Isosurface Visualization Using Augmented Reality for Improving Tumor Resection Outcomes

Niveditha Kalavakonda

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Electrical Engineering

University of Washington

2017

Reading Committee:

Blake Hannaford, Chair

Howard Chizeck

Program Authorized to Offer Degree:

Department of Electrical Engineering

University of Washington

Abstract

Isosurface Visualization Using Augmented Reality for Improving Tumor Resection Outcomes

Niveditha Kalavakonda

Chair of the Supervisory Committee:
Professor Blake Hannaford
Electrical Engineering

Tumor resection techniques have significantly improved survival rates in cancer patients. However, these require chemotherapy or radiation as follow-up procedures. With the advent of Mixed Reality technology, a wide range of possibilities have opened for the field of surgery. This work explores the use of mixed reality for maximizing tumor margin removed during surgical procedures. An isosurface reconstruction algorithm was integrated with the Microsoft HoloLens, a self-contained holographic computer, to enable visualization of Computed Tomography (CT) imaging superimposed in 3D on the patient. The results suggest that though the device has limitations at its current stage, the Microsoft HoloLens could be used for planning and overlaying the imaging information on the patient for excision of lesions in real-time. The modules developed could also be extended to other types of surgery involving visualization of Digital Imaging and Communication in Medicine (DICOM) files.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The author would like to thank Prof. Blake Hannaford for his continuous support in her research and study, for his patience, motivation and immense knowledge. Sincere appreciation is extended to Prof. Samuel Burden for his generosity in sharing research equipment without which this research would not have been possible. Finally, the author wishes to express gratitude to her family and friends who have been incredible mentors at different stages during this journey.

DEDICATION

To K. P. Vasudeva Reddy, my grandfather, who has been a source of constant encouragement, showing me there is nothing more valuable than knowledge, though he is no longer with us.

Chapter 1.

INTRODUCTION

1.1 MOTIVATION

Minimally Invasive Surgery (MIS) has revolutionized the way surgery is practiced. This has led to reduced visible scars on a patient, shorter stay at the hospital, improved recovery period after surgery and reduced trauma. MIS has conferred considerable advantages on the patient [1-2] but has also imposed additional difficulty for surgeons working in a limited workspace [3]. However, with the continual advancement of chip manufacturing, instrumentation, robotics and computer vision, there has been improved preoperative planning possibilities which enhance the skills of a surgeon and ameliorate complex procedures.

Image-guided surgery (IGS) has evolved with the advancing technology to augment a surgeon's ability to dexterously administer surgical procedures. It assists with decreasing the invasiveness of surgical procedures and improving safety and accuracy [4]. IGS tends to produce accurate results but it can be tedious for the surgeon to accurately perceive depth and avoid critical structures compared to an open surgery [5]. Since the view of 3D organs is projected on a 2D screen, there is information loss regarding the exact location of the tumor and critical structures. The visual displays of commercial navigation systems are generally situated away from the surgical field, adding to the problem. This presents surgeons with a difficult task of mismatched hand-eye coordination, thereby disrupting the surgical workflow [6]. Endoscopes offer a narrow keyhole field of view during the entire surgical procedure. The lack of depth perception limits the dexterity and maneuverability of a surgeon for tasks such as suturing and delicate dissection [7]. This intensifies the complexity of surgical procedures.

A number of patients are subjected to repeat surgeries due to the possibility of residual tumors. A significant percentage of the population (upto 71.6% for hepatocellular carcinoma and about 50% for intracranial meningioma) [7-8] are at risk after initial treatment for cancers affecting local regions. This increases cost further since locating the tumor and limiting the effects would be harder in the case of recurrences. An ideal solution is to minimize the risk of relapse and maximize conservation of healthy tissue.

Previous research suggests the potential benefits of using augmented or mixed reality devices to improve the precision of tumor margin excision [9-10]. However, the technologies developed were either too bulky, interfered with sterilization, lacked depth perception or had another significant limitation which restricted their use in the operating theater.

1.2 LITERATURE REVIEW

The Augmented Reality system presented in [12] performs extensive studies on the use of a live video camera feed to provide additional information about the patient's tumor for breast cancer surgery. The study analyzes the effects of tissue deformation that occurs between volume reconstruction from preoperative ultrasound scans and its use during surgery. The system required calibration between devices, similar to [13]. Reconstruction of the tumor is done by manual segmentation of images and is overlaid on the video feed. Though the tests in the phantom lab generated only a 2mm error, the clinical experiments proved that the system was not viable for real-time use. The patient was required to suspend breathing during the ultrasonic imaging to obtain scans with reduced error. This, however, resulted in errors during real-time use due to the inconsistent transformation matrices generated. These experiments demonstrated the potential usefulness of augmented reality for excision of the tumor.

Liao et. al [13] developed a 3D augmented reality navigation system by using an autostereoscopic imaging technique. This method reproduces three-dimensional images by making use of an array of convex lenses and a high-resolution flat display that has high pixel density. It makes use of various transformation matrices for the registration of the 3D display with the patient's anatomy by tracking and mapping of instruments. It additionally requires calibration of devices with respect to each other before every procedure and even multiple times during the same procedure. The device is placed at a 30-cm distance from the patient. Though the case studies suggest this would not affect the workflow, having an additional piece of equipment in close proximity could be an obstacle during emergencies.

The prototype built by Fuchs et. al [14] shows the benefits of using a Head Mounted Display device for providing guidance to surgeons in the case of minimally invasive laparoscopic surgery. The device aimed to provide the surgeon with a natural point of view, as presented during open surgery, by using a camera to create a vitual laparoscope by combining information from a conventional laparoscope and a projector emitting structured light (vertical stripes). The data captured provided

depth information for rendering true 3D structures, which can be aligned with the anatomy. UNC's optoelectronic ceiling tracker was used to estimate the user's head positions and orientations. Due to faulty depth extraction, depth calibration and the delays in image display, the system was not suitable for use during surgery.

Elmi-Terander et. al [15] developed an Augmented Reality Surgical Navigation (ARSN) system which integrates an intraoperative imaging system and optical tracking for thoracic pedicle screw placement using augmented reality. A study on cadavers using ARSN showed a 33% improvement in accuracy for matched complexity level. However, this information was still presented to the surgeon on a conventional display device.

In [16], Wen et. al, use a projection-camera system to track patient anatomy for registration of preoperative data for Radiofrequency ablation procedures. Once registered, the anatomy, however, could not be manipulated or interacted with directly. Change in positioning anatomy would require manipulating the software, which is tedious for a surgeon. But the results of static experiments verify the increase in accuracy of procedures while using augmented reality.

A team from Trinity College Dublin and Haptica [17] prototyped a box trainer integrated with a desktop-based virtual and augmented reality for training novice surgeons. Information is augmented on the simulation prototype for improving the training experience for novice surgeons. This is expensive since the phantoms in a box trainer cannot be reused multiple times. But the experimental evaluation shows the benefits of using augmented reality while training, even if only on a conventional system.

Another group evaluated the use of segmented CT images for registration with live camera feed from the operating room through experimental studies [18]. It evidenced the potential opportunities while using augmented reality for digestive surgery but disregarded the tracking error introduced into the system during needle insertion.

A mixed-reality headset, developed for visualization of anatomy during laparoscopic surgery, is described in [19]. This work is the closest-related prior work that uses hardware similar to the Microsoft HoloLens and has a native algorithm to reconstruct virtual surfaces for augmentation. However, their system had problems with respect to presenting reconstructed volumes because of computations taking as long as 30 minutes. This presents a long waiting period, which would add to the operating room cost. It also does not produce volumes which can be interacted with. This limit the amount of information that can be communicated during surgery.

1.3 PROJECT GOALS

1.3.1 *Long-term goal*

In the reviewed literature, only one study has mentioned the use of developing applications for reconstruction of imaging modalities to visualize isosurfaces. The general practice while using Mixed Reality appears to make use of additional medical software which requires processing information through multiple stages, proving to be tedious and time-consuming. The application developed using this process will eliminate these intermediate steps for visualizing virtual objects during surgery and will study the benefits of using augmented reality for tumor resection.

1.3.2 *Thesis goal*

This project proposes to integrate isosurface visualization on a mixed reality device and the incorporate the ability to interact with the data through multiple devices. Moreover, this project also proposes an improved performance time to view the reconstructed volumes during surgery without employing increased time to process information. The datasets used for research are made HIPAA compliant through an integrated module in the application. The chapters that follow describe the analysis of the problem, approach to the solution and a discussion of the results.

.

1.4 CONTRIBUTION

To the best of the authors' knowledge, this work is the first to integrate a native implementation of an isosurface visualization algorithm with mesh simplification on the Microsoft HoloLens device to test image registration for tumor resection. The works incorporates additional functionalities such as imaging data visualization, image registration and data anonymization.

Chapter 2.

BACKGROUND

2.1 Digital Imaging and Communications in Medicine

The Digital Imaging and Communications in Medicine (DICOM) standard has been established and widely accepted as a protocol for handling medical imaging and image-related information. It was developed by National Electrical Manufacturers Association (NEMA) and American College of Radiology (ACR) in 1993. DICOM provides compatibility between imaging modalities to reliably store information about patients, procedures, equipment, and to the medical images [20].
DICOM has been central to all medical imaging technologies for processing and conveying medical information reliably. This includes X-ray machines, Computed Tomography (CT) imaging, Magnetic Resonance Imaging (MRI) and Ultrasound amongst others. It is hierarchically structured and has a Client-Server architecture. This consists of data interchange protocol, image file or data format and network protocol architecture. These protocols can be used to retrieve information about the patient's imaging data during the surgery. Since it supports up to 65,536 shades of gray for monochrome image display, it offers increased image quality than other methods.
Currently, the DICOM Standard is divided into twenty parts [21]. These serve as a template for procedures and interface specification to enable connectivity between medical imaging technology and other systems. This allows data from different devices to be processed in an interoperable way. The standard is regularly modified to incorporate new technologies.
The data storage processes imaging data by linking it with the information of the patient. This is done by generating an examination number when image transfer is synced with Picture Archiving and Communications System (PACS). These are generally designated by the Radiology Information Systems (RIS) or Hospital Information System (HIS). In addition to patient data, each DICOM data object contains one attribute for pixel data. This usually relates to a single image for imaging systems although each attribute could allow storage of multi-frame data.
The DICOM committee defined a lookup table for consistently generating grayscale display images across devices. The devices used for viewing images can either use DICOM grayscale

standard display function (GSDF) or can be calibrated to the curve. Each of the grayscale images captured using imaging systems can be expressed as a continuous function $f(\boldsymbol{x})$. Here, $\boldsymbol{x}$ is the spatial coordinate vector and $f(\boldsymbol{x})$ is its intensity value. In this project, the $f(\boldsymbol{x})$ information is digitized to be processed by a computer.

