



iPhone Virtual Reality Training Model for Microsurgical Practice

BME 400

University of Wisconsin-Madison
Department of Biomedical Engineering
12/14/2022

Team Members:

Samuel Neuman - *Team Leader & BWIG & Communicator*
Haochen Wang - *BSAC & Co-BWIG*

Clients:

Dr. Ellen Shaffrey
Dr. Samuel Poore
Sahand Eftekari

Advisor:

Prof. Walter Block

Abstract

Current microsurgery training methods with surgical microscopes have seen their limitations. First, the microscopes are expensive and not portable, thus there is limited access to the devices and undermines the learning outcome for the trainees. Second, beginning during the COVID-19 pandemics, there is a rising need for remote surgical instruction. As smartphone cameras can provide sufficient resolution and magnification on the surgery station, this project aims to develop a microsurgery training solution based on common smartphone models that is capable of forming two visual channels to enable depth perception from the display. The team proposed a final design that reflects one visual channel but not the other and a corresponding code that can incorporate live-streamed video into a computer model of the virtual reality environment. Discrepancies between the two visual channels and the real-time video lag were quantified. This work facilitated the development of a timeline for the completion of this project and steps that need to be taken thereof. Future work includes implementation of our current code and image transformation algorithms into the VR environment to create a usable image for the training surgeon.

Table of Contents

Abstract	1
Table of Contents	2
1. Introduction	4
1.1 Motivation	4
1.2 Problem statement	4
2. Background	4
2.1 Surgical Microscopes	4
2.2 Smartphone Cameras	5
2.4 Clients	6
2.5 Product Design Specifications	6
2.6 Competing Designs	6
3. Preliminary Designs	8
3.1 Single Frame Design Concepts	8
3.1.1 Basis of Image Transformation	8
3.1.2 Single Frame Preliminary Design	10
3.2 Combined Frame Preliminary Design	11
4. Preliminary Design Evaluation	13
4.1 Design Matrix	13
4.2 Design Consideration	13
4.2.1 Effectiveness (Latency)	13
4.2.2 Sensitivity (Depth Perception)	14
4.2.3 Ease of Use	14
4.2.4 Cost	14
4.2.5 Compatibility	14
4.2.6 Frame Rate	15
4.2.7 Durability	15
5. Final Design	15
5.1 Timeline of the Final Design	15
5.2 Mirror Array	16
6. Fabrication/Development Process	19
6.1 Overview of the Development Plan	19
6.2 Video Acquisition	19
6.3 Video Processing	21
6.4 Testing	21
6.4.1 Testing Methodologies	21

6.4.2 Results	22
6.4.3 Calibration Parameters	23
6.4.4 Preliminary work on VR Application	24
7. Discussion	25
7.1 Future Work	26
Conclusions	26
References	27
Appendix	29
A. Product Design Specification	29
a. Performance requirements:	29
2. Production Characteristics	32
3. Miscellaneous	32
B. CameraScript.cs used in Image Acquisition Phase	33
C. CameraControl.cs to be used in Video Processing	35
D. CameraControlRight.cs to be used in Video Processing	37
E. Python code with OpenCV	37
F. CAD Drawing for the Smartphone Attachment	38

1. Introduction

1.1 Motivation

Hands-on experience with microscopes is an important aspect of microsurgery training, which not only enhances the learning outcome of medical education but also improves the surgical results of patients of the trainees in the future. Due to the expensiveness and immobility of surgical microscopes, opportunities for microsurgery practice are limited. Thus, there is a commercial gap to satisfy the need for a more portable and affordable alternative to surgical microscopes that allow for practice outside of the operating room. Cameras on smartphones can now provide sufficient magnification and high resolution for live streaming of videos and have therefore been proposed to fulfill this commercial gap. However, as microsurgery requires binocular vision to enable depth perception of the microsurgeon, most smartphone models can only activate one camera lens at a time, and therefore smartphone cameras alone are not feasible for microsurgery training. By combining a hardware design of smartphone attachment with a software design to process the video, the deliverable of this project will allow microsurgery trainees to practice at home via their smartphones while enabling live streaming of their practice for simultaneous instruction from the instructions.

