Towards 3D Medical Imagery Navigation and Alignment on a Human Body

Sean Konz
Stanford University
Department of Electrical Engineering
swkonz@stanford.edu

Daniel Henry
Stanford University
Department of Symbolic Systems
dhenry@stanford.edu

Abstract

In this work, we explore visualization strategies for medical imagery using Virtual and Augmented Reality. We utilize the Microsoft Hololens system and a Google Cardboard display system for positional tracking and visualization of real medical models. Our system allows users to explore medical models by walking around the model, and additionally allows users to drag, scale, and rotate models using hand gestures from the Hololens. We also detail methods for extending this work to allow for model alignment with patients in real world scenes. Overall, the hololens platform of visualization provides an interesting medium for 3D model visualization, but challenges stemming from the hololens field of view, hand tracking, and environmental precision are barriers for the advancement of medical use of AR systems.

1. Introduction

Medical imaging technology has enabled physicians to save millions of lives each year. Imagery technology such as Computed Tomography (CT) Scans or Magnetic Resonance Imagery Scans (MRIs) use radiation and magnetic fields in order to capture images of a anatomical subject. These scans capture cross-sectional images at specific depths into the subject, commonly known as tomographic slices. These slices are usually displayed as individual 2D images; however, modern imaging techniques utilize volume imaging in order to capture a 3D model of a subject. These imaging techniques can be thought of as capturing many slices of a subject with minuscule variations in depth (1mm - 5mm) in order to create a high resolution 3D image. Despite the 3D nature of volume imaging scans, the presentation of the resulting 3D models is still largely based on 2D image presentations. Although effective, this forces physicians to think about the depth dimension of these 2D images, often leading to loss of information, and practical errors.

Virtual and Augmented Reality (VR/AR) systems have become standard tools for visualizing 3D models and environments. VR is already being used in architecture and civil engineering to allow engineers to view and traverse through life-size models of buildings and structures. This evolution of the technology is beginning to become available in the medical space, at a much smaller, and more precise level.

Applications of Virtual Reality and Augmented Reality in gaming and architecture have paved the way for 3D medical imagery representation, enabling a significant improvement in how physicians can understand the medical challenges posed by their patients. The use of 3D models allows physicians to better understand the landscape of the body of a specific patient prior to beginning a surgery. Although most human bodies have a similar structure, slight variations in artery and organ placements can cause difficult problems for physicians while conducting surgical procedures. The 2D presentation of imaging models makes inferring the separation between body parts difficult across multiple imagery slices. Using a 3D representation of these models through VR or AR, physicians are able to see the relative distance between critical body structures which may be occluded in 2D images. Additionally, using VR can allow physicians to view, edit, and traverse medical scans through virtual reality. Through 3D representations of current 3D imagery such as MRIs and CT scans, physicians can spot problems before an operation, and can more rigorously explore the 3D models of their patient's bodies in order to make better diagnoses. This additional information can have a dramatic impact on a physician's preparation time, and surgery outcomes.

The visualization of a medical scans overlaid on a patient can also prove useful for during a surgical operation. Overlaying medical imagery on a live surgery can help surgeons gain make more precise movements around anatomical structures. This use case requires models to be represented in AR in order to augment real world objects. In this project, we will focus on supporting this use case by exploring the use of the HoloLens AR platform to visualize and align 3D medical models on real world patients.

2. Related Work

Viewing 3D medical models in AR/VR systems has focused on three major use cases: surgical training, operation planning, and surgeon augmentation during operations. The development of VR systems as a whole has driven the potential for VR/AR usage in the medical industry, with usage in architecture drawing a close parallel to medical use cases [14]. The use of VR/AR in architecture has begun to demonstrate the potential for 3D modeling, but many of the models in architecture do not require precise distances scales or alignment, since they are meant for viewing, rather than editing, manipulating, or surgical planning. In the medical field, 3D modeling has struggled due to precise measurement representations, anatomical alignment, and imaging difficulties such as persistently dynamic regions of the body, such as the abdomen.