2.2 COMPUTED TOMOGRAPHY IMAGING

Computed tomography imaging, commonly known as CT imaging, is a non-invasive imaging procedure that generates 2D cross-sectional images of the body [22]. In CT imaging a narrow beam of X-rays is used to scan internal regions of the body. Tomography is a method used to record the differences in effects on the passage of electromagnetic waves impinging on the 3D structures present in the body. CT imaging is also known as Computerized Tomography. The tomographic images/slices provide 3D information of the patient's anatomy, as opposed to the conventional X-ray. This includes information about bones, soft tissues, blood vessels and internal organs. Examples of sagittal, coronal and axial views generated using CT imaging are shown in Figure 1.

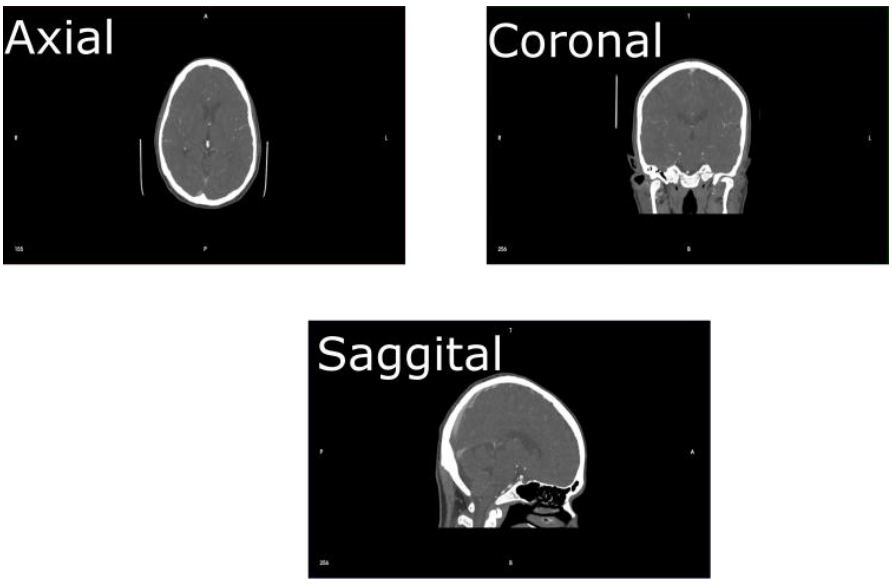


Figure 1: Sample Computed Tomography Image

CT scanners consist of two key components - an emitter and a receiver. The patient lies inside a tunnel-like equipment which is equipped with the scanner. The emitter is 180° across from the receiver. The patient is maintained in a stationary position while the scanner rotates around the

patient along the circumference of the tunnel, known as the gantry. X-rays are beamed and received at different points along this trajectory. Each time the bed moves to introduce a different anatomical structure, the scanner rotates to capture new information. A computer analyzes the information provided by the X-rays and constructs cross-sectional images of the body by transformation through mathematical reconstruction algorithms. Since different parts of the body absorb the radiation in varying degrees due to their radio-density, the image consists of differing shades of gray, to depict the level of absorption. This is crucial in identifying tumors and lesions certainly. The bones appear as bright regions as opposed to soft tissues, which are represented in shades of gray. Air appears black in a reconstructed image.

Biological Material	**Hounsfield Unit**
Air	-1000
Water	0
Soft tissue (Contrast)	100 - 300
Bone	700 - 3000
Muscle	35-55

Table 1: Approximate Hounsfield Unit for biological material

The CT scans used for this thesis consist of images of the head. They have a dimension of $512 \times 512 \times N$ voxels, where N is the number of slices in the image. The N values used during this project ranged from 50-350. These are generally stored by following the DICOM standard. The intensity values $f(\boldsymbol{x})$ represent the radio-density of the tissue and are denoted in Hounsfield Units [23], which represents a linear transformation of attenuation coefficients measured during tomography. These values range from -1024 to 3095.

A commonly occurring phenomenon in CT scans is the Partial Volume Effect (PVE) [24]. This is a result of a voxel enveloping more than one type of tissue due to low spatial resolution. The values are averaged, resulting in a blurry boundary between tissue images. Other commonly occurring issues are motion blurs from patient movement, streaks from metal implants and intensity noise. In this thesis, the images will be filtered to ensure robust processing of digital data.

When planning cancer surgery, surgeons use the cross-sectional images from a CT scan to determine the size and location of the tumors. The surgeons use multiple software to reconstruct

the multiplanar imaging information into 3D volumes. By interpreting this data, surgeons can determine approaches to the tumor even for complex procedures. This thesis generates the logic behind these mechanisms to model 3D objects using mixed reality to visualize the isosurfaces generated from a CT scan.

2.3 MIXED REALITY

Digital technology has been evolving at a rapid rate and has continuously integrated with our analog lives over last few decades. A development that is being extensively researched recently is the use of Head Mounted Display (HMD) devices. Though HMDs have been developed since the 1960s [25], it is only recently that the computing power has matched the needs of the user.

There are two major uses for HMD devices in development - Virtual Reality and Augmented Reality/Mixed Reality. Virtual reality (VR) is a computer-generated immersive user interface that consists of real-time simulation and interaction through multiple sensorial channels [26]. This could involve either exploring the environment or interactively modifying it. Virtual Reality devices include Oculus Rift, HTC Vive, Google Cardboard and others. Augmented Reality (AR) is the use of virtual objects to provide supplementary information about the real-world environment in real-time. These are provided using inputs such as visual display, sound, graphics or GPS data. A typical example of this is overlaying information about a monument of importance or providing directions based on relative position [27] in a form which can be superimposed on the user's normal experience. Mixed Reality (MR) is a combination of Virtual and Augmented Reality, where physical and virtual objects co-exist by interacting with each other. This is made possible by the deployment of holograms on the screen of the Head Mounted Device. Devices used for mixed reality are generally extended for use as augmented reality displays by generating information on the display along with other virtual objects. Some MR/AR devices that are currently studied are Glass 2.0, Microsoft HoloLens and Magic Leap. The Virtuality Continuum [28] is depicted in Figure 2. This was derived from the taxonomy defined by Paul Milgram and Fumio Kishino. They use three important factors to define Mixed Reality based on the medium of display used to provide visualizations. These are: (i) Extent of Knowledge, (ii) Reproduction fidelity and (iii) Extent of Presence Metaphor. The Extent of Knowledge determines the level of modeling involved in the development of the environment. Reproduction fidelity is the quality of the display device used to render the environment to provide the most realistic visuals. The extent

of presence metaphor ascertains the degree of perceived realism. For mixed reality devices, the user is not placed in a highly immersive environment and is aware of his or her surroundings.

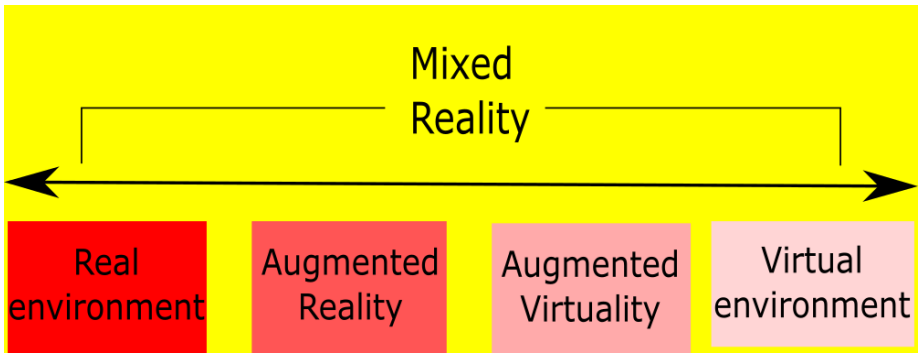


Figure 2: Simplified representation of Virtuality Continuum

The process of extending our reality has found a wide range of uses. These devices are currently used for research in medical, educational, industrial, military, and entertainment applications. Mixed Reality enhances the work environment by providing information to guide users through learning skills at a faster rate. The HMD devices for MR incorporate head tracking to ensure information is provided in desired locations. By using advanced computer vision algorithms and rendering processes, the user can engage with virtual objects in the real world.

Many of the devices being developed recently for Mixed Reality provide the complete processing power for image-based rendering onboard the device. This is the equivalent of using mobile devices for rendering. Hence, the performance capabilities of the device are not as high as devices tethered to a computer. The limitations of using a self-contained platform will be detailed in a later section.

Chapter 3.

METHODOLOGY

3.1 HYPOTHESIS

This project proposes to generate isosurfaces on the Microsoft HoloLens from 2D imaging data and smoothly render virtual objects with an initial high polygon count. The objects rendered will be used for interactive collaboration and aware of their environment.

3.2 MICROSOFT HOLOLENS

The Microsoft HoloLens, shown in Figure 3 is a mixed reality head-mounted display device that can render virtual objects in the real world. The “holograms” (term used by developers of the HoloLens) are projected onto the waveguide display, hardware that guides electromagnetic waves in the optical spectrum, which can then be manipulated and experienced by the user. The HoloLens then uses surface relief grating approach for extracting the images from the waveguide. It consists of combiner lenses behind tinted visor focus to display both 2D and 3D objects. The holographic lenses incorporate advanced optics, providing high luminance and enhanced resolution with low latency. The device also consists of two HD 16:9 light engines which enable 3D stereoscopic rendering, combined with the ability to automatically calibrate pupillary distance. The resolution of the device can drop to as low as 360p and has a maximum (default) of 720p. The HoloLens allows the user to visualize and work with digital content as part of the real world. This is made possible by the integration of various specialized components for holographic computing.

