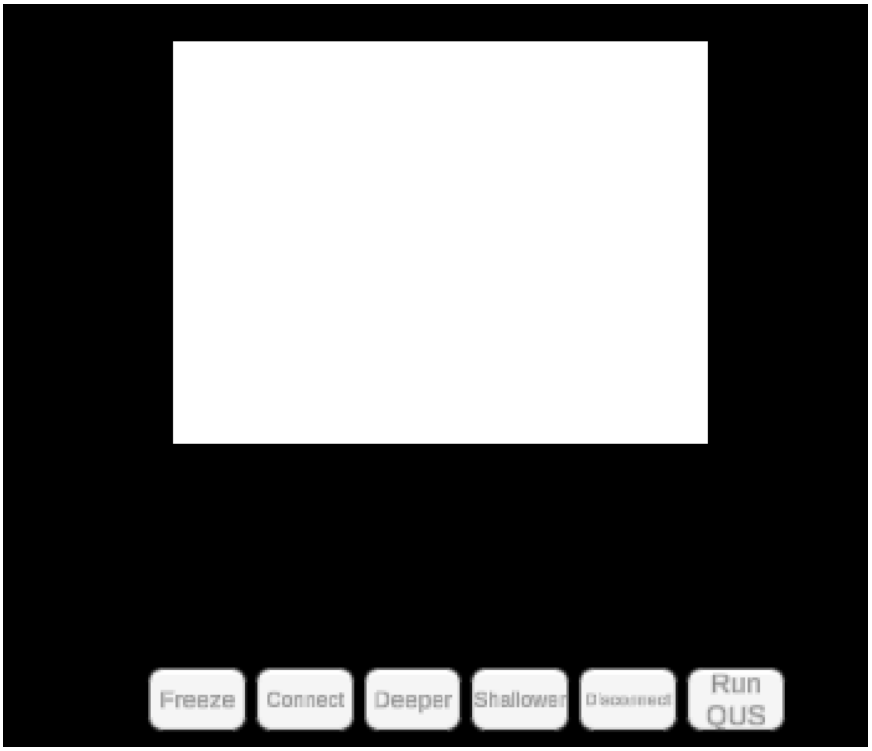

Figure 3.2: Unity UI with RawImage component and buttons

on-demand processing. The results of the MATLAB processing was then sent back to Unity for display on user demand.

3.2.2.4 Data Flow Integration The full software pipeline relied on synchronized data flow across devices:

- The Clarius probe streamed ultrasound frames to the PC via Wi-Fi.
- The Qt Caster server streamed these frames over TCP to the HoloLens 2.
- The Unity application decoded, textured, and aligned the images using ArUco marker tracking.
- The raw data was saved and sent to MATLAB for further processing, with results returned via TCP.

All inter-process communication was handled using TCP sockets for the availability of offline usage, and Unity's threading model was implemented through background image decoding processes and main-thread dispatch of rendering and UI updates [27].

3.2.2.5 Ultrasound Image Integration Pipeline Ultrasound image frames were rendered on the world canvas object and dynamically followed the pose of an ArUco marker attached to the ultrasound probe. A fixed spatial offset was applied to the canvas position so that the ultrasound image appears in virtual space "hovering" just in front of the patient, aligned with the probe's imaging plane and scaled based on the position of the canvas.

3.2.3 Parametric QUS Image Processing

To create parametric QUS images requires access to the the raw data, which was downloaded from the Clarius probe, transferred to the PC and processed offline in MATLAB. This data is in the form of in-phase and quadrature (IQ) signals prior to the application of post-processing steps that would typically be applied to reach a B-mode ultrasound image. For initial feasibility demonstration, we chose to create parametric Nakagami images [28, 29, 30]. However, using this pipeline, many types of QUS images could be implemented instead. The Nakagami distribution characterizes the scattering properties and heterogeneity of tissues [28]. The Nakagami distribution is defined by the probability density function:

$$p(r) = \frac{2m^m}{\Gamma(m)\Omega^m} r^{2m-1} \exp\left(-\frac{m}{\Omega} r^2\right), \quad r \geq 0 \tag{3.1}$$

where r is the envelope amplitude of the backscattered ultrasound signal, m is the Nakagami shape parameter describing scattering conditions, Ω is the scaling parameter, and $\Gamma(\cdot)$ is the Euler Gamma Function. Nakagami parametric images enhance tissue

characterization by highlighting variations in density and organization, useful in detecting pathologies such as tumors or fibrosis [30].

The specific steps to acquire and process the Nakagami parametric images are as follows:

1. **Data Acquisition and Decompression:** When a raw ultrasound data capture is triggered on the Clarius system, the device saves the data as a compressed archive file in `.tar` format. This archive contains both the raw IQ ultrasound data and accompanying header metadata necessary for processing. To optimize storage and transmission efficiency, Clarius compresses the raw IQ data using the Lempel-Ziv-Oberhumer (LZO) compression algorithm. While this significantly reduces file size, it requires decompression of data prior to analysis. Therefore, the MATLAB processing script first scans the designated raw data directory and identifies the most recently saved `.tar` archive. It then extracts the archive's contents, which include the compressed IQ data file (typically with a `.lzo` extension) and its associated metadata header file. The script utilizes an external decompression utility (`lzop.exe`) to decompress the LZO-compressed file, restoring the IQ data to its original raw binary format.

2. **Metadata Extraction:** The Yet Another Markup Language (YAML) configuration file associated with the raw data is read to extract imaging parameters such as depth, sampling rate, probe element count, and pitch. These parameters are used to define spatial axes for image display.

3. **Reading Raw IQ Data:** The header metadata is read to determine the number of frames. The script loads the final frame's IQ data, separating in-phase (I) and quadrature (Q) components to compute the envelope amplitude image:

$$\text{Env\_img} = \sqrt{I^2 + Q^2} \tag{3.2}$$

4. **Nakagami Parameter Estimation:** The Nakagami distribution is fitted to local windowed segments of the envelope data to estimate the shape parameter m, which reflects tissue scattering properties. This is performed via a sliding window approach to create a parametric image [29]. Specifically, a rectangular window of size 105×15 pixels is moved through the envelope image with 50.5% axial overlap and 53% lateral overlap. Each windowed region is multiplied by a Hanning window in order to taper edges smoothly to zero, reducing the influence of outlier pixels at the boundaries and emphasizing the central pixels, where the signal is more consistent [30]. The envelope data within the window is fitted to a Nakagami distribution using MATLAB's `fitdist` function, extracting the shape parameter m (`mu` in MATLAB). This value is then assigned to the entire window and the process is repeated as the window is translated throughout every possible axial and lateral position. The final resulting image is called a parametric Nakagami image.

5. **Output and Visualization:** The resulting parametric Nakagami image is stored in a matrix of the same spatial dimensions as the input ultrasound raw data image. This is then saved as a PNG image in the Unity project's resource folder for later visualization within the AR application.

3.3 Tracking and Registration

The system achieves spatially aligned AR visualization of ultrasound frames on the HoloLens 2 by combining intrinsic camera calibration and spatial calibration with marker-based tracking in Unity's 3D rendering pipeline.

3.3.1 Intrinsic Calibration

To estimate the marker's pose in real time, the HoloLens 2 photo-video (PV) camera is used to capture RGB frames. The complete calibration process is illustrated in Figure 3.3. These images are first undistorted using a precomputed camera intrinsic

matrix $\mathbf{K} \in \mathbb{R}^{3\times3}$ and distortion coefficient vector $\mathbf{D} \in \mathbb{R}^{1\times5}$, obtained via checkerboard calibration [26]. Let the intrinsic matrix be defined as:

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{D} = [k_1, k_2, p_1, p_2, k_3]. \tag{3.3}$$

Here (f_x, f_y) are the focal lengths in pixels (c_x, c_y) is the optical center, and $\mathbf{D}$ contains the lens distortion parameters: k_1, k_2, k_3 (radial distortion) and p_1, p_2 (tangential distortion) [31]. Radial distortion causes straight lines to appear curved, particularly towards the edges of the image while tangential distortion occurs when the lens and image are not perfectly parallel [31]. Removing these distortions ensures geometric measurements in the image accurately reflect the true projection of 3D points, which is critical for precise marker localization. Inaccurate distortion correction would result in systemic pose estimation errors, especially for markers near the edges of the camera's field of view.