1.2 Problem statement

Due to high cost and importability of surgical microscopes, there is limited access to the microscopes and thus restricted training opportunities for trainees. This limitation is more significant in less developed regions. This project aims to deliver a smartphone-based solution that displays the microsurgery station with depth perception, with minimal latency in video stream, with an affordable price and with high portability. Additionally, the design should not compromise the awareness of the user to the surrounding environment when performing microsurgery training. Ideally, the video can be simultaneously shared with instructors.

2. Background

2.1 Surgical Microscopes

Surgical microscopes are used to magnify the microsurgery station and the tissue of operation, allowing more precise operation on the patients for optimal treatment outcomes. Binocular microscopes first saw their appearance in operations rooms in 1922 [1]. Since then, advancements in optics fabrication enable microscopes to have high optic precision for less image distortion, high-power coaxial illumination for adjustable magnification levels, and optimal working distance that prevent obstruction on the operation [2].

As shown in Figure 1, the clients are currently using a model from Zeiss company that costs approximately \$300,000. The model has maximum 30x total magnification with two pairs of eyepieces, enabling practice with two surgeons at the same time. Meanwhile, exoscopes are more commonly used in contemporary operation rooms. Instead of looking through a pair of

eyepieces, surgeons can now view directly from a monitor via a pair of specially designed goggles. The Orbeye series from Olympus is an example of such exoscopes, which has 26x magnification with 4k display capability. The typical weight of an Orbeye exoscope is 20 kilograms, and the cost varies between \$200,000 to \$1 million [3, 4].



Figure 1. The Zeiss microscope that is currently used by clients in their lab (device in the middle with an extending arm). This demonstrates the importability of the microscope.

2.2 Smartphone Cameras

Phone camera quality has been improved to the point where they can be compared to the quality of standalone cameras. This, along with the inherent capability of processing/transferring data and the ability to interface with various types of software, makes smartphone cameras an ideal platform to build off of. While the quality is sufficient for the intended use, the lack of depth perception is one factor that leads to the biggest problem in applying the smart phone as a solution.

2.3 Preliminary Considerations in Stereoscopic Display Technology

Several methods of creating and viewing a stereoscopic display are currently employed. One of the most common methods of doing so is through the use of anaglyph filter. This consists of overlaying two chromatically different masks over an image. These filters are slightly offset to create two layered images. These colored masks are then perceived by an observer through the use of anaglyph glasses. The observer gets a sense of depth through the disparity between the images. The anaglyph image can be displayed on a normal monitor and allows for the observer

to be aware of their surroundings [5]. A second potential display platform is entry-level wearable virtual reality devices that process real time video via programming. Each eyepiece on the VR device can receive different visual inputs, thus effectively forming depth perception when the user views the real time video transmitted from the smartphone.

2.4 Clients

Our clients this semester are Dr. Ellen Shaffrey, Dr. Samuel Poore and Sahand Eftekari from the Division of Plastic Surgery at UW Madison.

2.5 Product Design Specifications

The first major specification from the client was in regards to depth perception. Our design must be able to give the user enough information to navigate their hands and medical instruments in 3D space. Next, the quality must be good enough so that the user is able to see and manipulate sutures that are 0.07 mm in diameter [6]. Since this design is intended to be an affordable training tool, the final design must remain cost effective and accessible for the end user. Lastly, it is important for the design to provide a low latency feed with a maximum delay of 0.5 seconds and minimal 30 fps in display. This ensures that the experience of using the design has a more natural and seamless feel.

2.6 Competing Designs

Currently there are some alternatives to traditional surgical microscopes. The Olympus Orbeye is a highly portable advanced surgical microscope. It offers the flexibility of displaying to a large, low latency 4k display. This maximizes the users ability to be aware of their surroundings and give others the opportunity to observe the procedure. Another advantage of this alternative is the ergonomics. The system of an external display being the main method of viewing means that the user can remain in a more natural position. This can be important for preventing fatigue in long surgical procedures.