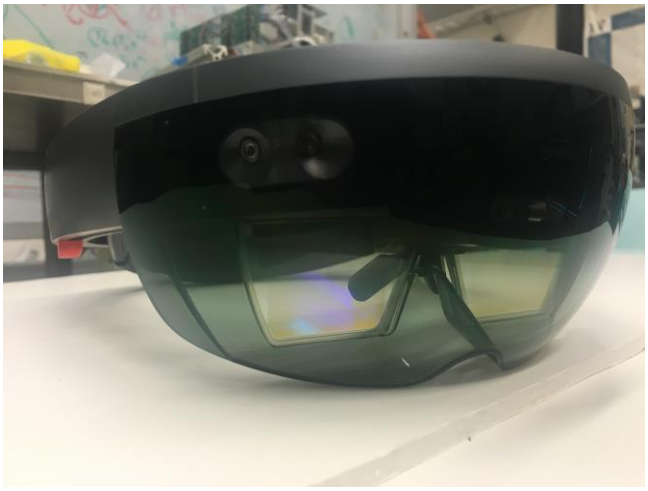


Figure 3: Microsoft HoloLens

The HoloLens optics system is backed by depth sensing and positioning technology on the device including an array of cameras, accelerometers, and gyroscopes for head pose estimation and gaze detection. This includes four environment-understanding cameras, two depth cameras, one 2.4 Megapixel photo/HD video camera, one ambient light sensor and an Inertial Measurement Unit (IMU). The depth camera uses the time of flight method [29] for active infrared scanning. The black and white wide Field of View (FOV) cameras (also known as environment-understanding cameras) are offset at an angle for positioning aspect in the device. The HoloLens does a spatial reconstruction of the environment and returns a 3D mesh. This provides the geometry of objects in the environment, in addition to the bounds and scale. The depth cameras and onboard sensors combined with optimized tracking algorithms on the HPU (Kalman filter- based sensor fusion are used for positional tracking of the user. The depth sensing capabilities also enable the detection of gestures by the user or others. These functionalities are referred to as sensor fusion on the HoloLens. The commonly used gestures are “airtap” and “bloom” Figure 4.

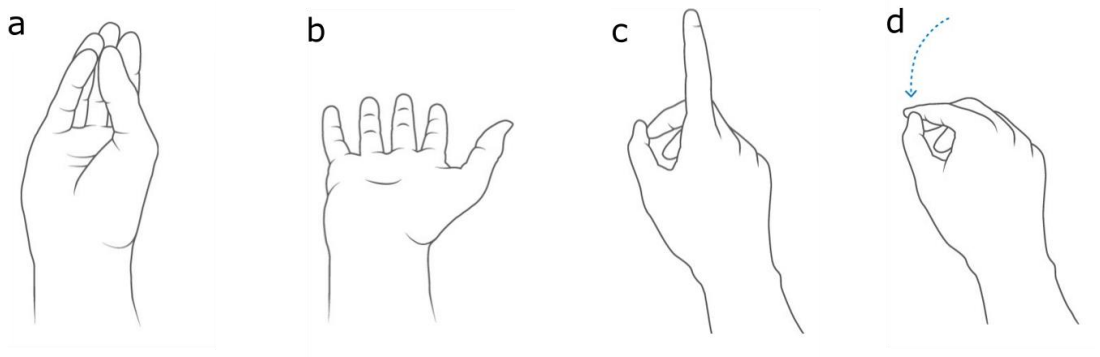


Figure 4: Gestures used in Microsoft HoloLens: (a)-(b) Bloom gesture; (c)-(d) Airtap Gesture

Four microphones and built-in speakers provide spatial audio functionality (Cortana) which allows parsing of the user's speech to commands and allows text-to-speech modalities. Unlike other speech processing devices, the user does not have to use wake words while communicating with Cortana on HoloLens. The speakers are binaural, which sounds like the sound source from virtual object rendered, simulating spatial effect. The use of gaze, gestures, and speech for interaction with "holograms" establishes a natural user interface, as opposed to having users adapt to external devices.

The HoloLens processes using an Intel 32-bit architecture with Trusted Platform Module 2.0 (TPM 2.0) support and a dedicated Holographic Processing Unit (HPU 1.0). It handles mesh processing, sensor fusion, and audio. This makes augmented reality practical since the processors can handle large amounts of data each second. It offers 2 GB RAM and 64 GB Flash. The device cannot process high polygon meshes or large textures (size over 2048) since the onboard processing is not powerful enough. HoloLens prioritizes a steady 60 frames per second to eliminate visual judders while rendering 3D objects.

The HoloLens runs Windows 10 software much like any other Windows device. Applications for the device can be developed on a Windows 10 computer with Visual Studio 2015 (or 2017) and Unity. The device has Wi-Fi 80.11ac and Bluetooth 4.1 LE wireless connectivity and a Micro-USB 2.0 port. Microsoft has also released an emulator for testing programs that can be deployed on the HoloLens device.

3.3 UNITY 3D

Unity3D is a game engine developed by Unity Technologies for developing cross-platform games in 2D and 3D. These could be deployed on Personal Computers (PC), consoles, mobile devices, Head Mounted Display devices and over the web. The personal edition of Unity is freely available for users to develop interactive content. It offers intuitive user interface tools and is customizable to fit any developer's needs. The Unity Editor provides highly optimized physically-based shading and enables the user to generate high-resolution graphics. It is not game-genre specific and is a feature rich game engine. It offers a great asset pipeline (the workflow for generating assets used in game development) and has a deep feature set.

For developing on Unity, the users can either use C# or JavaScript programming language. It natively offers Visual Studio and MonoDevelop integration. Unity supports iterative development as opposed to using DirectX or OpenGL and offers a relatively efficient engine in comparison to other game engines available. Unity3D is the most commonly used for developing games for Virtual Reality and Augmented Reality devices. Unity has been used for Universal Windows Platform (UWP) apps even before HoloLens was available publicly. UWP applications appear as 2D projections when deployed to the HoloLens.

Unity3D is the recommended platform for developing applications on the HoloLens. The user can start developing on the device after setting up the Windows SDK with Unity. The developer will need to switch between the Unity Editor and Visual Studio for completely prototyping an application package for the HoloLens. Once the Virtual Reality support is enabled on Unity, the Unity camera component handles the head tracking and stereoscopic rendering for HoloLens. The HoloToolkit made available on GitHub [30] provide commonly used prefabs for the HoloLens that can be imported into the Unity editor. These are helpful in rapidly developing applications for the device. Unity3D has an active ecosystem of assets and plugin creators and offers cross-platform integration for over 25 platforms. The Asset Store offers a wide variety of objects that can be used in projects for free or a minimal fee.

For this project, Unity3D was run on a laptop with 64-bit Windows 10 Student Edition, 16 GB RAM, Intel Core i7-4720HQ CPU 2.6 GHz Processor, and a NVIDIA GeForce GTX 960M graphics card. The development environment looks as shown in Figure 5.

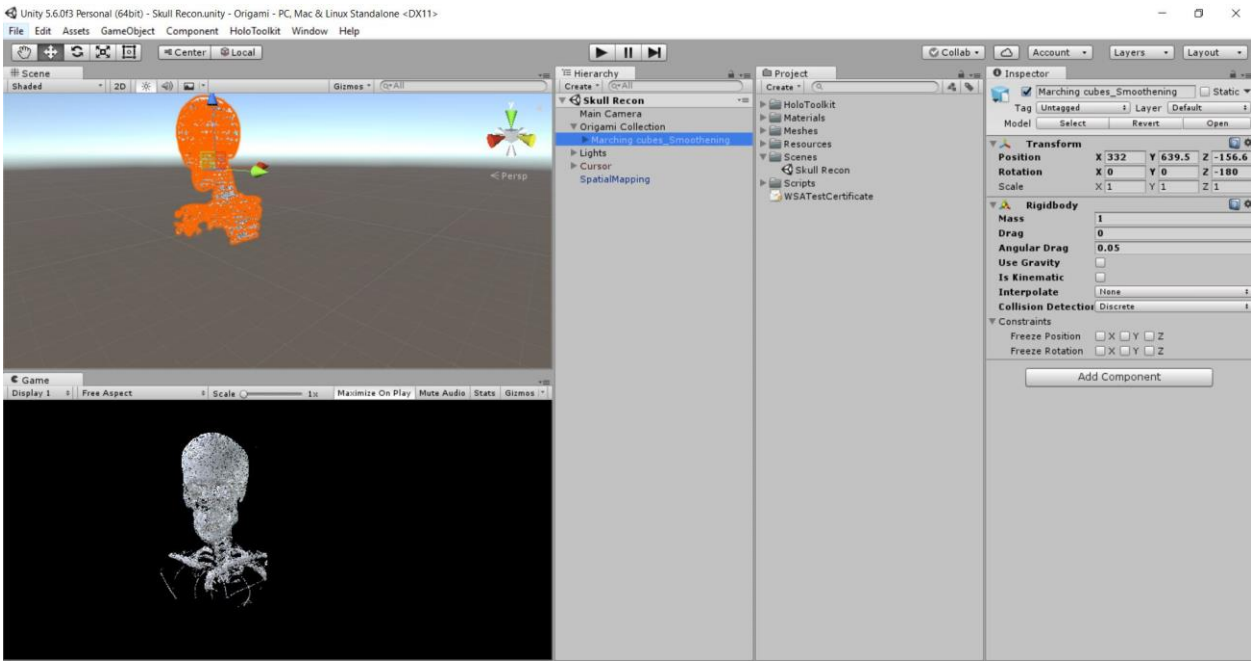


Figure 5: Sample of Unity3D Work Environment

3.4 WINDOWS PRESENTATION FOUNDATION

The Windows Presentation Foundation (WPF) is a Graphical User Interface (GUI) framework from Microsoft. It was initially released as a subset of the .NET framework 3.0. It uses Microsoft DirectX, a collection of Application Programming Interfaces (APIs) for game development, to render user interfaces. The GUIs generated using WPF can contain the traditional elements like textboxes and buttons that can be developed using C# and Visual Studio. The interfaces developed with WPFs provide more flexibility in comparison with WinForms by Microsoft.

The Windows Presentation Platform consists of a vector-based rendering engine which is independent of resolution. It creates desktop client applications. WPF supports features such as graphics, documents, data binding, controls, amongst others. It uses features of Extensible Application Markup Language (XAML), a declarative markup language, for developing applications. It allows developers to work independently on the User Interface (UI) and the application logic.

3.5 WORKFLOW ON HOLOLENS

The system developed in this thesis is an application module that acts an interface between the input DICOM CT images and the output HoloLens device which would be used by a surgeon. The system developed is depicted using a simplified architecture diagram in Figure 6.