Each undistorted frame is processed using the OpenCV `aruco` module to detect 2D image-space marker corners $\{\mathbf{x}_i\}$ corresponding to known 3D object-space coordinates $\{\mathbf{X}_i\}$. Pose estimation is then performed via the Perspective-n-Point (PnP) algorithm [32], solving for the rotation vector as shown in Figure 3.3 $\mathbf{r} \in \mathbb{R}^3$ and translation vector $\mathbf{t} \in \mathbb{R}^3$ such that:

$$\mathbf{x}_i = \pi(\mathbf{K} \cdot [\mathbf{R}|\mathbf{t}] \cdot \mathbf{X}_i), \tag{3.4}$$

where π is the camera projection operation, and $\mathbf{R} \in SO(3)$ is the rotation matrix obtained from $\mathbf{r}$ using Rodrigues' formula [33]. This yields the marker pose relative to the camera

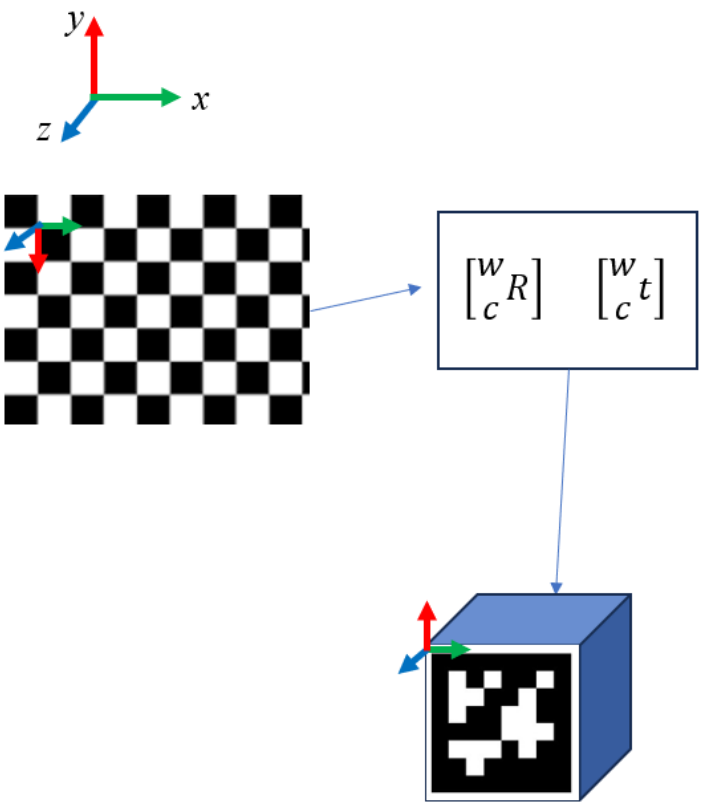

Figure 3.3: Calibration and ArUco marker–based pose estimation pipeline [1] used on the HoloLens2. Point correspondences between 3D board coordinates and 2D camera image points are computed using OpenCV's camera intrinsics and the perspective-n-point (PnP) algorithm by taking various frames of a checkerboard image. The resulting rotation (R) and translation (t) vectors align the virtual coordinate system of the AR overlay with the real-world marker position.

frame, represented as a homogeneous transform:

$$\mathbf{T}_{\text{marker\_cam}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4}. \tag{3.5}$$

To convert this pose into Unity world coordinates, we apply a series of transformations. The camera-to-world matrix $\mathbf{T}_{\text{cam\_world}}$, provided by the Unity XR API,

is multiplied with the marker pose in the camera frame:

$$\mathbf{T}_{\text{marker\_world}} = \mathbf{T}_{\text{cam\_world}} \cdot \mathbf{T}_{\text{marker\_cam}}. \tag{3.6}$$

As Unity uses a left-handed coordinate system while OpenCV uses a right-handed one, additional axis adjustments are applied. This involved negating the Y and Z axis to ensure proper orientation [34].

Once the marker's world pose is established, the ultrasound image is placed at a spatial offset from this marker to match the true location of the imaging plane. This offset is defined as a fixed transform $\mathbf{T}_{\text{offset}} \in \mathbb{R}^{4\times4}$, determined by visual alignment during an initial calibration procedure of measuring from the center of the ArUco marker to the end of the probe:

$$\mathbf{T}_{\text{ultrasound\_world}} = \mathbf{T}_{\text{marker\_world}} \cdot \mathbf{T}_{\text{offset}}. \tag{3.7}$$

The rotation and position extracted from $\mathbf{T}_{\text{ultrasound\_world}}$ are then used to update the transform of the World Space canvas in Unity each frame. This enables the ultrasound to appear fixed in space relative to the marker, even as the user or the probe moves.

3.3.2 Spatial Calibration

To accurately align the rendered ultrasound image with the physical world, two calibration steps were required: intrinsic calibration of the HoloLens PV camera and spatial calibration of the probe-to-marker offset. First, the FrameSaver tool [26] was used to collect RGB images and pose data from the HoloLens PV camera while imaging a printed checkerboard target. This data was processed offline using OpenCV's calibration routines to compute the camera's intrinsic parameters (focal length, principal point, and distortion coefficients), enabling accurate ArUco marker detection and pose estimation in Unity.

Second, spatial calibration was performed to align the ultrasound imaging plane with the ArUco marker coordinate system. To enable accurate visualization of the ultrasound image in AR, a spatial transform between the tracked ArUco markers and the imaging plane of the ultrasound probe was required. This calibration step ensured that the virtual ultrasound content appeared anchored at the correct physical location, providing accurate spatial correspondence during AR-guided imaging. In this system, four ArUco markers were affixed to the probe in a rigid geometric configuration—one on each visible face of the probe (front, back, and lateral sides). Each marker was a 50 mm $\times$ 50 mm square, printed on matte paper and attached such that its center had a known and fixed offset relative to the imaging base of the probe.

Each marker defines its own coordinate system, and the Unity application was configured to recognize any visible marker and use its pose to infer the probe's position. The transform from the detected marker pose $\mathbf{T}_{\text{marker}}$ to the ultrasound imaging plane was defined by a pre-measured transform $\mathbf{T}_{\text{offset}}$, such that the final imaging plane transform is given by:

$$\mathbf{T}_{\text{image}} = \mathbf{T}_{\text{marker}} \cdot \mathbf{T}_{\text{offset}}. \tag{3.8}$$

The transform $\mathbf{T}_{\text{offset}}$ was initially determined by manually measuring the physical distance between each marker's center and the bottom center of the probe's imaging face using a ruler. The front-facing marker was mounted 185 mm above the imaging face and had an offset of 25 mm laterally from the face of the marker to the probe's imaging plane, yielding a translation vector:

$$\mathbf{t}_{\text{offset}}^{(\text{front})} = \begin{bmatrix} 0 \\ -18.5 \\ -2.5 \end{bmatrix} \text{ cm}.$$

Similar offset vectors were computed for the remaining side and rear markers. Because the markers were rigidly attached to the probe body, these offsets remained constant across sessions and were hardcoded as fixed transformation matrices in Unity.

The system dynamically selected the most visible marker at runtime. When the user moved or rotated the probe, the tracking system detected whichever ArUco marker was in view and updated the pose accordingly. Because each marker had a known offset to the imaging origin, the ultrasound overlay remained locked to the physical probe's imaging plane across different orientations and viewpoints. This multi-marker layout improved robustness and ensured smooth transitions during probe manipulation.

3.4 System Evaluation and Testing

To evaluate the performance and clinical viability of the developed AR-guided multimodal ultrasound system, two primary tests were conducted: (1) system latency testing and (2) spatial registration accuracy testing. These evaluations were designed to quantify both the responsiveness and spatial alignment accuracy of the system, which are critical factors in ultrasound-guided interventional procedures.