While this alternative does offer a variety of beneficial features, the main concern is the cost. The Orbeye, shown in Figure 2, is similar in price to a traditional surgical microscope which makes it a viable competitor in that price range [7]. Since decreasing the cost of the design is one of the main concerns, this would not work as an option for the client. Another factor that makes the Orbeye a non-ideal alternative is the mobility. It is an upgrade when compared to current surgical microscopes, but it still is a relatively large design [8].



Figure 2. The Olympus Orbeye 4k Video Microscope

The next alternative design that was found was the Microsoft HoloLens 2, shown in Figure 3. This is a mixed reality system that can be repurposed for many use cases. It is also certified to be used in sterile environments and has a lot of development support. The high portability, intuitive use, and relatively low price point of \$4950 make the HoloLens 2 a viable alternative.

The main drawbacks of the HoloLens 2 are the lack of developed software for surgical use and the difficulty in integrating it into a teaching setting. The HoloLens 2 does not have a high enough resolution camera to produce a quality image for the user to use as a reference. Its ability to produce a mixed reality 3D space could help solve this issue but could result in unwanted latency issues. Its single user focused design prevents it from being used as an effective teaching tool as other observers cannot view the feed. The 3 hour battery life can also limit its usability [9].

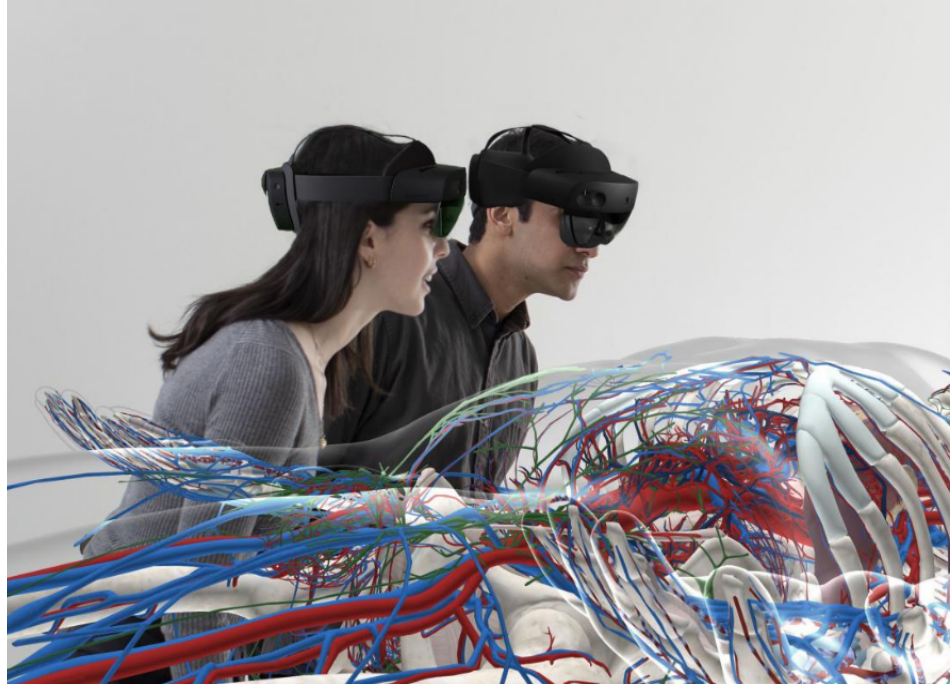


Figure 3. The Microsoft HoloLens 2.

3. Preliminary Designs

3.1 Single Frame Design Concepts

3.1.1 Basis of Image Transformation

The single frame view design aims to present the user with a single image at a time containing all information captured by the camera. First, the camera images the object through the mirror array. This splits the acquired image into two halves, with each having inclined above the flat reference plane (for instance, the tabletop). The information gained from this split-view image is not useful to the user, so certain manipulations need to be made to conjoin the images to create a meaningful picture. First, a linear transformation needs to be applied to the images to compress them in the horizontal plane and present a flattened image to the user. A schematic of the geometry of this can be seen in Figure 4.