2.1. Surgical Training

Surgical Training tools have focused on developing VR environments for training surgeons in a fully immersive environment. Challenges with surgical training environments center around the difficulty in effectively modeling tissue physics[6]. Additionally, the development of highly realistic haptic systems is critically important for effective training processes in order to properly replicate real world environments [10]. Modern AR/VR systems continue to rely on primitive geometries to represent real world objects, which often does not allow for the precision and realism required for medical imagery representation. These challenges are well defined, and tremendous advances have been made, as is evident by realistic simulators and high accuracy haptics for surgery as implemented in the popular Da Vinci remote surgery system [8].

2.2. Operation Planning

Visualizations for use in operation planning have been a large focus of development in order to utilize imagery navigation and manipulation in a 3D environment which can help prevent surgical complications through enhanced imagery inspection. Many of these tools are already in use in labs which have begun development on such systems. Soler et al developed tools for imaging the abdomen in order to better help surgeons prepare for surgery, they are currently under clinical testing [9]. Additionally, problems with 3D modeling of medical imagery have stemmed from a need to focus on the segmentation of body parts within a model [5]. This is of importance since removal of certain structures and components such as bone, nerves, and tissue, within an anatomical model can better allow physicians to view the structure of the patient's body. During surgical planning processes, the assessment of various imagery sources and their interactions helps to better spot sources of complications. This makes the alignment of numerous 3D models into a single scene desirable. Frohlich et al accomplished this using a GPGPU ray casting implementation, but generally focused on visualization, and not on accurate measurement representation [7]. Many of the model manipulation challenges faced by surgical planning use cases are shared across many use cases for VR/AR in virtual reality. The continued development of these use cases will only serve to benefit all adopters of medical VR/AR systems.

2.3. Intra-Operation Augmentations

Usage of AR systems for surgical operation augmentation still offers a large opportunity for further development. Most of these initiatives have struggled due to the need to align models to a constant reference location on the patient's body [15], but the means and formatting of imagery visualization in the operating room is still a topic of research. Von der Heide et al focus on using imagery captured during a surgery in order to aid actions taken by the surgeon [13]. The dominant usage of AR in the operating room has focused on surgical guidance, the act of understanding the surgical task at hand, and guiding a surgeon through the process by combining real world data and imagery data. Vassalo et al review surgical guidance systems which merge data in this manor, and introduce a novel technique for blood vessel classification based on data obtained during a surgery [12]. This work demonstrated the strong potential for AR to be used to augment surgeons by providing a method for intra-opertive blood vessel classification without the use of additional scanning techniques or dyes. The use of intraopertive drop-in imaging techniques has become increasingly popular during operations in order to aid surgeons in making precise movements and incisions. These techniques inherently require a means for visualization, which is currently accomplished using 2D screens [11]. Additionally, augmenting these scans with an ability to align models to the anatomical structures can provide a more direct route for guiding surgical incisions. Overall, these works can benefit from advanced visualization techniques, and the use of AR offers the most potential for combining visualization and alignment on humans of medical models for use in the operating room.

3. System Description and Development

In this work, we develop two systems for medical model visualization using AR and VR. Our applications in VR and AR both utilize the Unity game development platform, but we also comment on the use of alternative platforms for visualization through Microsoft DirectX. We additionally explore various options for converting medical scans from DICOM formats to 3D objects for use in visualization systems. Finally, we comment on our the ability for modern

AR platforms to accomplish the task of model alignment on anatomical structures based on our exploration of the sensor systems available on the Microsoft Hololens.

3.1. DICOM format and Conversion to 3D

The standard medical imagery format is the Digital Imaging and Communications in Medicine (DICOM) format and is used by nearly all medical scanning techniques. The format consolidates the medical information of a patient, and a patient's actual medical scan into a single file format to ensure the patient information is never separated from their medical imagery. For this reason, anonymizing the data is required, which requires scrubbing the personal information from each frame. This was accomplished through the use of the pydicom library [4].