3.4.1 Latency Testing

System latency was measured to assess the end-to-end communication and processing delays between the hardware and software components, including the Clarius probe, Qt application, MATLAB processing pipeline, and Unity AR visualization on the HoloLens 2. Custom logging scripts were developed to capture time-stamped events at key stages of the image acquisition, transmission, processing, and display pipeline.
For each trial, the following timestamps were recorded:

- Frame sent from Qt to Unity (B-mode and elastography frames)
- Frame received by Unity from Qt (B-mode and elastography image display)
- Unity UI trigger to send a raw data download request to Qt

- Raw data request received by Qt
- Raw data acquisition completed by Qt
- Command sent from Qt to MATLAB to start QUS processing
- Command received by MATLAB
- QUS image sent from MATLAB to Unity
- QUS image displayed in Unity

Ten trials were completed, each consisting of approximately 5 seconds of B-mode imaging, 5 seconds of elastography imaging, followed by manual selection of the "Run QUS" option within the HoloLens user interface. Latency measurements were computed for five key intervals (1) B-mode latency (i.e., time between frame sent from Qt to B-mode image display), (2) Elastography latency (i.e., time between elastography frame sent from Qt to elastography image display), (3) Raw data download (i.e., time between raw data request to start of raw data processing), (4) QUS processing (i.e., time between command sent from Qt to begin processing, and QUS image transmission) and (5) QUS display (i.e., time from QUS image sent to image display). For each interval, the mean and standard deviation latency values were calculated across all trials.

3.4.2 Registration Accuracy Testing

3.4.2.1 Experimental Setup

To evaluate the spatial registration accuracy of the AR-ultrasound system, phantom-based experiments were conducted using three different phantoms: (1) a custom-made water phantom for controlled point-based accuracy measurements, (2) the CIRS Model 054GS General Purpose Ultrasound Phantom, and (3) the CIRS Model 049 Elasticity Phantom (CIRS Inc., Norfolk, VA). Each phantom provided complementary testing conditions to assess system performance across a range of imaging depths, tissue properties, and modalities. Figure 3.4 shows the validation setup including the three different phantom types.

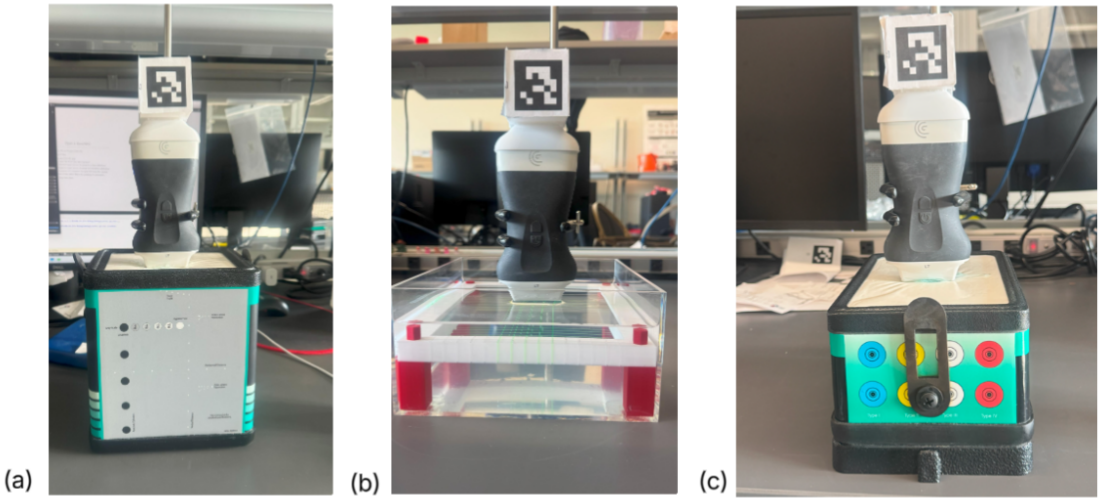

Figure 3.4: Phantom validation setup using the (a) general purpose phantom, (b) point target phantom and (c) elastography phantom.

The custom water phantom was constructed to contain embedded point targets at precisely known depths. It was used to quantify depth-dependent registration accuracy under idealized conditions. Specifically, 3D printed posts were used to suspend a 3D printed frame surrounding the edge of a plastic container, as shown in Figure 3.4. The frame contains several notches 1 cm apart, and the thickness of the frame is 1.5 cm. Within the 5 central notches in both x- and y- directions, green nylon thread was secured in a loop, creating point targets for ultrasound.

The ultrasound probe was tested under two imaging depths: 2 cm and 4 cm. At the 2 cm imaging depth, three well-isolated point targets were visible at 1.5 cm; at the 4 cm imaging depth, three additional deeper targets became visible at 3 cm. These targets served as ground truth references for spatial alignment. For each imaging depth, ten independent trials were conducted with the probe positioned at the same location and orientation to isolate overlay consistency. In the 2 cm imaging depth, only the top 3 points

were selected. In the 4 cm imaging depth, only the bottom 3 points were selected, representing deeper imaging targets. Only B-mode imaging was used for this phantom.

The CIRS 054GS phantom was used to simulate realistic clinical B-mode imaging with static anatomical targets. The probe was positioned in ten predefined locations to capture distinct circular structures distributed throughout the phantom. These structures had well-defined, high-contrast boundaries, enabling consistent selection of fiducial landmarks for error analysis.

The CIRS 049 Elasticity Phantom was used to assess registration accuracy using the elastography imaging mode. This phantom contains inclusions of known size, shape, and stiffness, allowing evaluation of overlay performance during elastography imaging. In each of ten trials, the probe was placed to image one or more inclusions, while the probe was used to manually compress the tissue and introduce deformation required for strain elastography. Videos were captured in both B-mode and elastography modalities to examine consistency across imaging modes and under tissue-like deformation.

For all phantom tests, a 3–5 second video was recorded from the HoloLens 2 perspective, capturing the AR overlay from multiple viewing angles to simulate realistic use. The phantoms were positioned at approximately eye level to minimize perspective distortion. Still frames were extracted from each video, and landmark points were manually annotated using a custom MATLAB script.

3.4.2.2 Annotation and Error Analysis To compute registration error between the rendered ultrasound overlay and the actual target position in the phantom, a custom MATLAB tool was developed for post-hoc annotation and error analysis. The first step involved a manual calibration procedure to convert pixel distances to physical units (mm) as shown in Figure 3.5. Specifically, the ultrasound array was present in each image and represents a known physical length of 38mm. Selecting the two ends of the array and measuring the distance in pixels, enables the definition of a pixel-to-mm conversion factor

for that particular example. This calibration factor was used for all subsequent distance measurements within the same image.

Two types of error measurements were implemented depending on the phantom used. For the general-purpose and elastography phantoms, which contain circular inclusions with known geometry, a circle-to-circle method was used. Demonstrated in Figure 3.6, the center and edge of the visible inclusion was selected for each the known phantom target and the overlaid B-mode or elastography image. The Euclidean distance between the center of these circles was calculated in both pixel and millimeter space. This approach allowed for repeatable measurement of spatial registration error across fixed anatomical structures. Each trial output included the center coordinates, radii, and inter-center distance in mm, which was stored in a CSV log for further statistical analysis.

For the custom water phantom with embedded point targets, a point-based error method was applied. Demonstrated in Figure 3.7, the ground truth and overlay points from each image frame were each selected. Between one and eight pairs of landmarks were selected per trial, and per-point Euclidean errors were computed. The root mean square error (RMSE) was calculated across all selected points and scaled into millimeters using the previously determined calibration factor. These values were logged across trials.