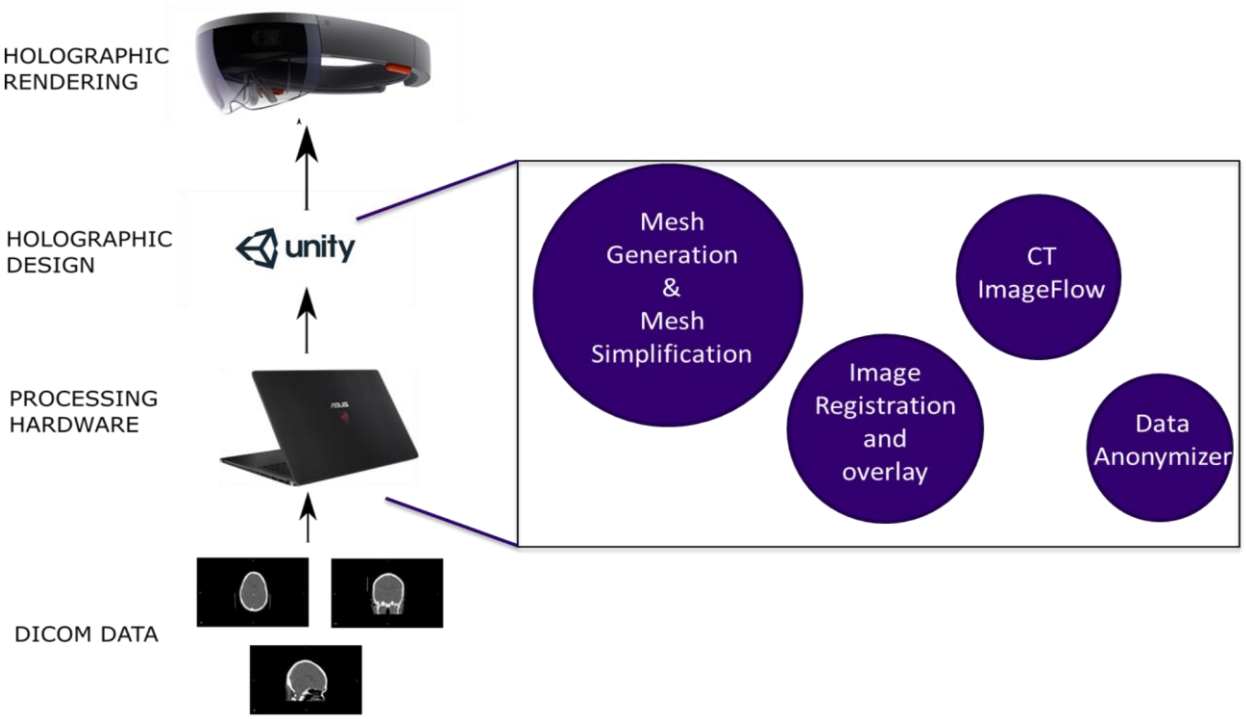


Figure 6: Sequence diagram of system developed

The DICOM images are provided as an 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. The mesh rendered is simplified using mesh reduction techniques described in a later section. The 3D mesh is registered to a fiducial marker location and is overlaid on a test setup mimicking a human patient. In addition to providing the user with the ability to view the 3D surface, a process is developed to view all the slices from a patient's CT scan directly on the HoloLens device, instead of using an intermediate monitor. This is defined as a "CT ImageFlow" in Figure 6. Additionally, to protect the privacy of the individuals during research and development of the system, a data anonymizer is developed to ensure the information cannot be traced back to

an individual. The scripts were tested using the UI elements of Windows Presentation Foundation before being adopted into Unity.

3.5.1 *Marching Cubes*

Three-dimensional images are routinely used during surgery to visualize the anatomy of a patient. These images can be enhanced (but in the operating room, they often still are not) with the help of volume rendering techniques. There are two broad categories for volume rendering– Direct Volume Rendering (DVR) and Indirect Volume Rendering (IVR). This project made use of an IVR technique called Marching Cubes [31] which is a sequential-transversal method. An isosurface is a surface of constant value, α that satisfies a scalar field $F(P)$, where F is a scalar function in $\mathbb{R}^3$. Marching Cubes (MC) is an isosurface visualization algorithm that converts two-dimensional images into 3D volume by generating a mesh with triangles of constant density.
Here, the constant value α represents the surface of choice, in our case bone or soft tissue. The algorithm considers two slices of the dataset at a time to generate triangles for individual voxels. Each CT slice has pixels of constant scalar value which are used as vertices for constructing a grid cell. The corresponding row and column values for each pixel are chosen in adjacent slices to form a solid cube. This is represented in Figure 7.

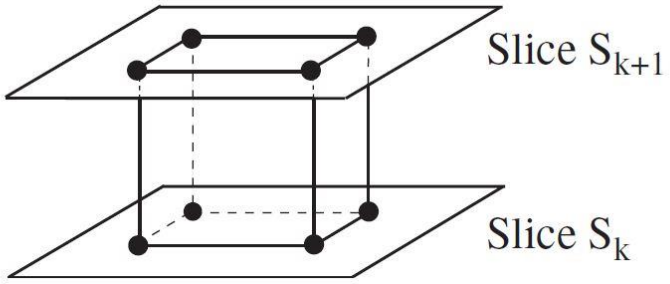


Figure 7: Grid cell from two CT slices

Each lattice point has a scalar density value. By comparing each of the vertex values, an 8-bit array index is created for each grid cell. The value represents if each vertex is either above or below an isovalue/threshold. By using binary labeling for each voxel, the bit representation for a vertex is assigned 1 if the value is greater than the isosurface value (outside the surface) or 0 if the value is less than the isosurface value (inside the surface). By making use of this index, it can be evaluated

which edges intersect with the isosurface. Though there is a possibility of obtaining 256 different cases of edge intersection, by using properties of symmetry and rotation, we can use 15 different base cases to determine edge intersections. These are as shown in Figure 8.

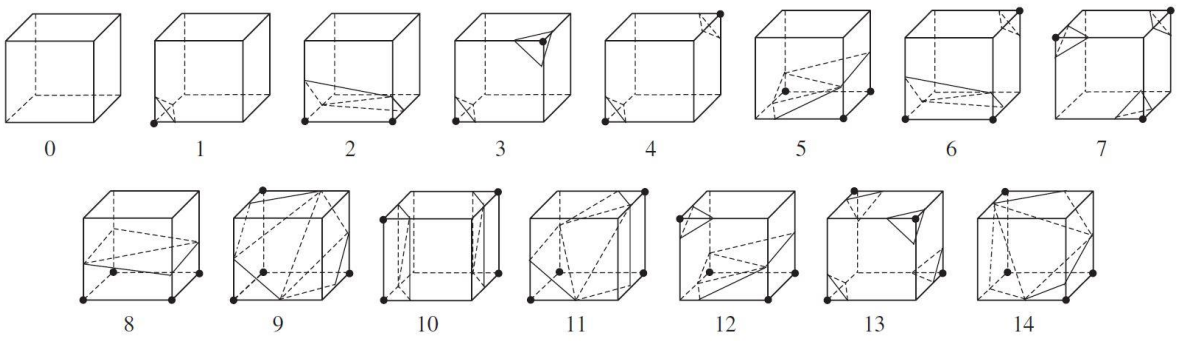


Figure 8: Bases cases in Marching Cubes [27]

Information about the various edge intersections are stored in a lookup table. The index for a grid cell acts as a pointer to obtain information about the various intersections. This information is used to generate possible triangles for each cube. For each triangle, the vertex locations are determined by the process of linear interpolation of voxel values.
This is given by the formula:

$$x = i + \left(\frac{T - V[i]}{V[i+1] - V[i]}\right) \tag{1}$$

where T is the isovalue, i is the vertex number and $V[i]$ is the value of the Voxel at given vertex. Since the triangles produced are restricted to less than 5 in each grid cell, using any other form of interpolation offers little improvement. Finally, the normals for each triangle vertex is determined to produce shaded images. This ensures that the triangles are calculated in a predefined direction and will always be visible to the user. The gradient at each of the cube vertices is determined through central differences. The equation for the method is given in equation (2).

$$G_{(axis)}(i,j,k) = \frac{D(i+1,j,k) - D(i-1,j,k)}{\Delta(axis)} \tag{2}$$

where G is gradient, D is density and (axis) is x,y or z direction.

3.5.2 *Mesh Simplification*

The meshes generated by using Marching Cubes produce a large number of vertex points for most datasets. These ensure a high-resolution mesh but the resulting large number of polygons reduces performance when used on a device such as the HoloLens. To increase the ease of use of HoloLens in real-time, it is essential to reduce the number of triangles rendered at a given instant. For this purpose, mesh reduction components were implemented.

Before mesh simplification techniques were adopted, simulations on Unity evidenced the excessive number of draw calls for the meshes drawn using the marching cubes algorithm. This was a result of the restriction imposed by Unity 3D to limit each mesh to a maximum of 65,534 vertices. However, the meshes generated consisted of 1.4 million vertex points, which were broken into 30 different mesh components. For drawing each one of these game objects, the engine issues an individual draw call to the graphics API. Every draw call is resource-intensive and causes performance overhead on the CPU side. The number of draw calls for the mesh generated from marching cubes was a total of 35. This is computationally expensive for a single mesh being statically rendered. Hence, the meshes were combined to offer a reduced number of draw calls when being viewed statically on the HoloLens device. The combination of meshes while rendering was implemented by using the process of tagging components on Unity 3D. Since the components have the same material properties, it is easier to render them together. This method renders discrete meshes as a single mesh and is useful if individual meshes are unlikely to cull. Figure 9 shows that for the mesh generated, the isosurface game object can be rendered with a reduced number of only 3 draw calls.