In order to utilize the real medical data within a game engine such as unity, we needed to convert the scan to a 3D model. We accomplished this using 3D Slicer, which is an open source visualization and conversion tool for DICOM imagery [1]. Figure 1 shows the slice representations of an MRI brain scan. Figure 2 shows the progression of the 3D model from the slices, and shows the resulting 3D model after the extraction of the skull structure. Extracting specific components from the same MRI scan was done by selecting regions of interest (ROIs) within a single slice, setting threshold values based on the grayscale pixel values within the ROI, and propagating through the other slices to generate boundaries for the structures within the ROI. In MRI scans, dense bone, water, and air are dark, while in CT Scans the colors are inverted. The slices generated through MRI and CT scans map intensity results to 0-255 range in order to store them as 8 bit grayscale images. Following the generation of the 3D models, it's necessary to re-scale and crop the resulting model file since empty space is still included in the resulting model since the 2D slices are not perfectly white balanced or scaled. These post processing edits make the model more manageable once within the game engine environment. The left and right images in figure 2 are results of this process for the bones and brain tissue within an MRI. The resulting models from this process varied in precision, but had polygon counts of upwards of 40,000.

3.2. Virtual Reality System

Our Virtual Reality scene included a 3D rendering of ribs extracted from a CT scan using the aforementioned process. We used the GoogleVR SDK for generating a VR application within Unity, and allowed for users to view, navigate around, and physically align a 3D object model with a manikin within the scene [2]. The player within the scene consisted of a GoogleVR camera that could be moved around the scene using the arrow keys to get different angles of the model ribs and manikin. We used the custom VRduino platform for IMU information and allowed for viewer movement using a keyboard.

3.3. Augmented Reality System

Our Augmented Reality scene used the extracted brain and skull models displayed in figure 2. The scene was constructed using Unity and the Microsoft HoloToolkit for deployment on the Hololens [3]. The application allows users to view imported 3D models, move them around, scale them, and rotate them using hand gestures. Spatial awareness and spatial anchoring was accomplished using the Holotoolkit APIs. This allowed us to design our scene in unity using where the viewer camera was located at boot-up as the origin, and the model to be viewed was positioned at a point relative to that origin. When the application is loaded into the hololens, the system scanned the surrounding environment, and sets the origin location within the room to the location which the headset is in at the start of the application. This allows the viewer to move around in the room while the coordinate system of the physical and virtual world remain constant. Using this setup, a user is able to walk 360 degrees around the objects in the scene. The Unity app was compiled into a DirectX project, which was then ported to the Hololens. Figure 3 shows a view of our application in use.

3.4. Alternative DirectX Visualization System

Care was taken to explore alternative application development routes for the Hololens, which led to our development of a DirectX application environment. Usage of DirectX for building a graphics environment provided an additional level of freedom to access sensor suites and information from the hololens, however, this system was ultimately not realized due to time constraints.

4. Testing and Results

We qualitatively tested our applications within our team as well as with an actual surgeon and collected comments about the performance of the applications. From figure 2, it's possible to see the low resolution of the resulting 3D models from the MRI data. This is due to model smoothing required in order for the hololens to effectively display the models at a reasonable frame rate, but also due to the fact that minute details which are present in the 2D slice imagery are not effectively converted, nor represented in these 3D models. This is due to shaders and textures which do not

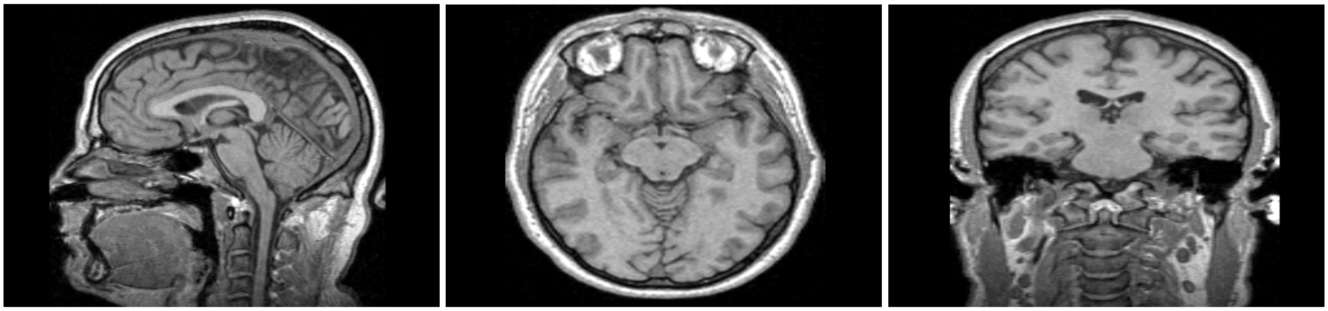

Figure 1: Examples of DICOM slices for a brain MRI representing the Sagittal, Axial, and Coronal representations, respectively. Each representation consisted of 250 slices at a slice depth of 1mm.