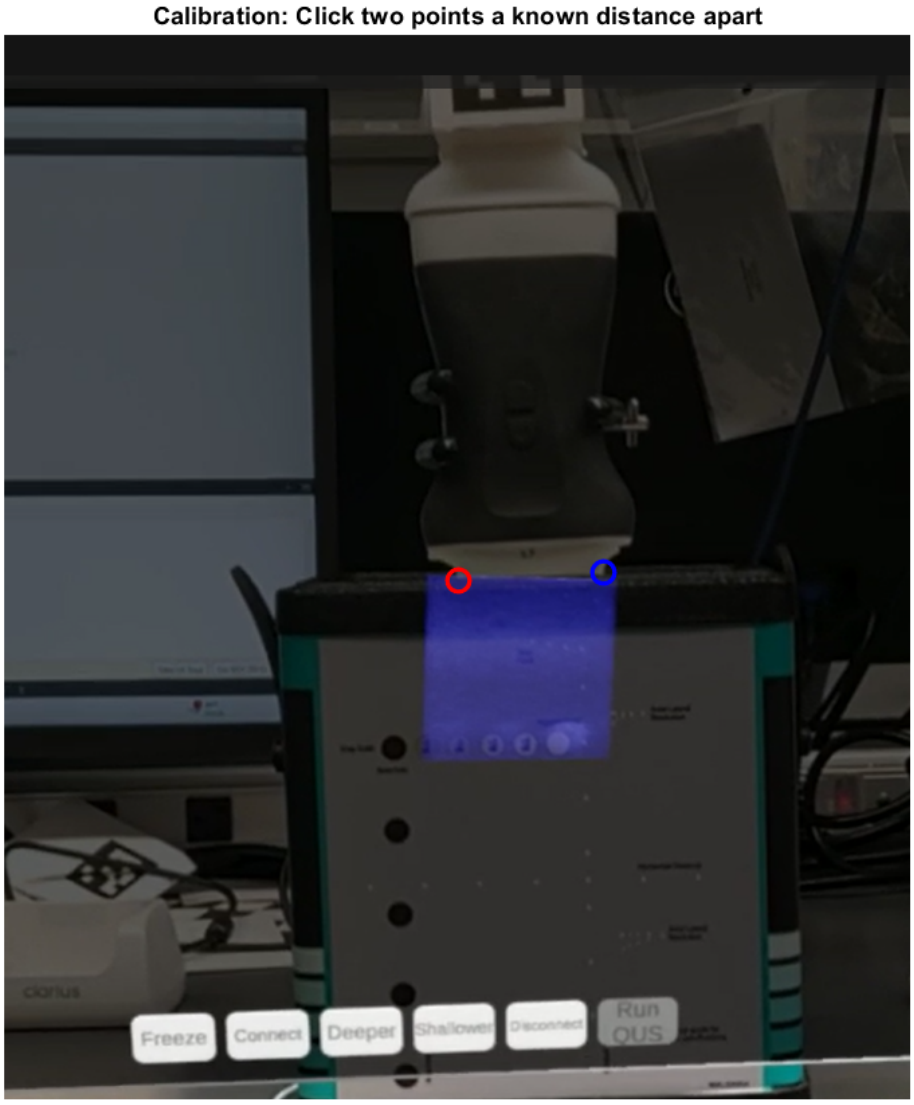

Figure 3.5: Pixel-to-millimeter calibration process. Two points with a known real-world distance were selected from the ultrasound probe base (38 mm) to compute a pixel-to-mm conversion factor. This scale was used for all subsequent registration error calculations.

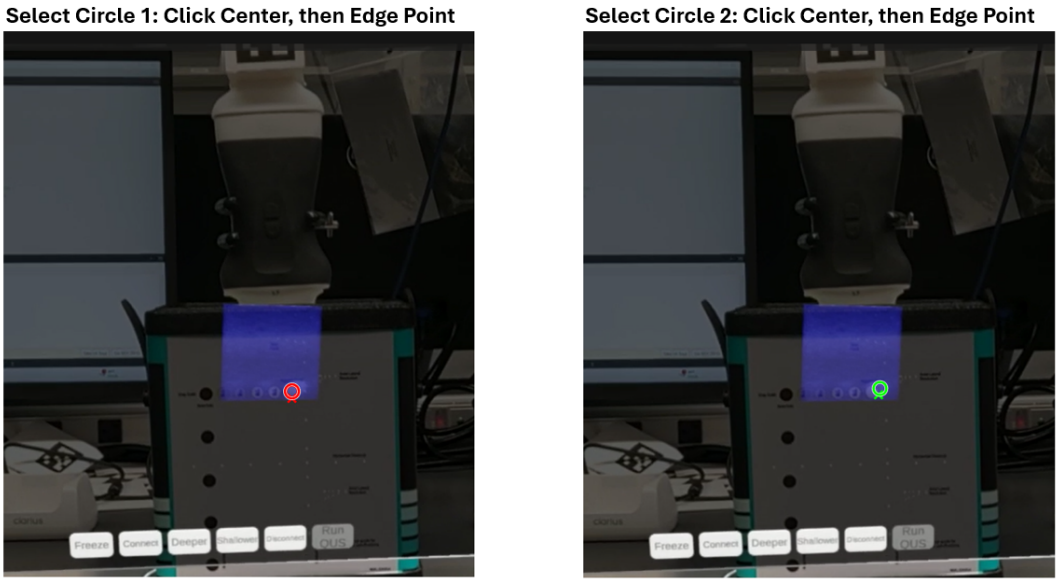

Figure 3.6: Circle-based registration error annotation. The center and edge of two inclusions were identified in both the real world (left) and AR-estimated (right) ultrasound images. The distance between circle centers was used to calculate inter-target registration error.

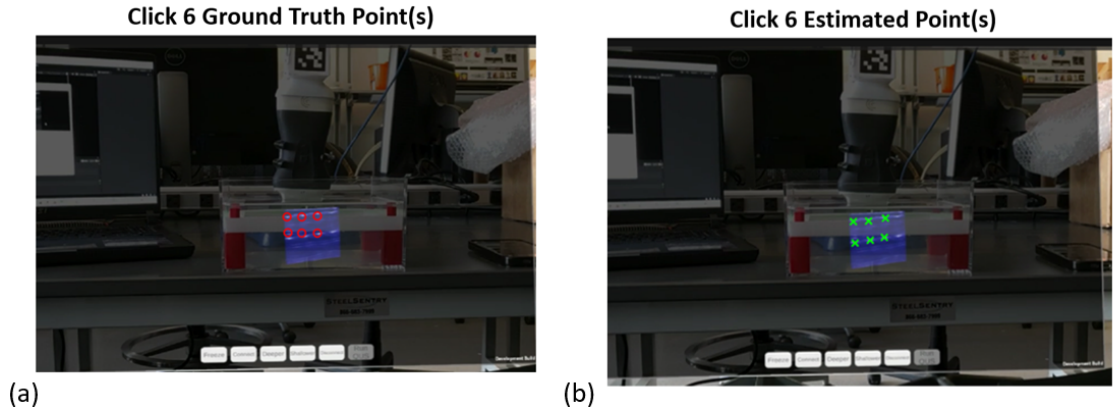

Figure 3.7: Point-based registration accuracy measurement. Anatomical landmarks were manually identified in the (a) real world space and (b) the AR-rendered ultrasound image. Euclidean distances between point pairs were used to compute per-trial RMSE.

CHAPTER FOUR

RESULTS

4.1 Demonstration of System

Figure 4.1 shows the multimodal functionality of the developed AR ultrasound guidance system running on the HoloLens 2. The figure demonstrates three distinct imaging modes: (a) B-mode imaging, (b) elastography imaging, and (c) QUS imaging. The elasticity phantom was used to generate visible tissue displacement by applying manual compression, providing a clear elastography image suitable for AR visualization. This multimodal capability highlights the system's flexibility in adapting to different clinical imaging needs, with each mode offering complementary information about tissue structure and mechanical properties. The ultrasound image is spatially tracked and scaled correctly within the AR environment based on the position of the tracked probe, while its placement in depth is user-defined to match the desired viewing perspective.