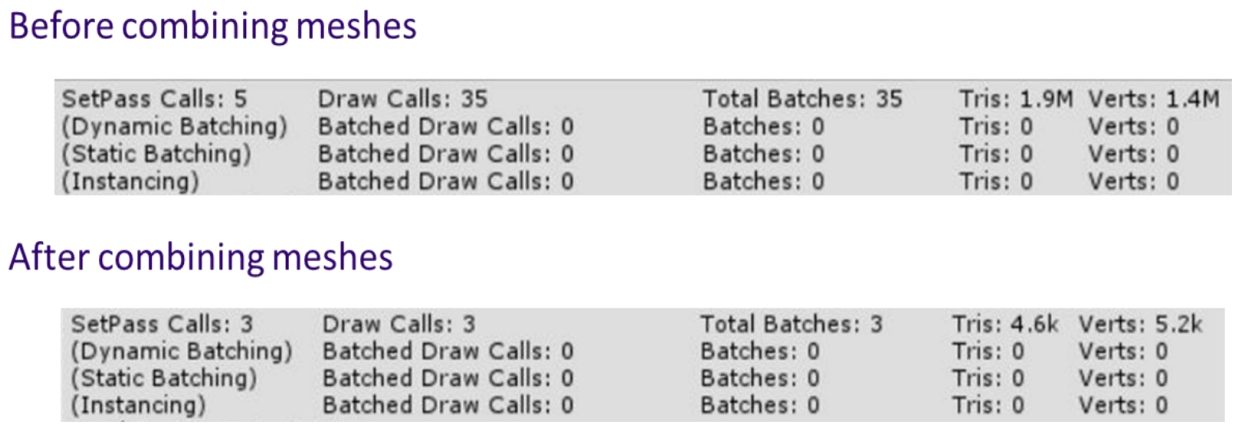


Figure 9: Render call comparison

Though this method is useful while rendering information about static meshes, it does not prove highly effective if the meshes need to be taken apart for analysis. For this purpose, we further reduce the mesh size by evaluating and modifying the properties of the triangles. The parameters used for mesh reduction are edge length and curvature of the surface. The surgeon is provided with the ability to choose the level of reduction for meshes in terms of percentage of original vertex count.

The simplification of the mesh is first processed by identifying and ensuring there are no repeated values for vertices. The mesh is taken as input and an internal dataset is constructed for ensuring the vertex data are truly different by identifying the UV values and color in addition to spatial locations. The entire process is developed my making the functions serializable in Unity. A class called UniqueVertex handles the processing of the vertex points of the mesh. By generating new faces and vertices, the repetition of vertex points in the 3D space is overcome. The bone weights of all the vertex points contribute to obtaining information for mesh simplification.

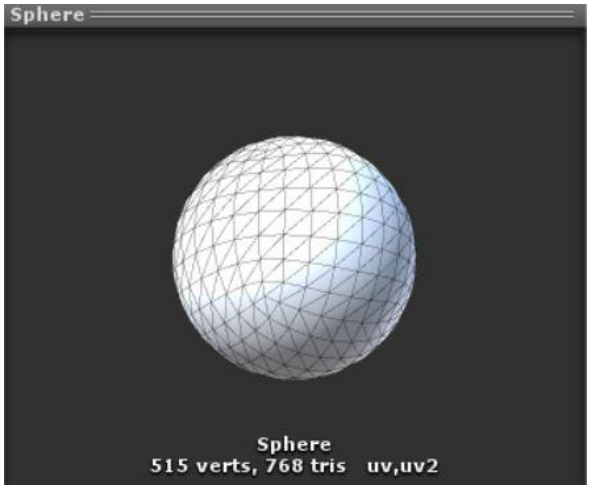
The original mesh is simplified by the process of skinning mesh components. The program recursively moves through each mesh in each tree to reduce all the child nodes equally. The mesh parameters used are curvature and edge length. The curvature determines the angles between the triangles and eliminates cases which produce a smoother transition between surfaces, that is, have a large value. This method is employed to ensure meshes in each surface plane can be combined to produce one large triangle instead. The edge length configuration eliminates a percentage of the small edges that form triangles to be combined with rather larger edges instead. This ensures the final mesh obtained would produce lesser triangles than otherwise generated. A cost function from

[32] is used for ensuring a maximum number of short edges are eliminated without affecting the surface rendered. For consolidating the mesh after simplification, the new vertex values are obtained and added to the mesh based on their existence in the previous list. If the value is not available, a new index is created and the triangles are generated based on the relation of the newly generated vertex points with respect to each other.

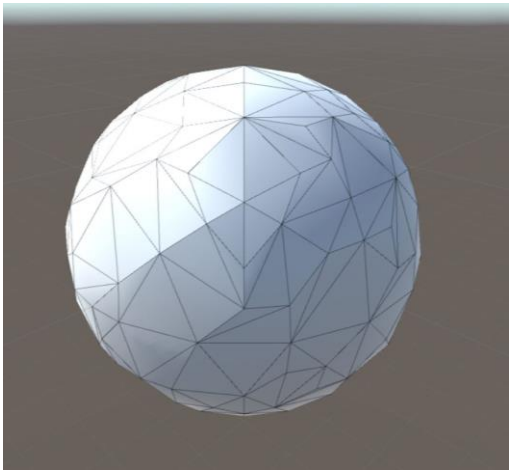
For analysis, the experiment was tested on a simple sphere with 515 vertex points. The results obtained while choosing the two parameters individually to study the effect on vertex count are represented in Table 2. By denoting the extent of curvature change and dependency on edge length for modification, mesh reduction for the parameters are shown in Figure 10.

Curvature	Edge Length	Vertex Count
Yes	Yes	437
Yes	No	429
No	Yes	373
No	No	346

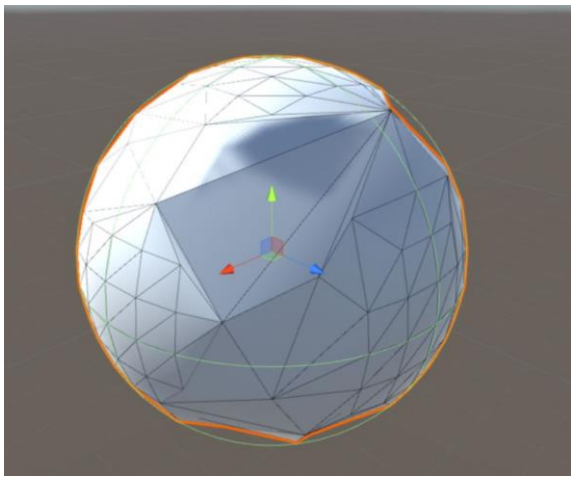
Table 2: 50% reduction of mesh with 515 vertices



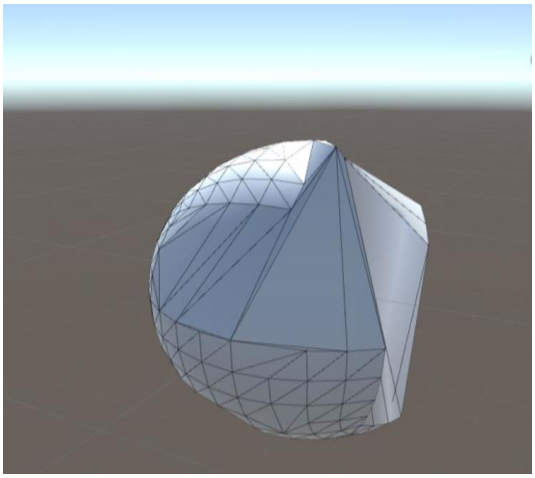
(a)



(b)



(c)



(d)

Figure 10: Mesh simplification on sphere: (a) Sphere with 515 vertices, (b) 50% simplification with curvature, no edge length elimination, (c) 50% simplification with edge length elimination, no curvature, (d) 50% simplification with edge length elimination and no curvature.

3.5.3 *Volume Registration*

The generated isosurface should be registered to the patient for preoperative planning by surgeons or for use during surgery while visualizing with respect to the patient. This can be done using one of two methods - the 3D surface could either be manually overlaid on the patient using tap gestures to pick and place virtual objects, or fiducial markers could be used to place them in a certain location as defined by patient position and orientation.

The HoloLens incorporates a feature called Spatial Mapping which provides a detailed representation of the environment in the format of a mesh. This information is used to generate 3D game objects which are fully aware of their environment and hence provide interactions similar to the real world. Spatial Mapping uses ‘Spatial Surface Observer’ and ‘Spatial Surface’ within a world space coordinate system. This information is generated every time the HoloLens device is turned on. To reduce the rendering cost, the spatial map of the environment can be mapped, stored and loaded for processing. Since the surgeon will not be dynamically moving between rooms while using the HoloLens and will not require spawning virtual objects frequently, the 3D map of the

operating room can be stored and loaded for reducing computational complexity. Figure 11 represents the map generated for the room used during experimentation.

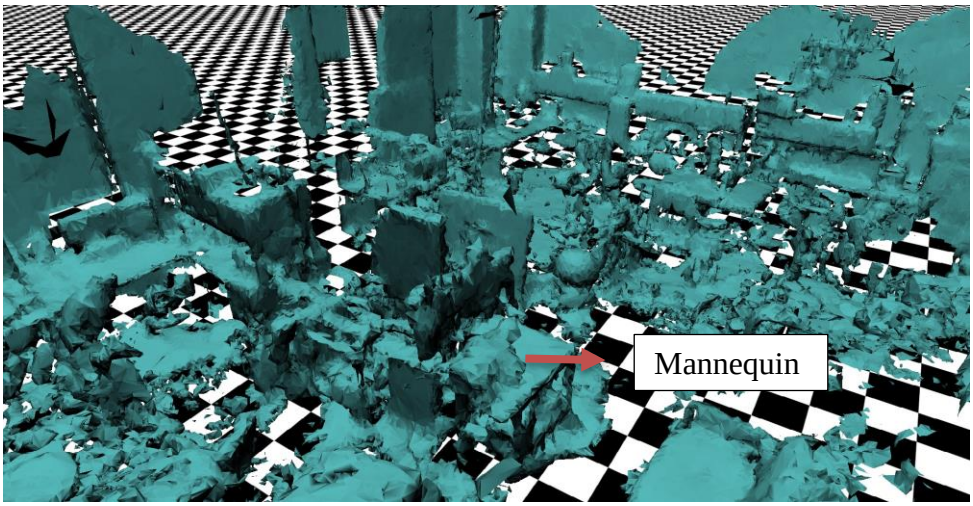


Figure 11: Spatial map of room used for experiments. Test subject (Mannequin) is located at the center of the room

The spatial map has five different purposes that the information is commonly used for - occlusion, visualization, placement, physics, and navigation. The spatial map offers users the ability to place hologram on top of surfaces and develop features with more realism. The spatial map is generated immediately after starting an application. While using the device in an operating theater, this includes a map drawn around the patient combined with the rest of the environment. However, this prevents the placement of holograms by overlaying it on the patient and are restricted to be placed atop the spatial map instead. To overcome this problem, a bounding volume was generated to interact with the spatial map around the region of interest. A sample bounding sphere is generated to depict the use in Figure 12.