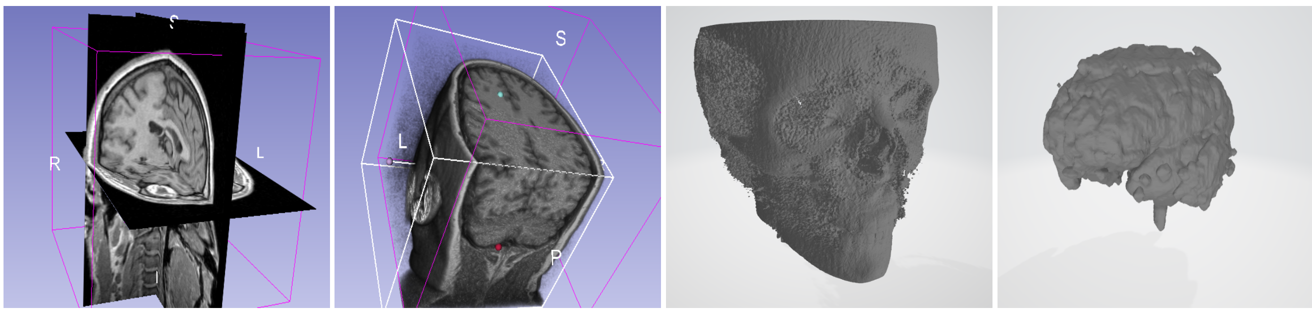

Figure 2: First the slices are aligned about their centers as slices (left). Then they are converted to a volume representation which can be cut into to view each slice in 3D (left-middle). Lastly, we convert the volume representation to a 3D model and extract regions of interest (right-middle, right)

properly render the details of the object. Additionally, our 3D models are generated as shells in order to keep complexity at a minimum. This results in a loss of information for all the contents within each object which can still be visualized within a 2D slice. Furthermore, the threshold approach for generating the 3D models resulted in some regions being included in the output model which were inaccurate since the boundaries within the slices are not always precise.

In addition to challenges around the fidelity of our resulting models, we saw issues around the usability of the control mechanisms in both AR and VR. In VR, our system was dependent on the use of the keyboard and mouse. Although this was intuitive, model manipulation sometimes reacted in inaccurate ways due to inaccurate sensor readings. However, these issues could be rectified using a more advanced VR system that included position tracking and handheld controllers. In our AR system, we were completely tied to the use of hand gestures for all manipulation. The hololens gesture tracking proved too noisy to provide for smooth model manipulations which negatively impacted the overall usability of the system. In addition to the input controls for the hololens, the field of view, spatial awareness problems, and resolution of the display made visualizations on the device less informative, and challenging to use. The spatial