4.2 Latency Testing

Table 4.1 summarizes the system processing time and latency measurements collected across multiple trials for each stage of the pipeline, including (1) B-mode image display latency, (2) raw data request and download latency, (3) QUS processing time, and (4) QUS display latency. Each value represents the mean ± standard deviation (SD) measured across ten trials. The overall mean B-mode latency was 159.3 ms, representing real-time performance, with minor delays in elastography at 704 ms. For QUS imaging, these measurements capture the full end-to-end delay from image capture on the ultrasound probe to image display on the HoloLens 2. Notably, the raw data request and download latency shows high variability (158816.0 ± 133601.6 ms), suggesting potential network or processing bottlenecks that could be addressed in future optimizations. In contrast, the relatively low QUS display latency (208.0 ± 396.0 ms) and QUS processing

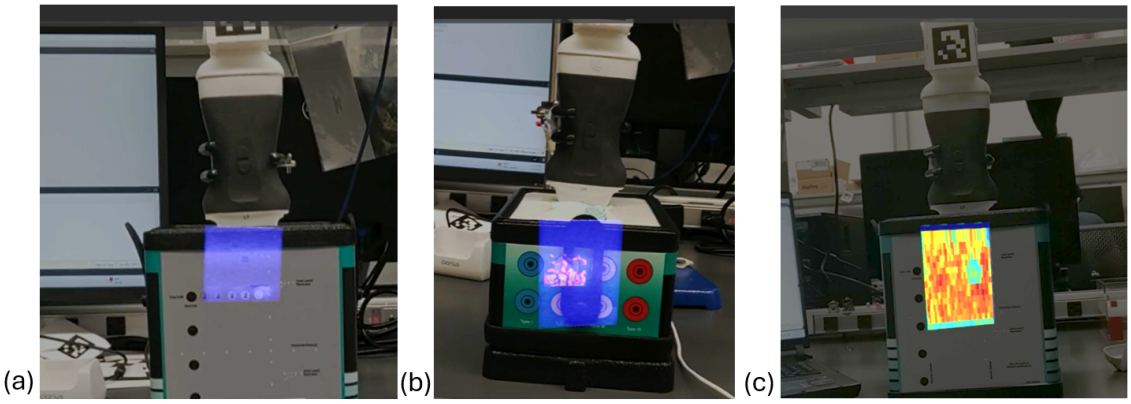

Figure 4.1: HoloLens screenshot demonstrating the multimodal functionality of our AR system showing (a) B-mode imaging, (b) elastography imaging, and (c) QUS imaging.

Table 4.1: Summary of system processing time.

Stage	**Time (ms)**
B-mode latency	159.3 ± 161.5
Elasography latency	704.1 ± 6442.6
Raw data download	158816.0 ± 133601.6
QUS processing	7968.2 ± 3980.2
QUS display	208.0 ± 396.0
Total QUS latency	166992.2 ± 131956.5

time (7968.2 ± 3980.2 ms) indicate that the primary time constraint lies in data acquisition and transmission rather than processing or rendering.

Figure 4.2 shows boxplots representing the latency measurements per trial for each stage of the imaging pipeline. The red line indicates the median, the blue box indicates the interquartile range, and the red crosses indicate outliers. Notably, the raw data download stage exhibits the widest range and largest variability, as reflected by the larger interquartile range. In contrast, the QUS processing and display latencies show

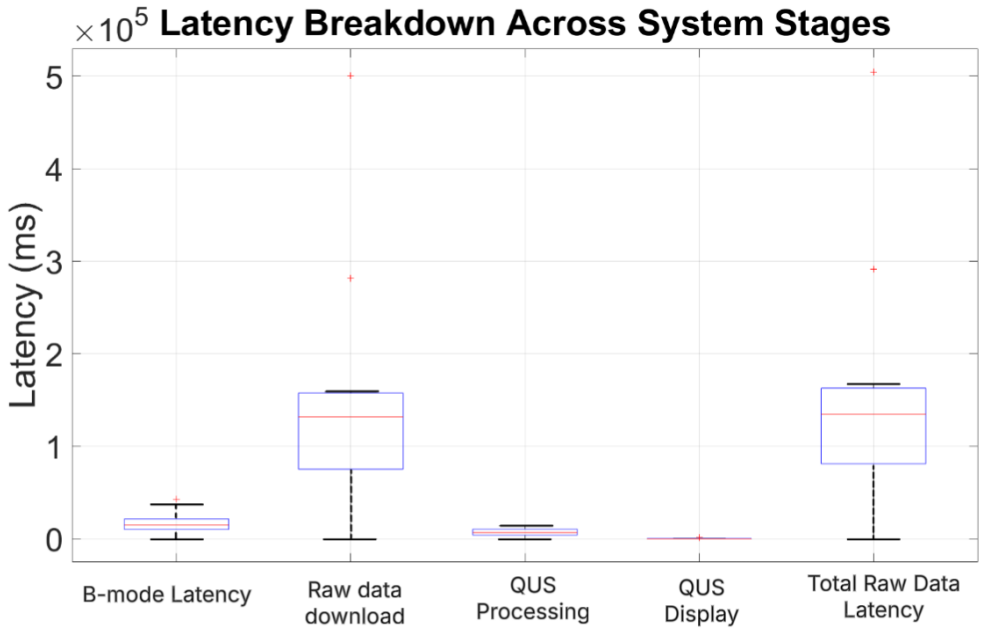

Figure 4.2: Latency measurements per trial for the B-mode image display, the raw data request and download, QUS processing, and QUS display.

tighter distributions and lower overall latency values. This figure highlights the variability in data acquisition latency and the relatively stable performance of QUS processing and visualization.

In addition to the stage-wise breakdown, Figure 4.3(a) and (b) present both the overall system latency and effective frame rate for B-mode, elastography, and QUS imaging modes, respectively. The latency plot highlights the significant increase in total processing time for QUS (1.67e+5 ms) compared to B-mode (159.3 ms) and elastography (704.1 ms), primarily due to the processing bottleneck in raw data request with the external Qt and MATLAB processing pipeline for QUS image reconstruction. This marked difference in latency underscores the trade-off between advanced imaging capabilities and real-time performance. The corresponding frame rate plot further illustrates this trade-off, showing that B-mode achieved the highest frame rate of 6.28 fps,

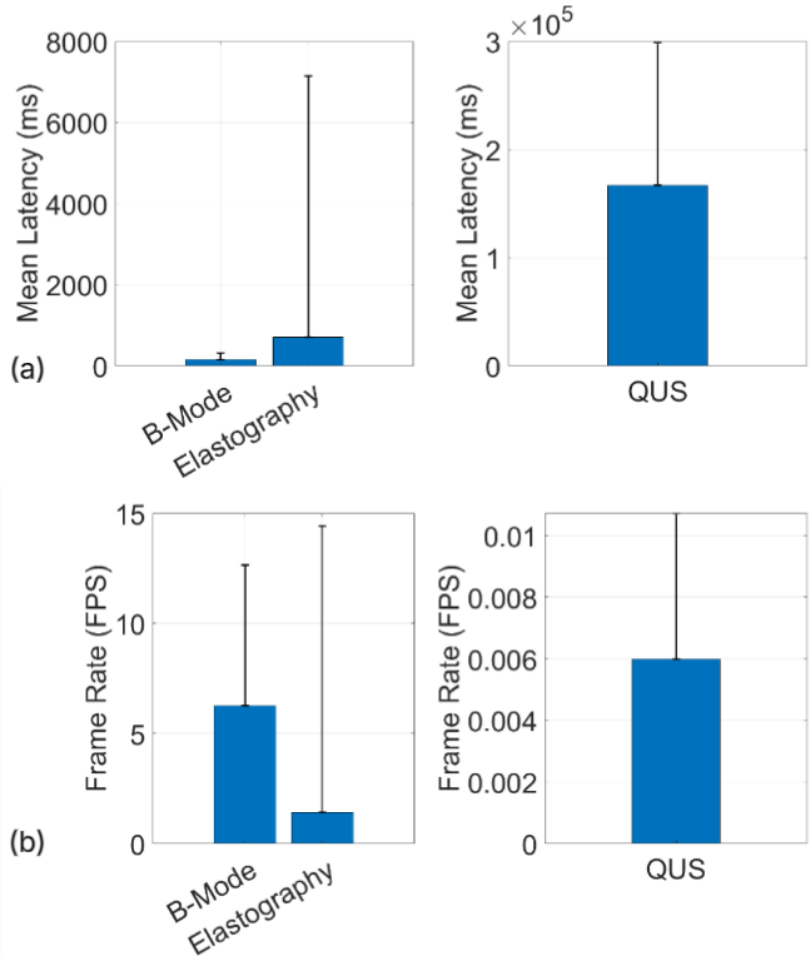

Figure 4.3: Overall system (a) latency measurements and (b) frame rate for B-mode imaging, elastography imaging, and QUS imaging.

followed by elastography at 1.4203 fps, with QUS operating 0.006 fps. These low frame rates reflect the current limitations in the system and highlight areas for future hardware or algorithmic improvements.