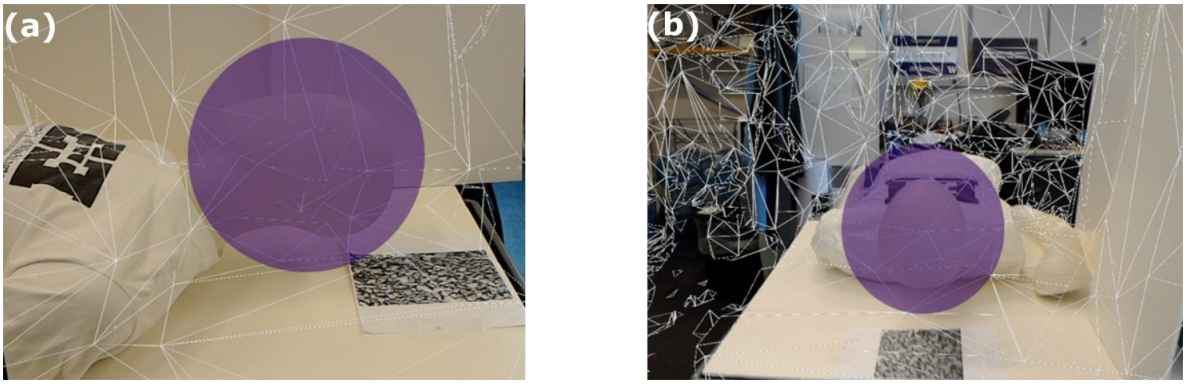


Figure 12: Bounding sphere (a) Lateral view (b) Axial view

The vertices that fell within this bounding volume were removed to generate a new set of vertices for spatial mapping. This allowed users to place holograms on surfaces that were initially classified as surface planes.

For overlaying holograms by using fiducial markers, Vuforia Software Development Kit (SDK) [33] was used. The SDK uses QCAR algorithm to track feature points in an image. By choosing an image with multiple feature points, the camera on the HoloLens device can be used to track markers quickly. Each 3D game object is associated with a certain fiducial marker. This allows us to store information about the multiple imaging data (preoperative or intraoperative) generated for one patient to different fiducial markers, or even information for different patients to certain fiducial markers. The test setup with the fiducial marker used for experimentation is provided in Figure 13. The fiducial markers could be replaced with barcodes or ArUco markers if desired.

Figure 13: Test setup with fiducial marker

3.5.4 *DICOM CT ImageFlow*

Surgeons tend to view CT slices intermittently during surgery in addition to the 3D data provided to them. This happens due to the traditional methods of training that surgeons undergo, which increases their reliability on such devices. For this purpose, a mixed reality-based CT image viewer was developed as a part of the application for isosurface visualization. This allows the surgeons to view CT slices directly on the HoloLens device in addition to the isosurface mesh rendered. In this viewer, surgeons make use of gestures for controlling and viewing the CT images without having to interact with traditional interface devices, which may not be sterile or require sterilization. The images are imported from folders in a location specified by the user. A slider element is provided in the UI to offer ease of interaction since most software used by surgeons use sliders. Two functions, MoveToPreviousImage() and MoveToNextImage(), transfer the current view to either the previous CT slice or following CT slice respectively. A test prototype for the user interface script was developed using the Windows Presentation Foundation. The screen also shows the slice and file number of the desired image. A screen capture of this interface is provided in Figure 14.

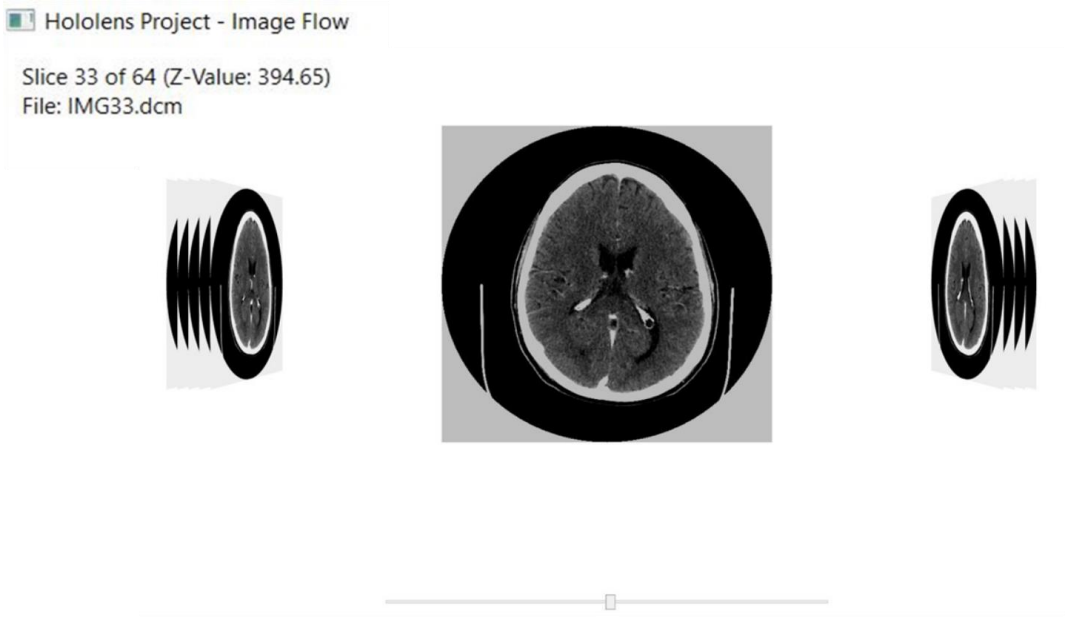
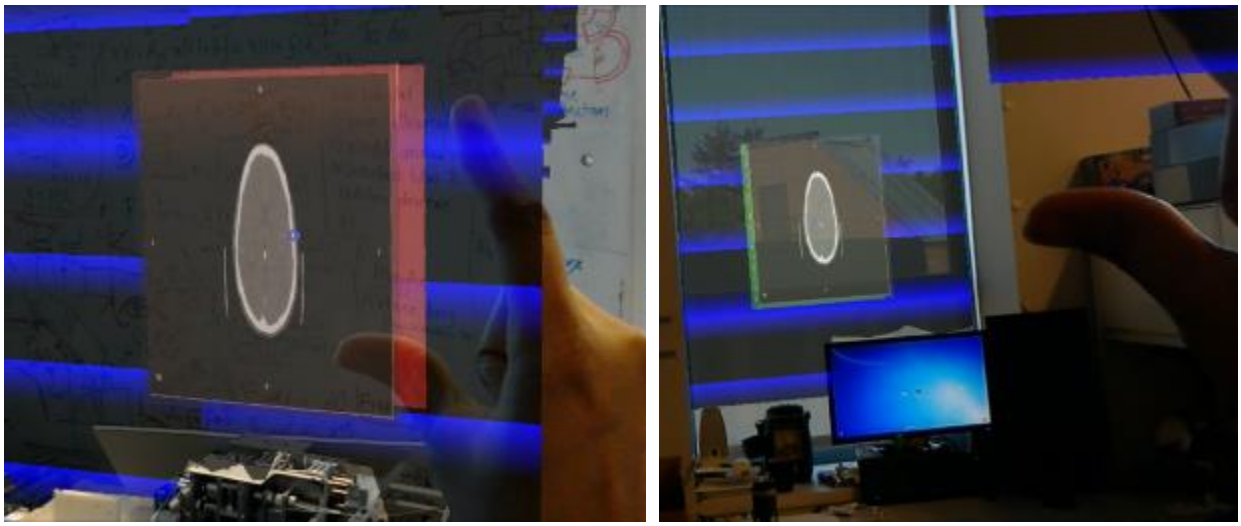


Figure 14: CT viewer tested using Windows Presentation Foundation

A canvas game object in Unity3D acts as an interface element for dynamic viewing of scans. The model developed on the HoloLens device also aids surgeons in choosing locations in the spatial

map for the canvas with favorable viewing conditions. The spatial map is created to represent the walls of the operating room using a SurfaceMeshestoPlanes script. This script allows the developer to generate planes for ceilings/walls/floor and is a part of HoloToolkit. A SpatialMappingManager component allows the definition of walls using distinct material properties. Since the map is generated in advance, processing this information does not require additional memory. When placing the CT canvas in the mapped OR, the surgeon can move the canvas using tap gestures.
The Placeable() script uses the plane finding data and provides feedback on whether a surface will fit a desired game object. This is determined by using raycasting from the center and four corners of a bounding cube. A red shadow is generated when the hologram cannot be placed on a wall surface and a green shadow denotes placeable surfaces. The results of the Placeable function are shown in Figure 15.



(a) (b)

Figure 15: CT Image Canvas (a) Non-placeable surface (b) Placeable surface

3.5.5 *DICOM Information Anonymizer*

The Health Insurance Portability and Accountability Act (HIPAA) is a law in the United States of America that safeguards the patient's medical information by providing standards of privacy and security for processing healthcare transactions. HIPAA Privacy Rule requires compliance with the Protected Health Information (PHI) regulation, which protects the health information of an

individual from being uniquely identifiable. PHI includes information about the common identifiers, such as the name of a patient, their social security number, address and birthdate, and provision of healthcare and payment by the individual that make them uniquely identifiable.

Since the patient information will be stored in the files that we use and load into the HoloLens, it is important to protect this information from being hacked and misused. For this purpose, the information loaded into the application will first be anonymized by either eliminating or substituting value fields and stored under an alias for future reference by the hospital. By following Sections 164.514 (b) and (c) of the HIPAA Privacy Rule [34], the de-identification standards are incorporated into the software. The standard also defines specifications for re-identification of health information to assign codes to the patient health information. This minimizes privacy risks to individuals and ensures the datasets do not provide a reasonable basis to identify an individual.