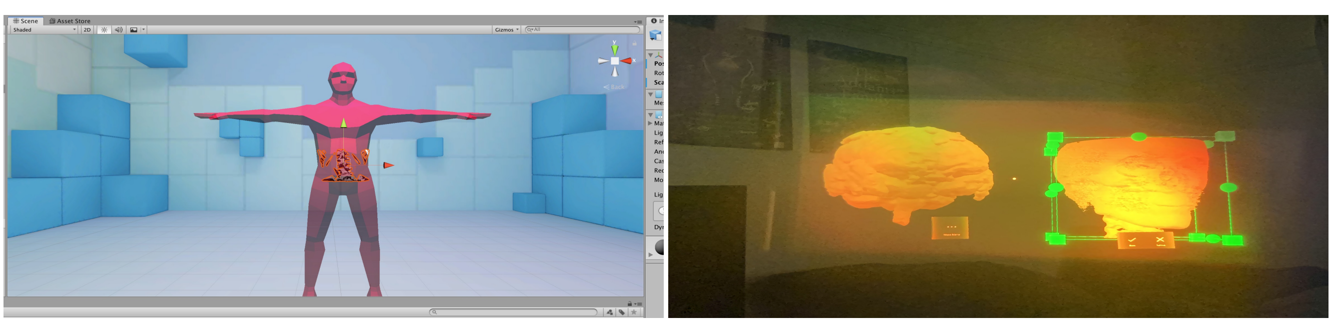

Figure 3: Visualizations from our Virtual Reality System (left), and our hololens application (right)

awareness system of the hololens relies on the detection of planes in the surrounding room in order to place holograms into the world for the user. In smaller, darker rooms, this works well, and the system is able to detect planes, and anchor objects on those planes reliably. However, in larger rooms, the anchoring is less reliable, and some objects will drift within their plane as the user looks around, making it difficult to focus on the object. Despite the shortcomings of our system, we saw success in experimenting with different visualization processes for medical models and in attempting new visualization techniques. Visualization by separating regions allowed to increase a physician's ability to view sub-structures of a scan in 3D, but the imprecision of the resulting model had a negative impact on a physician's ability to extract useful information from the model. Additionally, From our demos with physicians, it's evident that AR and VR systems for medical imagery visualizations need to provide physicians with significant additional information and ease of use in order for their adoption to become worthwhile, and this goal cannot be achieved through simple visualization tasks. However, our experiments around model alignment with humans show greater potential.

5. Discussion and Future work

This project serves as a foundation for the development of more complex AR/VR systems for use in visualizing 3D medical imagery. From our tests with physicians, its clear that VR systems will be challenged to provide additional information to physicians in order to warrant their adoption. In contrast, AR systems show significant potential for use in medical imagery visualization due to the potential for physicians to use these systems during surgery, and while interacting with patients. This exaggerates the importance of advancing the components of this project which were points of particular difficulty. As mentioned above, it's evident that 3D model shells generated from DICOM imagery discard considerable amounts of useful information which physicians need in order to effectively do their jobs. These results are common for both the VR and AR systems, and are more dependent on the methodology used to generate a 3D model from the original DICOM file. As a result, it's necessary to alter the DICOM representation pipeline in order to generate a 3D model of the DICOM data while retaining an ability to visualize individual slices of the model. This can likely be achieved through cropping of the 2D slices to match the anatomical shape being scanned, and then representing each slice individually within a virtual environment, rather than compiling all the slices into a single 3D model for use in a game engine-based application. Additionally, the complexity of these resulting representations must be brought into consideration when the intended use is in AR systems. This motivates the use of alternate graphics environments rather than the use of game engines such as Unity. The use of alternate graphics environments allows for more control over the precision of model measurements within a scene in addition to allowing for optimizations of specific object geometries which may be specific to medical imagery (i.e. slice representations). In order to support the adoption of AR systems in medical imagery visualization, interactions with the real world are critical. AR shows potential for use in the operating room and other real world environments, but the ability for applications on platforms like the Hololens to utilize the sensors available on the device will be critical for the advancement of these use-cases. Currently, access to the device sensors is restrictive and burdensome to work with, however; developments are being made to make this process simpler in the future. The availability of sensor information in custom applications will assist in the ability to align models on real world objects. This use case appears to be highly desirable to physicians, and to be the most reasonable to achieve in the near future.

Finally, the gesture and input mechanisms explored in this project make evident the need for seem-less interactive gestures for AR applications. Gesture detection and reliable hand tracking can significantly improve the user experience with AR applications, and the need for these kinds of reliable interactions is only exaggerated for use-cases in the medical field. Similar to the ability for external developers to utilize device sensors, these features will likely improve with further iterations in AR devices like the Hololens.