4.3 Registration Accuracy Testing

Figure 4.4 shows the registration error measured using the water-based point target phantom at two imaging depths 2 cm and 4 cm. At the 2 cm depth, the mean registration error was 2.77 mm $\pm$ 0.31 mm (mean $\pm$ SD) across 10 trials. For the 4 cm depth, the error

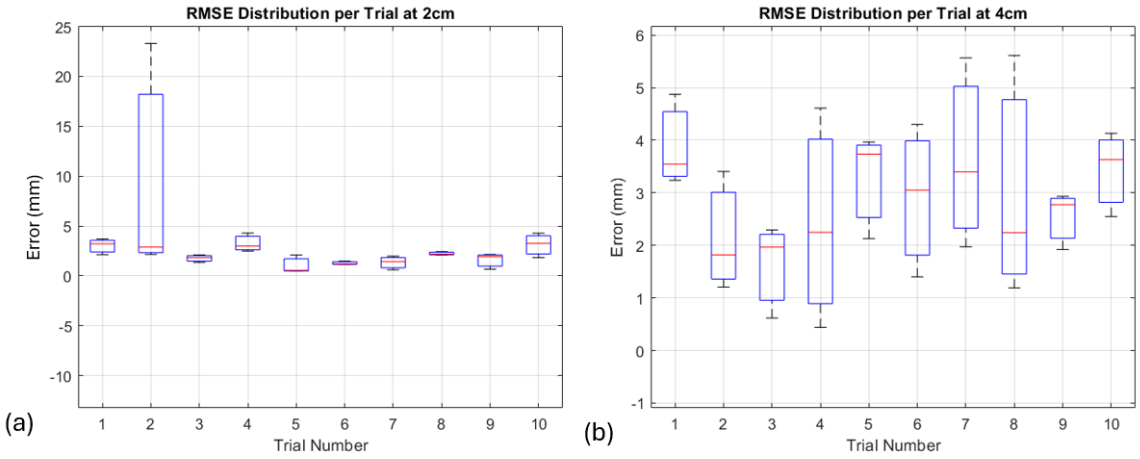

Figure 4.4: Registration error using the point target phantom at (a) 2 cm depth and (b) 4 cm depth.

increased slightly to 2.82 mm $\pm$ 0.28 mm, reflecting a minor degradation in spatial alignment accuracy with increased imaging depth. This trend is consistent with the expected decrease in spatial resolution away from the focus in ultrasound images. However, the relatively small difference between depths indicates robust registration performance across clinically relevant imaging ranges.

Figure 4.5 shows boxplots for each trial representing the center-to-center error of targets from the general purpose phantom. The mean registration error across all 10 trials was 10.80 mm $\pm$ 4.32 mm (mean $\pm$ SD).This variation likely stems from several factors, including minor inconsistencies in marker placement, fluctuations in the camera's pose estimation accuracy, and slight variations in ultrasound probe positioning or angulation during imaging. Additionally, the phantom's surface properties may influence the visibility and detectability of the ArUco markers, which can further impact spatial alignment.

Figure 4.6 presents the registration error measured from the elastography phantom across 10 trials, with a RMSE of 16.99 mm $\pm$ 4.55 mm (mean $\pm$ SD) over all trials. The

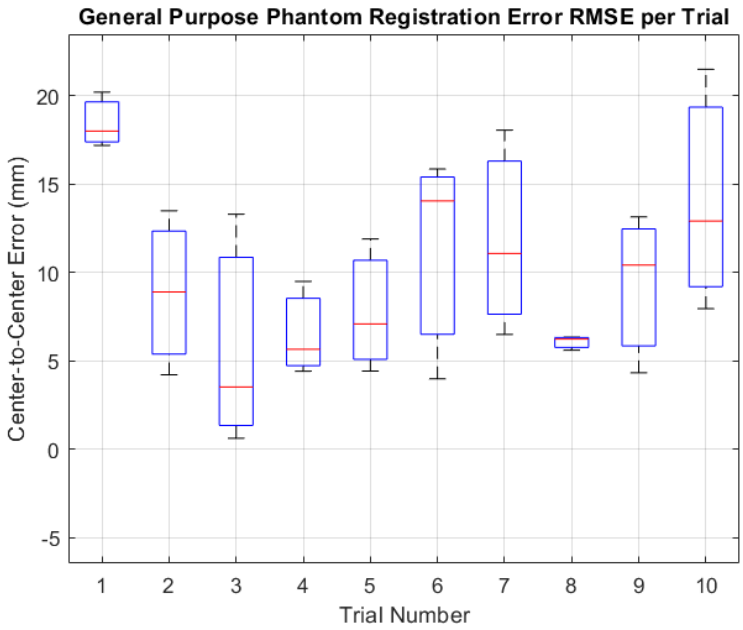

Figure 4.5: Registration error using the general purpose phantom.

error was higher than that observed with the general-purpose phantom, likely due to minor phantom deformation and increased complexity in feature boundaries when using elastography. These results highlight the challenges of maintaining precise spatial registration when imaging elastic structures, reinforcing the importance of accounting for tissue compliance in guided interventions.

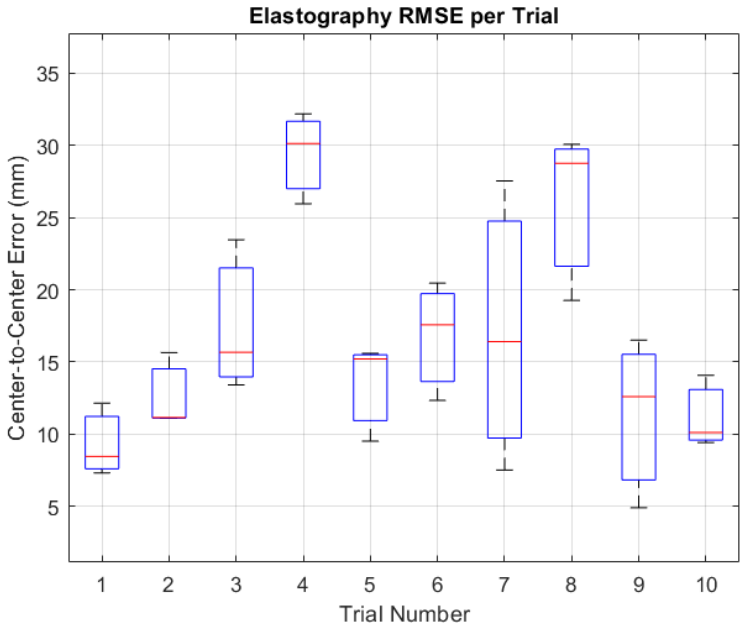

Figure 4.6: Registration error using the elastography phantom.

CHAPTER FIVE

DISCUSSION

This study presents a novel AR-based ultrasound guidance system that integrates multimodal imaging, wireless probe technology, and real-time visualization within an AR environment. Compared to prior AR-ultrasound platforms, which typically rely on B-mode imaging alone and are constrained by wired setups and limited interactivity [7, 13, 19], the developed system offers several technical and clinical advantages.

5.1 Multimodal Imaging in AR Context

One key strength of the presented system is its successful integration of multimodal ultrasound data streams, including conventional B-mode imaging, elastography, and QUS modalities. This multimodal approach allows for improved tissue characterization and provides clinicians with more comprehensive diagnostic information during procedures. By integrating multiple ultrasound parameters within the AR interface, the system overcomes one of the main limitations of previous AR-ultrasound solutions, which often present only conventional B-mode ultrasound images without additional diagnostic context [8]. This richer visualization has the potential to improve targeting accuracy and procedural decision-making during ultrasound-guided biopsies.