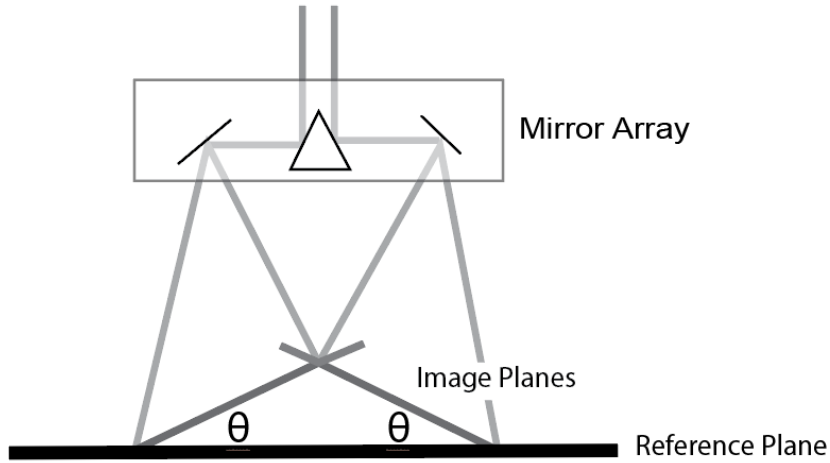


Figure 4. The geometry of the inclined image planes created by the mirror array. The offset of the mirrors in the mirror array will be set such that the camera focuses only at a fixed distance from the mirror array. With this known distance, the angle θ can be derived as the angle between the two image planes and the reference plane. A trigonometric relationship exists between the image planes and the reference plane that will govern the amount of horizontal compression needed to create a meaningful image to the user.

Once the angle θ between the image planes and reference planes is calculated, a mathematical representation between the image planes and the reference plane can be defined. One simple way to do this is by placing the x and y coordinates into a 2x2 matrix and applying the transformations shown in Eq. 1.

$$\{\mathbf{A} \in \mathbb{R}\} = \begin{bmatrix} \cos(\theta) & 0 \\ 0 & 1 \end{bmatrix}$$

(Eq. 1)

Since matrix A has non-zero determinant and occupies the set of real numbers, no voxels of the transformation can be lost by this manipulation. Thus, the only loss of resolution will be the result of exceeding the resolution limit of the computer screen or computing power of the device operated by the user. The transformed matrix is shown in Eq. 2.

$$\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \times \mathbf{A} = \begin{bmatrix} x\cos(\theta) & 0 \\ 0 & y \end{bmatrix}$$

(Eq. 2)

As a proof of principle, two images of the same object were taken using a phone camera. In matlab, they were treated with the linear transformation to flatten them onto the same plane. The two images were cropped and stitched using grayscale averaging to maximize the overlap. The final image is shown in Figure 5. The full application of this code took roughly 30 seconds to process the two 1080p images, suggesting possible bottlenecks in future working codes.



Figure 5. *A proof of principle of the linear transformation. Both images were taken at roughly 45 degrees from the normal. Disparities between the images may be due to inaccuracies in the actual angle of imaging and alignment of the camera with the objects of interest.*

3.1.2 Single Frame Preliminary Design

The Single Frame View preliminary design operates by presenting the user with the left and right contributions of the final image simultaneously. Thus, this design is allowed to operate at our lower bound of frame rate of 60 fps. To capture the three-dimensional form of the image being captured, three manipulations will need to be performed. First, the linear transformation will be applied to both the left and right images to compress them into normal-looking images. Next, the images will be cropped such that an equal area of each image is removed to create the final image; cropping will occur at an equal distance from the ‘center’ of the left and right images. Finally, an anaglyph filter will be applied to the images. This will create the perception

of three dimensions in the images to allow for ascertainment of tissue macrostructure. The concept of the single frame view is given in Figure 6.