6. Conclusion

This project demonstrates a starting point for developing VR and AR applications for visualizing 3D medical data based on MRI and CT scans. We leverage the Hololens platform and the Google Cardboard SDK in order to build virtual environments for visualizing medical imagery data. Additionally, we detail a process for generating 3D models from MRI and CT scan DICOM data. We discuss the shortcomings of our approach, which most commonly lie in our model generation and representation approach, and suggest numerous routes for improvement and additional development. In general, we found that the use of common game-engines lack the freedom and fidelity required in order to create highly precise environments for medical data representation, and we suggest the use of more custom environments through DirectX in the future. We view this project and our approaches as a step towards development of mixed reality medical imagery visualization, and our tests with physicians suggest growing acceptance for these kinds of technologies.

7. Acknowledgements

We'd like to thank Medweb Inc for providing us access to collections of DICOM image data. Additionally, we'd like to thank Dr. Peter Killcommons for testing and providing feedback on our work.

References

[1] 3D Slicer 3d slicer is an open source software platform for medical image informatics, image processing, and three-dimensional visualization. `https://www.slicer.org/`. Accessed: 2019-06-7.

[2] Google Cardboard SDK google vr provides sdks for many popular vr development environments. `https://developers.google.com/vr/`. Accessed: 2019-06-7.

[3] Microsoft Mixed Reality Toolkit mrtk-unity provides a set of foundational components and features to accelerate mr app development in unity. `https://microsoft.github.io/MixedRealityToolkit-Unity/README.html`. Accessed: 2019-06-7.

[4] PyDICOM library and information. `https://pydicom.github.io/`. Accessed: 2019-06-7.

[5] S. González Izard, J. A. Juanes Méndez, P. Ruisoto Palomera, and F. J. García-Peñalvo. Applications of virtual and augmented reality in biomedical imaging. *Journal of Medical Systems*, 43(4):102, Mar 2019.

[6] U. Khnapfel, H. akmak, and H. Maa. Endoscopic surgery training using virtual reality and deformable tissue simulation. *Computers Graphics*, 24(5):671 – 682, 2000. Dynamic Medical Visualization (special topic).

[7] A. D. A. W. N. C. P. H. C. S. V. B. K. Magali Frohlich, Christophe Bolinhas and S. Gobron. Applications of virtual and augmented reality in biomedical imaging. *Journal of Medical Systems*, 43(4):102, Mar 2019.

[8] C. Perrenot, M. Perez, N. Tran, J.-P. Jehl, J. Felblinger, L. Bresler, and J. Hubert. The virtual reality simulator dv-trainer is a valid assessment tool for robotic surgical skills. *Surgical Endoscopy*, 26(9):2587–2593, Sep 2012.

[9] L. Soler, S. Nicolau, J. Schmid, C. Koehl, J. Marescaux, X. Pennec, and N. Ayache. Virtual reality and augmented reality in digestive surgery. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 278–279, Nov 2004.

[10] O. A. J. van der Meijden and M. P. Schijven. The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review. *Surgical Endoscopy*, 23(6):1180–1190, Jun 2009.

[11] M. N. van Oosterom, H. G. van der Poel, N. Navab, C. J. van de Velde, and F. W. van Leeuwen. Computer-assisted surgery: virtual-and augmented-reality displays for navigation during urological interventions. *Current opinion in urology*, 28(2):205–213, 2018.

[12] R. Vassallo, H. Kasuya, B. W. Y. Lo, T. Peters, and Y. Xiao. Augmented reality guidance in cerebrovascular surgery using microscopic video enhancement. *Healthcare Technology Letters*, 5(5):158–161, 2018.

[13] A. M. von der Heide, P. Fallavollita, L. Wang, P. Sandner, N. Navab, S. Widert, and E. Euler. Cameraaugmented mobile carm (camc): A feasibility study of augmented reality imaging in the operating room. *Medical Robotics and Computer Assisted Surgery*, 14(2):2587–2593, Dec 2017.

[14] J. Whyte, N. Bouchlaghem, A. Thorpe, and R. McCaffer. From cad to virtual reality: modelling approaches, data exchange and interactive 3d building design tools. *Automation in Construction*, 10(1):43 – 55, 2000.

[15] M.-L. Wu, J.-C. Chien, C.-T. Wu, and J.-D. Lee. An augmented reality system using improved-iterative closest point algorithm for on-patient medical image visualization. *Sensors*, 18(8), 2018.