5.2 Implications for Clinical Workflow and Training

Another significant advantage of this system is its wireless design. The use of a Wi-Fi-based Clarius ultrasound probe eliminates the need for tethered hardware setups, allowing for greater mobility during procedures. This wireless configuration also enhances system adaptability across different clinical environments. In contrast, many existing AR-ultrasound systems depend on wired probes and external tracking hardware, limiting their portability and ease of deployment [3].

Moreover, the AR interface has meaningful implications for clinical education and remote guidance. By overlaying ultrasound data on the patient anatomy with spatial consistency, the system can assist novice users in understanding probe positioning and interpreting imaging features in context. In a training environment, expert clinicians could observe and annotate within the same AR space, providing just-in-time guidance [5]. Similarly, this spatial alignment could support telemedicine scenarios, where remote specialists assist bedside practitioners in performing Ultrasound (US)-guided procedures.

5.3 Latency and System Responsiveness

Preliminary testing showed a latency of approximately 159 ms for B-mode imaging and 704 ms for elastography. For reference, [16] and [18] reported a latency of 143.5 ± 51.5 ms and 122.49 ± 11.6 ms, respectively in their AR-guided ultrasound systems, suggesting that the B-mode latency in the system is comparable and may be sufficient for clinical tasks. However, the larger latency observed in elastography imaging exceeds typical thresholds for real-time interaction and may hinder its use in time-sensitive procedures. This increased delay is likely due to the added computational burden required for elastography on the Clarius probe paired with the required data transfer. Potential improvements include optimizing data transfer protocols, compressing raw data more efficiently, and integrating processing directly on the host device to reduce round-trip latency.

The largest latency was observed for QUS imaging, primarily due to the added time required for raw data acquisition over the wireless network and the external MATLAB-based image processing pipeline, which introduces significant computational and data transfer delays. The notably high latency and low frame rate for QUS currently limits its suitability for direct real-time clinical use, especially in scenarios demanding immediate feedback such as intraoperative guidance. However, the latency bottleneck is largely attributable to network transfer speeds and the external processing framework.

Optimizations such as implementing more efficient data compression techniques, utilizing on-device processes for image reconstruction, and streamlining communication protocols could substantially reduce this delay. With these improvements, the system could achieve clinically acceptable latency thresholds, enabling the practical deployment of QUS for enhanced diagnostic and interventional applications.

5.4 Registration Accuracy and Phantom-Based Insights

Registration accuracy demonstrated varying performance across phantoms with different properties and complexities. Using the point target phantom, mean registration errors remained low and consistent across imaging depths (2.77 ± 0.31 mm at 2 cm and 2.82 ± 0.28 mm at 4 cm; Figure 4.4), reflecting robust spatial alignment in controlled conditions. These values align well with prior AR-ultrasound registration studies [16] and are within acceptable margins for needle guidance within targeted tissue regions.

In contrast, the general-purpose ultrasound phantom and elastography phantom exhibited substantially higher registration errors, with mean RMSEs of 10.80 ± 4.32 mm and 16.99 ± 4.55 mm, respectively (Figures 4.5 and 4.6). The increased error in the elastography phantom likely arises from tissue-mimicking inclusions that undergo deformation and have more complex boundaries, which challenge both the tracking system and registration algorithms. The increased error observed when imaging either the general purpose phantom or the elastography phantom as opposed to the water phantom is likely due to inaccuracies in estimating the true center of the circles, especially under image distortion from the HoloLens. Unlike point targets, which offer a precise reference, the circular targets tested require center and edge estimation during error analysis, which can introduce an additional level of uncertainty.

CHAPTER SIX

CONCLUSION AND FUTURE WORK

This thesis presents the first AR-based ultrasound-guided biopsy system that incorporates multimodal information from elastography and QUS to include critical diagnostic information into the AR visualization. We demonstrate technical feasibility in aligning multimodal ultrasound data with AR-tracked spatial markers and streaming images wirelessly to a head-mounted display. Preliminary phantom-based testing indicated consistent spatial alignment under controlled conditions and showed that the system is capable of rendering ultrasound images within the AR environment with minimal perceived lag. The system establishes a foundation for future AR-guided ultrasound platforms, with potential extensions toward clinical training tools and remote procedural support as spatial accuracy, latency, and usability are further improved.

While challenges remain, including latency, tracking robustness, and the potential for expanded modality support, this work represents a significant step toward AR-assisted ultrasound guidance systems with increased tissue data acquisition capability. With further technical refinement and clinical validation, the proposed system has the potential to improve procedural accuracy, support novice clinician training, and expand access to image-guided interventions in both urban and rural healthcare settings.

There are five key areas where this system could be expanded to enhance its clinical utility, technical performance, and research impact. First, reducing system latency remains a priority for enabling real-time interactivity, particularly in quantitative imaging modes like QUS. Future work should focus on optimizing the processing pipeline through parallel computing, GPU acceleration, and improved data transfer protocols. Eliminating dependence on external MATLAB-based processing and integrating computation directly

into the Unity application—or into the HoloLens runtime once Clarius SDK support for ARM64 becomes available—would significantly improve both latency and portability.

Second, while the current system relies on ArUco marker-based tracking, this approach is vulnerable to occlusion and environmental factors such as lighting or angle of view. To improve robustness, future versions should explore hybrid or markerless tracking strategies, including deep learning-based object tracking for probe localization [35]. In addition, integration with the Mixed Reality Toolkit (MRTK) will be pursued to improve spatial tracking accuracy and enable more immersive interaction. MRTK's advanced spatial mapping, hand tracking, and input systems could enhance registration fidelity by refining the alignment between virtual overlays and physical anatomy [36]. This integration also opens opportunities for more intuitive user interfaces that leverage gesture-based control, speech control and spatial anchors, potentially reducing user error and improving procedural flow.

Third, the system could benefit from the integration of artificial intelligence (AI) algorithms into the imaging workflow. Real-time tissue segmentation, lesion classification, and automated needle detection could reduce the cognitive load on clinicians and streamline decision-making [37]. These features would also support automation of key components of the image analysis pipeline, potentially improving consistency and reducing user dependence.

Fourth, the integration of photoacoustic imaging (PAI) would provide complementary functional and molecular information. PAI captures optical absorption contrast, enabling visualization of tissue composition, hemoglobin concentration, or malignancy-specific markers that are not visible with conventional ultrasound. This modality could significantly enhance lesion detection and targeting during biopsies [38]. However, incorporating PAI would require addressing several technical challenges, including synchronizing acquisition between ultrasound and optical sources, and

developing real-time AR rendering methods to integrate PAI overlays within the current visualization pipeline [39].

Finally, although the system has demonstrated strong registration accuracy under phantom testing, it has not yet undergone human-subject validation. Future work should involve clinical and user-centered studies to evaluate usability, learning curves, and the system's impact on procedural accuracy and workflow efficiency [5, 40]. These studies are essential to assess adoption readiness and establish the clinical value of the AR-guided ultrasound platform. Controlled evaluations involving both expert clinicians and novice users will provide valuable insights into real-world performance and help identify any remaining gaps in the system's translational potential. By further improving and developing the initial system, it has the potential to evolve from a research prototype into a clinically deployable AR guidance tool suitable for a wide range of ultrasound-guided interventions.

REFERENCES

[1] M. Doughty, "ArUco Marker Detection with HoloLens 2 and Research Mode." https://doughtmw.github.io/posts/ArUco-Detection-HoloLens-4.

[2] A. Dinh, A. L. Yin, D. Estrin, P. Greenwald, and A. Fortenko, "Augmented reality in real-time telemedicine and telementoring: Scoping review," *JMIR mHealth and uHealth*, vol. 11, p. e45464, Apr. 2023.

[3] A. Lastrucci, Y. Wandael, A. Barra, R. Ricci, G. Maccioni, A. Pirrera, and D. Giansanti, "Exploring augmented reality integration in diagnostic imaging: Myth or reality?," *Diagnostics*, vol. 14, no. 13, p. 1333, 2024.