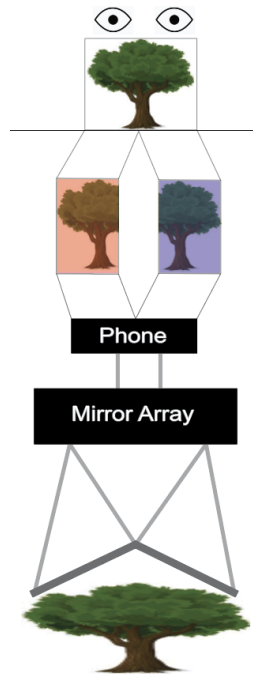


Figure 6. The schematic of the single frame view. First, the camera acquires the images of the two inclined planes of the object. The phone first applies the linear transformation to flatten the image and crops out the overlapping regions of the two halves of the image. Finally, the phone applies an anaglyph filter (red on the leftmost shadows of the objects, blue on the rightmost shadows of the objects) to give the perception of three dimensions to the user. The left and right anaglyph contributions are overlaid onto a single image that is presented to the user.

A major concern for this design, as well as the second design, is the feasibility of computationally expensive calculations to be made in real-time to resolve meaningful images to the user. Crude proof-of-concept code in MATLAB showed that two individual 1080p greyscale images can undergo linear transformation and cropping in approximately 30 seconds. While steps can certainly be taken to streamline these computations, these results show that extra attention should be paid to the Python functions used in the final code such that the data structures allow for the fastest runtime and thus least amount of bottleneck and skipped frames.

Additionally, this design will require the user to wear anaglyph glasses. The implementation of the anaglyph filter and glasses will result in a loss of some perceived color from the image presented to the user. This is likely to be a hindrance to translation, and steps need to be taken to minimize this.

3.2 Combined Frame Preliminary Design

The combined frame, like the single frame design, involves the simultaneous capture of both hemispheres through the mirror array. The two hemispheres will be cropped such that each image contains the same landmarks and no distal parts of the images, but the two hemispheres will not be stitched together. Instead, they will be presented in an alternative fashion i. e. the left image will be presented quickly while the right image is not presented, then the right image will be presented without any semblance of the left image. Ideally, the images will be nearly identical with the exception of the lack of the linear transformation to flatten the images. The goal of this method is that, subliminally, the brain will fail to resolve differences between the images and will process each image as the contributions from each eye independently and then combine information from the two hemispheres to create the perception of three dimensions. In order for this to occur, a frame rate of 120 fps, or double the frame rate of the single view design, would need to be achieved. A schematic of this design is shown in Figure 7.

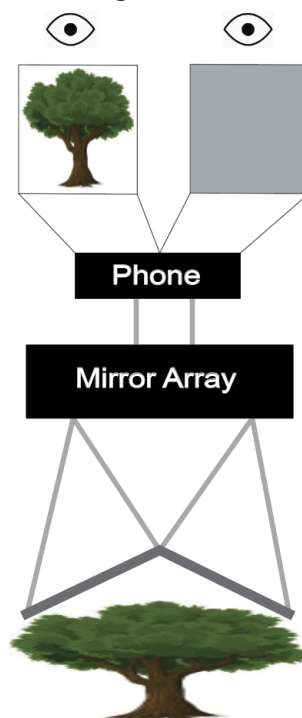


Figure 7. *The combined view preliminary design. Both hemispheres of the images will be acquired simultaneously and cropped by the software to contain the same landmarks and thus be the same shape and encompass the same features. The phone screen will display only one of the left or right images to both eyes at one time in rapid succession, creating the illusion of each eye contributing its perspective to resolve a three dimensional image.*

Even more than the single frame view design, hardware capabilities will be a significant constraint on this design. Operating at a frame rate of 120 fps is optimistic, albeit in excess of the specifications of most computer monitors and phone screens. Otherwise, no accelerations of the computations will be needed, and this design concept is less computationally expensive due to the lack of the linear transformation of the images.

4. Preliminary Design Evaluation

4.1 Design Matrix

The design matrix is shown in Table 1. This matrix helped the team evaluate the two preliminary designs; this serves the additional purpose of helping the team identify aspects of each design that succeed in certain criteria set forth by the PDS when considering the final design.

Table 1. The design matrix of the combined frame view and single frame view designs. Each design was given a score between one and five for each of seven categories. The seven categories were weighted to sum to 100. Cells highlighted in yellow received the higher score for that category between the two designs, and the green cells indicate the winning design concept of this matrix.