All the information from the DICOM files is stored in an XDocument after it is loaded into the application developed. XDocument is a .NET class that represents an XML Document and additionally provides support for Language Integrated Query (LINQ). The Information Object Descriptions (IOD) are recipes of items that define instances of objects. The various DICOM attributes are defined in modules, such as Patient Module (Patient name, Patient ID, etc.,) and Study Module (Study Unique Identification Number, Date, etc.,). In the metadata, DICOM attributes have defined tags, value representation (VR), value length and a value field. The information that should be anonymized can be retrieved using the tags and provided with either a random code or erased to protect patient privacy. After de-identification, the information would look like the representation provided in Figure 16. The information about the Referring Physician's name and other identifiers were cleared to have no attribute values. The Patient Name and Patient ID were provided with a random value to make the dataset uniquely identifiable for internal use. This process would be carried out for all other common identifiers.

(0008,0090)	Not in Dictionary	
(0008,1010)	Station Name	
(0008,1030)	Study Description	CTA HEAD AND NECK WITHOUT AND WITH CONRAST
(0008,1032)	Procedure Code Sequence	
(0008,103E)	Series Description	Head W
(0008,1050)	Not in Dictionary	
(0008,1060)	Name of Physician(s) Reading Study	
(0008,1070)	Operators' Name	
(0008,1090)	Manufacturer's Model Name	Revolution CT
(0008,1110)	Referenced Study Sequence	
(0008,3010)	Not in Dictionary	1.2.840.113619.2.416.203607755739041121679600211215836707449
(0009,0010)	Private Creator	GEMS_IDEN_01
(0009,0011)	Private Creator	GEIIS
(0009,1001)	Not in Dictionary	43 54 5F 4C 49 47 48 54 53 50 ....
(0009,1002)	Not in Dictionary	52 45 56 31
(0009,1004)	Not in Dictionary	52 65 76 6F 6C 75 74 69 6F 6E ....
(0009,1027)	Not in Dictionary	BD 16 9A 57
(0009,10E3)	Not in Dictionary	
(0009,1110)	Not in Dictionary	FE FF 00 E0 C6 0F 00 00 28 00 ....
(0010,0010)	Patient's Name	002
(0010,0020)	Patient ID	1002109429

Figure 16: Data Anonymizer viewed using Windows Presentation Foundation

3.5.6 *Sharing Service*

The Microsft HoloToolkit offers a library named Sharing for setting up communication between multiple HoloLens devices. The Sharing Service executable runs on a device that acts as the server for the entire communication process. A client library is included in the application to run simultaneously on multiple devices. Communication between the server and client modules is established using sockets. Figure 17 provides an overview of the server-client implementation.

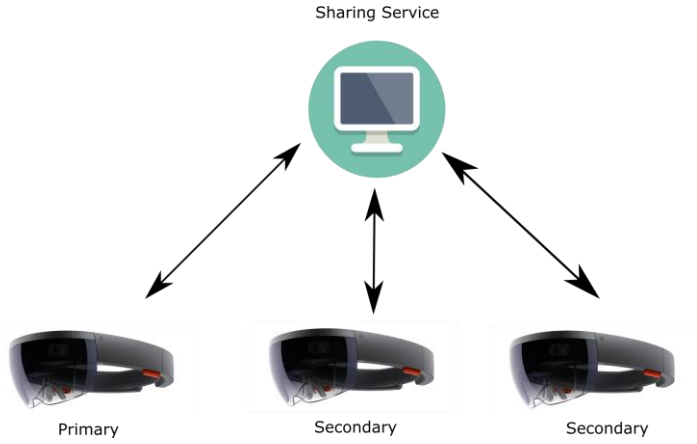


Figure 17: Server-Client module on HoloLens

A primary device starts the application and renders the virtual object representing the isosurface. This is placed at the desired location using gestures. The “holograms” could also be anchored using registration through fiducial markers. The world space coordinates of virtual object placement are then uploaded onto the server. Any secondary device connecting to the shared experience can

download and view this isosurface from a different perspective. This would allow users to visualize the same isosurface in 3D to communicate and plan for surgeries better than they could otherwise.

Chapter 4.

RESULTS

The marching cubes algorithms with the limitation of the number of triangles generated produce an isosurface to represent the desired anatomy of the patient. Figure 18 represents an isosurface with an isovalue, α value of 500 Hounsfield units to represent bone.

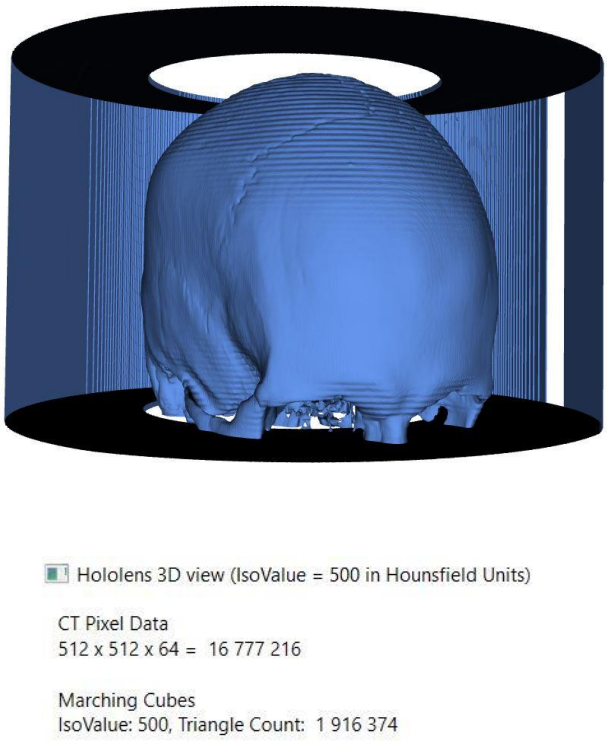


Figure 18: Marching Cubes – Bones (Isovalue: 500 Hounsfield Units)

Figure 19 represents an isosurface for α of -800 Hounsfield units that constitutes soft tissue information. These models can be used in combination while rendering on the HoloLens since they would maintain the same coordinate system with respect the each other. A tumor consists of a different isovalue from healthy tissue. This would allow the tumor to be segmented in the exact location as obtained from the imaging data and be included in the visualization.

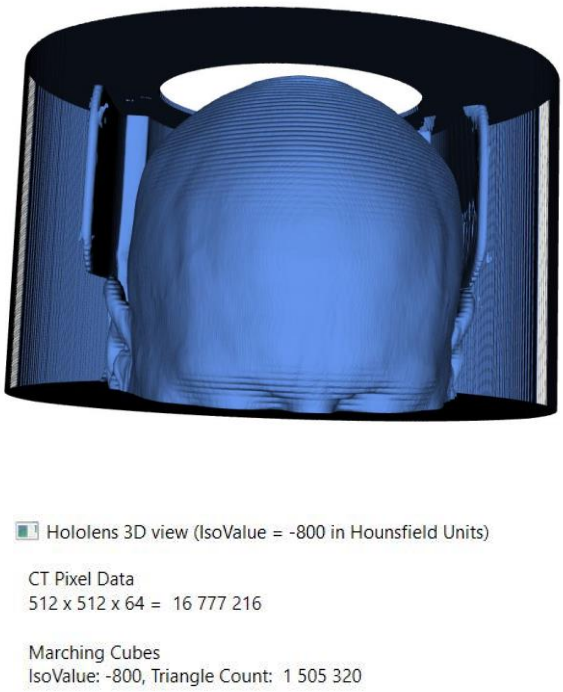


Figure 19: Marching Cubes – Soft Tissue (Isovalue: -800 Hounsfield Units)

The number of vertices and triangles generated for both the isosurfaces are over a million points in the 3D plane. This generated mesh was then reduced using mesh simplification techniques using two primary parameters – curvature and edge length. The percentage reduction allows the simplification of the mesh down to 1% of the original size. The result of mesh simplification to varying percentage levels are documented in Table 3.

Vertex Percentage (%)	Vertex Count
100	1,437,115
50	689,343
30	429,658
10	145,105
1	13,763

Table 3: Mesh simplification to reduce vertex count

In addition to reduction in number of draw calls as mentioned in 3.5.2, the edge length and curvature parameters are used to further specify and reduce the number of vertices and triangles

generated for mesh rendering. This information is provided in Table 4 for a 50% reduction of the original mesh.

Curvature	Edge Length	Vertex Count
Yes	Yes	689,343
Yes	No	689,337
No	Yes	689,342
No	No	689,149

Table 4: 50% reduction of bone isosurface

Though the vertex count for the cases considered is different only by a small margin, the representation of the surfaces can be as starkly different as shown in Figure 10 for the spheres. The algorithm ensures the surface stays intact and does not have any holes generated due to the mesh reduction process. This is the prime reason for the reduction of vertex count to be different from the exact number obtained mathematically. For example, 50% reduction a default sphere from Unity (with 515 vertex points) results in more than 258 vertex points, as shown in Table 2. By decreasing the amount of penalty imposed on removing a percentage of edges with shorter length, it would be possible to further reduce the number of triangles generated. However, this would also mean the surface generated may not provide refined information as is currently obtained when the curvature and edge length parameters are used. Downsampling the image datasets would reduce the number of pixels available for generation of meshes, thereby resulting in lesser number of triangles when using the Marching Cubes algorithm.

By using gestures, the surgeons can interact with the model generated and communicate with others involved in the shared experience. The hologram can be placed in the desired location and rotated while being anchored. Another form of registration is using fiducial markers with multiple feature points. Using a reduced number of features would result in a delay of volume rendering and hence was not preferred. A fiducial marker will be placed at a predetermined distance and orientation with respect to the patient/test subject. Once the HoloLens detected the feature points of the object loaded into the program as a fiducial, the associated isosurface object would be rendered and overlaid on the patient/test subject. The user can make use of gestures to move the surface and anchor in the desired location, if necessary. Figure 20 represents the output of volume registration to a mannequin used as a test subject.

Figure 20: Image overlay using fiducial markers

The surgeons can move around the rendered isosurface and visualize the CT slices if desired on the same device. The sharing service setup also provides surgeons the ability to communicate about the anatomy and surgical approach visually using multiple HoloLens devices during surgery without disrupting the workflow.

Chapter 5.

DISCUSSION

The results obtained through mesh rendering show the Level of Detail a generated mesh could possess about the anatomy of a patient. The 3D volume reconstructed from the 2D scans relay fine details about the cracks, folds, and any tissue deformation present in the anatomy reliably. This can be visualized using the HoloLens device for either preoperatively planning surgeries or using the 3D model intraoperatively to gain enhanced information about the location of critical structures. However, tests on the HoloLens device evidenced the need for reduction of the size of meshes rendered. This is a result of the HoloLens being a mobile device with limited onboard processing capabilities. With mesh simplification techniques, the number of vertex points required for defining a structure is reduced for smooth rendering.