[4] A. State, M. A. Livingston, W. F. Garrett, G. Hirota, M. C. Whitton, E. D. Pisano, and H. Fuchs, "Technologies for augmented reality systems: Realizing ultrasound-guided needle biopsies," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp. 439–446, 1996.

[5] S.-C. Liao, S.-C. Shao, S.-Y. Gao, and E. C.-C. Lai, "Augmented reality visualization for ultrasound-guided interventions: a pilot randomized crossover trial to assess trainee performance and cognitive load," *BMC Medical Education*, vol. 24, no. 1, p. 1058, 2024.

[6] 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.

[7] C. Rüger, M. A. Feufel, S. Moosburner, C. Özbek, J. Pratschke, and I. M. Sauer, "Ultrasound in augmented reality: a mixed-methods evaluation of head-mounted displays in image-guided interventions," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, pp. 1895–1905, 2020.

[8] E. M. W. Ong, "Translating new breast ultrasound techniques into clinical practice: evaluating their intended uses and describing other unexpected uses for them," *Translational Breast Cancer Research*, vol. 4, p. 23, 2023.

[9] G. Cloutier, F. Destrempes, F. Yu, and A. Tang, "Quantitative ultrasound imaging of soft biological tissues: a primer for radiologists and medical physicists," *Insights into Imaging*, vol. 12, no. 1, p. 127, 2021.

[10] J. H. Catani, "Sonographic based imaging: Ultrasound, color doppler, elastography, and automated breast imaging," in *Modern Breast Cancer Imaging*, pp. 97–130, Springer, 2021.

[11] H. Kwon, S. Oh, M.-G. Kim, Y. Kim, G. Jung, H.-J. Lee, S.-Y. Kim, and H.-M. Bae, "Artificial intelligence-enhanced quantitative ultrasound for breast cancer: Pilot

study on quantitative parameters and biopsy outcomes," *Diagnostics*, vol. 14, no. 4, 2024.

[12] R. M. Sigrist, J.-H. Liau, A. E. Kaffas, M. C. Chammas, and J. K. Willmann, "Ultrasound elastography: review of techniques and clinical applications," *Theranostics*, vol. 7, no. 5, pp. 1303–1329, 2017.

[13] M. Kanbara, T. Okuma, H. Takemura, and N. Yokoya, "A stereoscopic video see-through augmented reality system based on real-time vision-based registration," in *Proceedings IEEE Virtual Reality 2000 (Cat. No. 00CB37048)*, pp. 255–262, IEEE, 2000.

[14] T. Okuma, K. Kiyokawa, H. Takemura, and N. Yokoya, "An augmented reality system using a real-time vision based registration," in *Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170)*, vol. 2, pp. 1226–1229 vol.2, 1998.

[15] 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.

[16] N. Costa, L. Ferreira, A. R. de Araújo, B. Oliveira, H. R. Torres, P. Morais, V. Alves, and J. L. Vilaça, "Augmented reality-assisted ultrasound breast biopsy," *Sensors*, vol. 23, no. 4, p. 1838, 2023.

[17] Y. Sato, M. Nakamoto, Y. Tamaki, T. Sasama, I. Sakita, Y. Nakajima, M. Monden, and S. Tamura, "Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization," *IEEE Transactions on Medical Imaging*, vol. 17, no. 5, pp. 681–693, 1998.

[18] H. Li, W. Yan, J. Zhao, Y. Ji, L. Qian, H. Ding, Z. Zhao, and G. Wang, "Navigate biopsy with ultrasound under augmented reality device: Towards higher system performance," *Computers in Biology and Medicine*, vol. 174, p. 108453, 2024.

[19] N. Cattari, S. Condino, F. Cutolo, M. Ferrari, and V. Ferrari, "In situ visualization for 3D ultrasound-guided interventions with augmented reality headset," *Bioengineering*, vol. 8, no. 10, p. 131, 2021.

[20] B. S. Garra, "Imaging and estimation of tissue elasticity by ultrasound," *Ultrasound Quarterly*, vol. 23, no. 4, pp. 255–268, 2007.

[21] M. L. Oelze and J. Mamou, "Review of quantitative ultrasound: Envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 2, pp. 336–351, 2016.

[22] Microsoft Corporation, "HoloLens Research Mode - Mixed Reality." https://learn.microsoft.com/en-us/windows/mixed-reality/develop/advanced-concepts/research-mode, 2024.

[23] OpenCV.org, "Pose Estimation." https://docs.opencv.org/4.x/d7/d53/tutorial_py_pose.html.

[24] Perk Lab, Queen's University, "PLUS Toolkit Model Catalog." http://perk-software.cs.queensu.ca/plus/doc/nightly/modelcatalog/.

[25] OpenCV, "Detection of ArUco markers." https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html.

[26] M. Doughty, "HoloLens Camera Calibration Tools." https://github.com/doughtmw/HoloLensCamCalib/tree/master.

[27] P. de Witte, "UnityMainThreadDispatcher." https://github.com/PimDeWitte/UnityMainThreadDispatcher/tree/master/Runtime.

[28] P. M. Shankar, "A general statistical model for ultrasonic backscattering from tissues," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 47, no. 3, pp. 727–736, 2000.

[29] P.-H. Tsui and C.-C. Chang, "Imaging local scatterer concentrations by the nakagami statistical model," *Ultrasound in Medicine & Biology*, vol. 33, no. 4, pp. 608–619, 2007.

[30] S. Muhtadi, R. R. Razzaque, A. Chowdhury, B. S. Garra, and S. Kaisar Alam, "Texture quantified from ultrasound Nakagami parametric images is diagnostically relevant for breast tumor characterization," *Journal of Medical Imaging*, vol. 10, no. S2, pp. S22410–S22410, 2023.

[31] G. Madejski, S. Zbytniewski, M. Kurowski, D. Gradolewski, W. Kaoka, and W. J. Kulesza, "Lens distortion measurement and correction for stereovision multi-camera system," *Engineering Proceedings*, vol. 82, no. 1, p. 85, 2024.

[32] OpenCV.org, "Pose Estimation - solvePnP Function." https://docs.opencv.org/4.x/d5/d1f/calib3d_solvePnP.html.

[33] OpenCV.org, "Camera Calibration and 3D Reconstruction: Rodrigues Function." https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html#void%20Rodrigues(InputArray%20src,%20OutputArray%20dst,%20OutputArray%20jacobian).

[34] U. Technologies, "Quaternion and euler rotations in unity." https://docs.unity3d.com/6000.1/Documentation/Manual/QuaternionAndEulerRotationsInUnity.html.

[35] N. Uke, P. Futane, N. Deshpande, and S. Uke, "A review on deep learning-based object tracking methods," *Multiagent and Grid Systems*, vol. 20, no. 1, pp. 27–39, 2024.

[36] L. Chen, H. Zhao, C. Shi, Y. Wu, X. Yu, W. Ren, Z. Zhang, and X. Shi, "Enhancing multi-modal perception and interaction: An augmented reality visualization system for complex decision making," *Systems*, vol. 12, p. 7, 12 2023.

[37] V. Kumar, J. M. Webb, A. Gregory, M. Denis, D. D. Meixner, M. Bayat, D. H. Whaley, M. Fatemi, and A. Alizad, "Automated and real-time segmentation of suspicious breast masses using convolutional neural network," *PLOS ONE*, vol. 13, no. 5, p. e0195816, 2018.

[38] N. Nyayapathi and J. Xia, "Photoacoustic imaging of breast cancer: a mini review of system design and image features," *Journal of Biomedical Optics*, vol. 24, no. 12, p. 121911, 2019.

[39] C. Yang, H. Lan, F. Gao, and F. Gao, "Review of deep learning for photoacoustic imaging," *Photoacoustics*, vol. 21, p. 100215, 2021.

[40] N. A. Farshad-Amacker, R. A. Kubik-Huch, C. Kolling, C. Leo, and J. Goldhahn, "Learning how to perform ultrasound-guided interventions with and without augmented reality visualization: a randomized study," *European Radiology*, vol. 33, no. 4, pp. 2927–2934, 2023.