Criteria	Weight	Combined Frame View		Single Frame View	
		Raw Score	Score	Raw Score	Score
Effectiveness (Latency)	25	3/5	15	5/5	25
Sensitivity (Depth Perception)	20	4/5	16	5/5	20
Ease of Use	20	5/5	20	4/5	16
Cost	15	3/5	9	2/5	6
Compatibility	10	4/5	8	2/5	4
Frame Rate	5	4/5	4	2/5	2
Durability	5	4/5	4	5/5	5
Total	100	27/35	76	25/35	78

4.2 Design Consideration

4.2.1 Effectiveness (Latency)

Effectiveness was defined as the extent to which the images presented to the user replicate the images produced by the surgical microscope. This includes the perception of three dimensions produced by the images, the minimization of bottleneck and frame skips based on the theory of image processing, and the minimization of image compression as a result of computational constraints. The single frame view won in this category because the combined frame view exceeded the hard limit of screen refresh rate found in commercial computer and phone screens; in order to achieve 120 fps, more advanced screen components would be needed. The single frame view received a score of 5 in this category compared to a score of 3 for the combined frame view design. More generally, the single frame view holds a negligible disadvantage over the combined frame view due to its use of the linear transformation, and this is not expected to impact the user experience.

4.2.2 Sensitivity (Depth Perception)

Sensitivity was defined as the degree to which the user can subjectively perceive three dimensional macrostructures of the object being imaged. The single frame view excelled in this category due to its use of the anaglyph filter and was awarded a score of 5. It is important to note that the anaglyph filter concept has uncategorized drawbacks, such as the loss of color resolution and the necessity of anaglyph glasses to be worn by the user. We predict that the alternative frames of the combined frame view design would be incoherent to the user as the ‘retinal disparity’ or angle of the image planes would be too large and thus the two images would be too disparate. The combined frame view received a score of 4 in this category. By holding the image to one single frame incorporating the full context of the image, we believed the single frame view concept would provide the greatest three dimensional sensitivity to the user.

4.2.3 Ease of Use

Ease of Use relates to the amount of time we estimate it would take to write the code for each design. This estimation was made purely on a theoretical basis. The combined frame view design received a score of 5 due to its lack of a linear transformation, while the single frame view design received a score of 4 as it included the linear transformation and the application of the anaglyph filter before the merging of the two images.

4.2.4 Cost

Costs associated with both devices did not include the costs of 3D printed components but the cost of hardware associated with the upscaling of both designs. The combined frame view design received a score of 3 in this category because its relatively streamlined computational framework is unlikely to require advanced processing capabilities when applied to a larger scale, though it would require a higher-end screen to display the alternative images at 120 fps (monitors are generally cheaper than processors). The single frame view design received a score of 2 in this category because its less streamlined computational framework would require upscaling in terms of computing power, which is generally more costly than improvements in screen technology.

4.2.5 Compatibility

Compatibility was defined as the extent to which experience provided by each design replicates the surgical setting. The combined frame view does not require anaglyph glasses and thus allows the user more comfort and the ability to gain more sensory context from the environment, such as looking down onto the object without obstruction by a pair of glasses. The combined frame view design received a score of 4 in this category. The single frame view did, however, require anaglyph glasses, earning it a score of 2 in this category.

4.2.6 Frame Rate

Frame rate was evaluated in the hypothetical context in which no hardware limitations exist. Practically, when controlling for computation time, the combined frame view would be able to display twice the frame rate as the single frame view design with the same number of computations. Thus, the combined frame view design earned a score of 4 in this category, while the single frame view earned a score of 5 in this category.

4.2.7 Durability

Durability was defined as the hardware strain that the final codes of each design were theorized to implement. To maximize the resolution and framerate of the captured images, governors would have to be coded to prevent frame lag and bottlenecks. Because it was theorized that the single frame view design would be more computationally expensive, more governors would be needed for this design compared to the combined frame view design. Thus, more safety nets would exist in the code for the single frame view design than the combined frame view design, and thus the single frame view design is expected to be safer for the computer hardware when fun at full speed; this paradox persists in many aspects of computer science. The single frame view design received a score of 5, and the combined frame view received a score of 4 in this category.