The process of combining meshes using the tagging process reduces the number of draw calls required for visualizing the surface information. By the process of manipulating the spatial mapping mesh, it is possible to overlay the isosurface generated on the target object, here, a mannequin. The surgeon can generate meshes to a certain degree of simplicity based on the level of detail desired in the final mesh rendered. By using the technique used in [32] the surgeon can specify the sections of the mesh that need to have a higher value of detail as opposed to a different section of the same mesh which could contain lesser information.

The application provides surgeons the ability to view the rendered surfaces and interact with the models by incorporating a range of techniques that have not previously been introduced onto the HoloLens device. This work is the first in the literature, to the best of our knowledge, to incorporate an array of functions that will be beneficial for the surgeon, including visualization and interaction with isosurfaces, registration to patient anatomy and on-board CT image viewer. Since the test methods were developed for smooth rendering of 3D models, it is not restricted to any one type of surgery. This allows us to extend the use of our current application to any type of surgery that could benefit from receiving augmented information. The information obtained from the datasets are also made HIPAA compliant and hence do not violate any standards set to protect patient privacy.

Chapter 6.

FUTURE WORK

The algorithm to reduce mesh size of isosurfaces explores the advantages of generating a low polygon count while rendering objects. However, the size of generated mesh is still large for a mobile device like the Microsoft HoloLens to handle. Supplementary algorithms for improving the mesh performance can be implemented to benefit the use of the HoloLens device for a much smoother interaction during real-time. The Delaunay triangulation [35] method restricts the number of points by avoiding points inside the circumcircle of triangles and can be explored to verify effect on mesh simplification. In addition to testing the variations of Delaunay triangulation, other methods for mesh simplification will also be explored. The retaining of anatomical information with finer detail will be provided a higher cost while reconstructing meshes for visualization as opposed to regions which could be handled with lower mesh sizes. Due to the exploratory nature of the project, the use of the device was only tested with a mannequin. This will be extended for use during live surgery for studying the benefits of HoloLens in the operating theater.

In addition to using gestures for controlling holograms, voice commands will be incorporated to provide surgeons a secondary method of communication with the HoloLens. Voice commands can be used to start the holographic rendering process, for scaling objects, manipulating data or for object placement. The robust voice recognition system employed on the HoloLens device consists of multiple microphones which can perform well despite the background noise from equipment present in operating rooms.

Using multiple fiducial markers would allow improved accuracy for registration. Currently, the system has a delay of approximately one second when the user shifts view to an entirely different cross-section of the 3D object. This could also be a result of the large mesh size. A user interface will be developed to provide a simple interaction module that surgeons can seamlessly navigate and use throughout the surgery. This will be mindfully developed by surveying and including UI factors that would provide a steady transition for surgeons to view, dissect and study the volumetric data.

A software for preoperative planning of skull base surgery is under development at the Biorobotics Lab. The models generated will be introduced onto the HoloLens device to enable surgeons to collaboratively visualize surgical plans and use this information to increase the quality of tumor margin excision through augmented reality. The use of the HoloLens device can also be extended for object tracking to provide information about the distance between instrument tip and tissue in real-time. The large deformations that occur due to breathing need to be accounted for during registration.

BIBLIOGRAPHY

[1] J. C. Hu *et al.*, "Comparative effectiveness of minimally invasive vs open radical prostatectomy," *Jama*, vol. 302, no. 14, pp. 1557–1564, 2009.

[2] K. Nagpal *et al.*, "Is minimally invasive surgery beneficial in the management of esophageal cancer? A meta-analysis," *Surg. Endosc.*, vol. 24, no. 7, pp. 1621–1629, 2010.

[3] D. W. Easter, "A surgeon's perspective on laparoscopic cholecystectomy," *AJR Am. J. Roentgenol.*, vol. 157, no. 2, pp. 241–242, Aug. 1991.

[4] U. Sure, O. Alberti, M. Petermeyer, R. Becker, and H. Bertalanffy, "Advanced image-guided skull base surgery," *Surg. Neurol.*, vol. 53, no. 6, pp. 563–572, Jun. 2000.

[5] S. DiMaio *et al.*, "Challenges in Image-Guided Therapy System Design," *NeuroImage*, vol. 37, no. 0 1, pp. S144–S151, 2007.

[6] P. Breedveld, H. G. Stassen, D. W. Meijer, and L. P. S. Stassen, "Theoretical background and conceptual solution for depth perception and eye-hand coordination problems in laparoscopic surgery," *Minim. Invasive Ther. Allied Technol.*, vol. 8, no. 4, pp. 227–234, 1999.

[7] A. F. Durrani and G. M. Preminger, "Three-dimensional video imaging for endoscopic surgery," *Comput. Biol. Med.*, vol. 25, no. 2, pp. 237–247, Mar. 1995.

[8] K. Ikeda *et al.*, "Risk factors for tumor recurrence and prognosis after curative resection of hepatocellular carcinoma," *Cancer*, vol. 71, no. 1, pp. 19–25, Jan. 1993.

[9] A. B. Adegbite, M. I. Khan, K. W. E. Paine, and L. K. Tan, "The recurrence of intracranial meningiomas after surgical treatment," *J. Neurosurg.*, vol. 58, no. 1, pp. 51–56, Jan. 1983.

[10] L.-M. Su, B. P. Vagvolgyi, R. Agarwal, C. E. Reiley, R. H. Taylor, and G. D. Hager, "Augmented Reality During Robot-assisted Laparoscopic Partial Nephrectomy: Toward Real-Time 3D-CT to Stereoscopic Video Registration," *Urology*, vol. 73, no. 4, pp. 896–900, Apr. 2009.

[11] D. Teber *et al.*, "Augmented Reality: A New Tool To Improve Surgical Accuracy during Laparoscopic Partial Nephrectomy? Preliminary In Vitro and In Vivo Results," *Eur. Urol.*, vol. 56, no. 2, pp. 332–338, Aug. 2009.

[12] Y. Sato *et al.*, "Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization," *IEEE Trans. Med. Imaging*, vol. 17, no. 5, pp. 681–693, 1998.

[13] H. Liao, T. Inomata, I. Sakuma, and T. Dohi, "Three-dimensional augmented reality for mriguided surgery using integral videography auto stereoscopic-image overlay," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1476–1486, 2010.

[14] H. Fuchs *et al.*, “Augmented reality visualization for laparoscopic surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1998, pp. 934–943.

[15] A. Elmi-Terander *et al.*, “Surgical Navigation Technology Based on Augmented Reality and Integrated 3D Intraoperative Imaging: A Spine Cadaveric Feasibility and Accuracy Study,” *Spine*, vol. 41, no. 21, pp. E1303–E1311, Nov. 2016.

[16] R. Wen, C. B. Chng, C. K. Chui, K. B. Lim, S. H. Ong, and S. K. Y. Chang, “Robot-assisted RF ablation with interactive planning and mixed reality guidance,” in *2012 IEEE/SICE International Symposium on System Integration (SII)*, 2012, pp. 31–36.

[17] G. Lacey, D. Ryan, D. Cassidy, and D. Young, “Mixed-Reality Simulation of Minimally Invasive Surgeries,” *IEEE Multimed.*, vol. 14, no. 4, pp. 76–87, Oct. 2007.

[18] L. Soler *et al.*, “Virtual reality and augmented reality in digestive surgery,” in *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2004, pp. 278–279.

[19] V. Ferrari*, G. Megali, E. Troia, A. Pietrabissa, and F. Mosca, “A 3-D Mixed-Reality System for Stereoscopic Visualization of Medical Dataset,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 11, pp. 2627–2633, Nov. 2009.

[20] P. Mildenberger, M. Eichelberg, and E. Martin, “Introduction to the DICOM standard,” *Eur. Radiol.*, vol. 12, no. 4, pp. 920–927, 2002.

[21] W. D. Bidgood, S. C. Horii, F. W. Prior, and D. E. Van Syckle, “Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging,” *J. Am. Med. Inform. Assoc.*, vol. 4, no. 3, pp. 199–212, 1997.

[22] D. J. Brenner and E. J. Hall, “Computed tomography—an increasing source of radiation exposure,” *N. Engl. J. Med.*, vol. 357, no. 22, pp. 2277–2284, 2007.

[23] U. Schneider, E. Pedroni, and A. Lomax, “The calibration of CT Hounsfield units for radiotherapy treatment planning,” *Phys. Med. Biol.*, vol. 41, no. 1, p. 111, 1996.

[24] M. Soret, S. L. Bacharach, and I. Buvat, “Partial-Volume Effect in PET Tumor Imaging,” *J. Nucl. Med.*, vol. 48, no. 6, pp. 932–945, Jun. 2007.

[25] R. Holloway and A. Lastra, “Virtual environments: A survey of the technology,” 1993.

[26] J. Steuer, “Defining Virtual Reality: Dimensions Determining Telepresence,” *J. Commun.*, vol. 42, no. 4, pp. 73–93, Dec. 1992.

[27] R. T. Azuma, “A Survey of Augmented Reality,” *Presence Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 355–385, Aug. 1997.

[28] P. Milgram and F. Kishino, “A taxonomy of mixed reality visual displays,” *IEICE Trans. Inf. Syst.*, vol. 77, no. 12, pp. 1321–1329, 1994.

[29] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt, “3D shape scanning with a time-of-flight camera,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 1173–1180.

[30] *HoloToolkit-Unity: This is effectively part of the existing HoloToolkit, but is the repo that will contain all Unity specific components*. Microsoft, 2017.

[31] W. E. Lorensen and H. E. Cline, “Marching Cubes: A High Resolution 3D Surface Construction Algorithm,” in *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 1987, pp. 163–169.

[32] “Asset Store.” [Online]. Available: https://www.assetstore.unity3d.com/en/#!/content/43658. [Accessed: 06-Jun-2017].

[33] “Vuforia | Augmented Reality.” [Online]. Available: https://www.vuforia.com/. [Accessed: 06-Jun-2017].

[34] U. D. of Health, H. Services, and others, “Summary of the HIPAA privacy rule,” *Wash. DC Dep. Health Hum. Serv.*, 2003.

[35] J. R. Shewchuk, “Delaunay refinement algorithms for triangular mesh generation,” *Comput. Geom.*, vol. 22, no. 1–3, pp. 21–74, 2002.