5. Final Design

5.1 Timeline of the Final Design

This semester's work resulted in both significant progress in the development of our final design concept as well as realization of a realistic timeline for completion of this project. Our timeline can be divided into two main phases. Phase 1, completed this semester, aimed to develop a processing pipeline and apply the work to present an image taken by a smartphone onto a computer screen. To do this, we used a mirror array to create a reflection of the image in one hemisphere. The smartphone then records and livestreams a video to the computer over a tethered connection. The basic algorithm executed by the computer in phase 1 involves first splicing of the left and right visual channels. Calibration parameters are then applied to the spliced images, and a linear transformation is then applied to the images, as defined during calibration, to create two coherent hemispheres of the image. The two channels are then displayed on the computer screen for the user to view. An outline of phase 1 is shown in Figure 8.

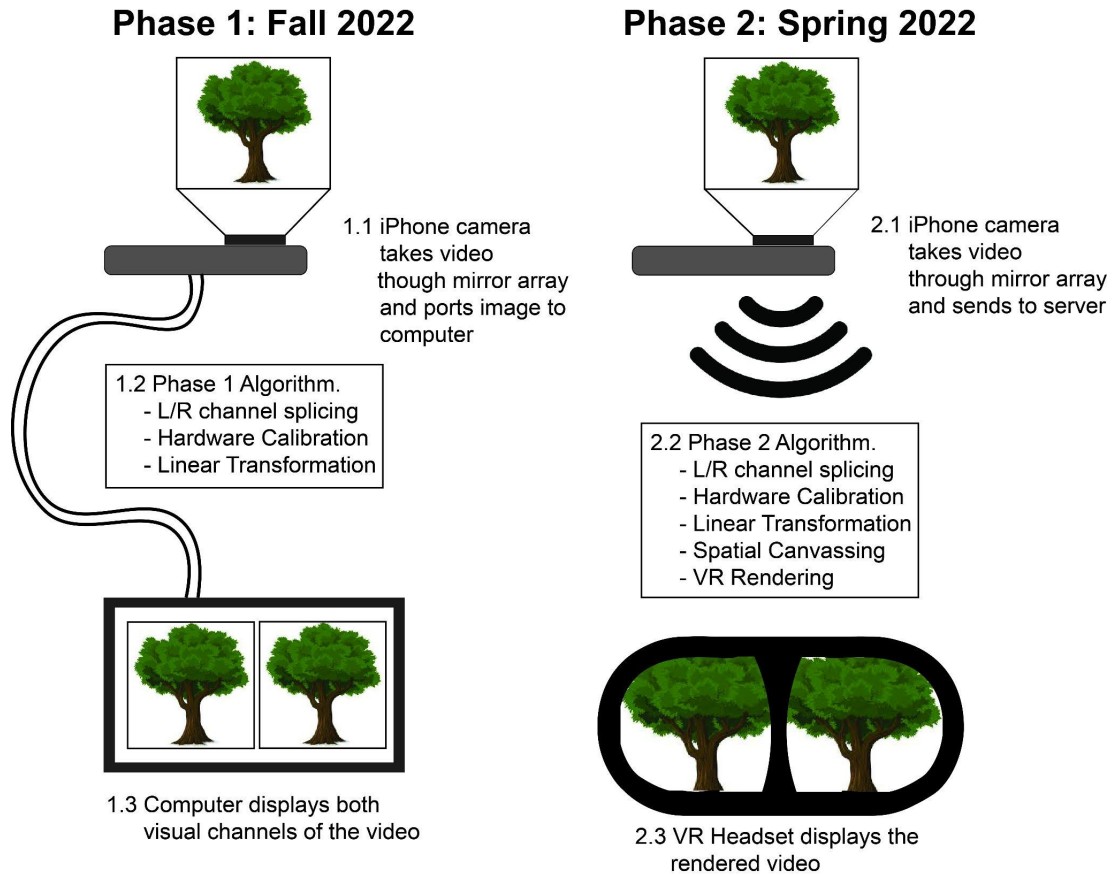


Figure 8. Schematic of the two phases of the project. Phase 1 involves the development of an algorithm for image transformation and execution of the transformations by a computer. Phase 2 involves the application of this concept to a virtual environment.

Phase 2 of the project expands upon the achievements of phase 1; instead of using a tethered connection to stream videos from the smartphone to the computer, a wireless connection will be used to connect the smartphone to a virtual reality headset. Additionally, due to the nature of the virtual environment, additional steps will need to be added to the transformation algorithm to account for changes to the shapes of the images as they are warped. Phase 2 will also require the development of code to allow the virtual reality headset to receive live streamed images from the smartphone.

5.2 Mirror Array

In order to capture two channels of the same object, a mirror array was developed that can reflect one of the channels with an angle of disparity θ . The existence of this angle means that the two channels acquired in the two hemispheres of the image will not be identical--discrepancies in two images will be used by the user to achieve the perception of three dimensions. We opted to use configuration in which one of the visual channels is not reflected and instead passes through the mirror array onto the object of interest. The second visual channel, however, is first reflected by a mirror in a direction perpendicular to the focal ray and

then reflected by a second mirror downward towards the object of interest. The two mirrors were spaced such that the distance between the origin of the focal rays in the real image visual channel and virtual image visual channel are spaced 60 mm apart, or the distance between adult pupils. The angle of disparity between the two focal rays and between the reference plane on which the object of interest would be placed and the inclined image plane of the reflected focal ray visual channel necessitates a transformation to one of the hemispheres to create coherent, similar images in both hemispheres. A schematic of the ray tracing diagram is given in Figure 9 below.

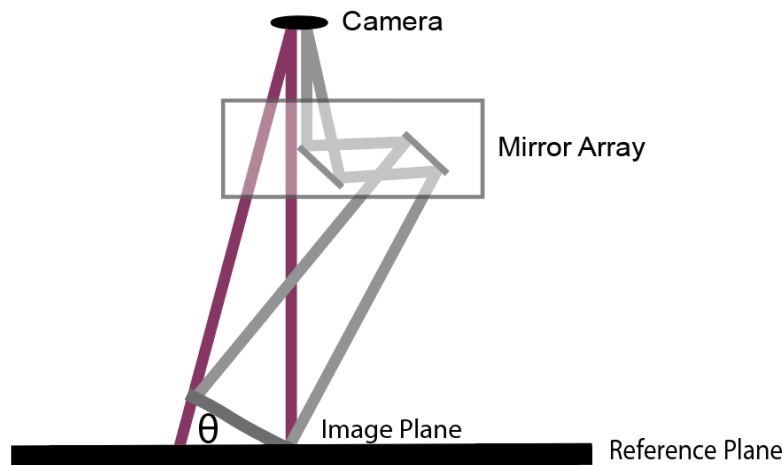


Figure 9. The ray tracing diagram of the final design concept of the mirror array. One visual channel, shown in red, is not reflected and passes through the mirror array to image the object of interest directly. The other visual channel, shown in gray, is reflected orthogonally to a distance of 60mm and then reflected downward towards the object of interest. The image plane of the reflected visual channel is inclined at an angle θ to the reference plane on which the object of interest rests.

Due to uncertainty in the resolution of available 3D printers, our team opted for a design in which the mirrors could be adjusted to calibrate the device and thus account for changes in focal distance. This was achieved through the use of pins on which the mirrors could be fastened, and the pins could be turned using a screwdriver for adjustment. The mirror array also includes holes for a velcro strap to be inserted for securing the mirror array to the smartphone. An image of the 3D-printed mirror array is shown in Figure 10.



Figure 10. The 3D printed mirror array final design. The two pins can be turned to adjust the angle of the mirrors for calibration. A velcro strap can be wrapped around the smartphone, fed through the plastic loop shown at the top of the array, and folded onto itself to secure the smartphone.

This mirror array was designed in SOLIDWORKS; formfactor drawings are shown below in Figure 11.

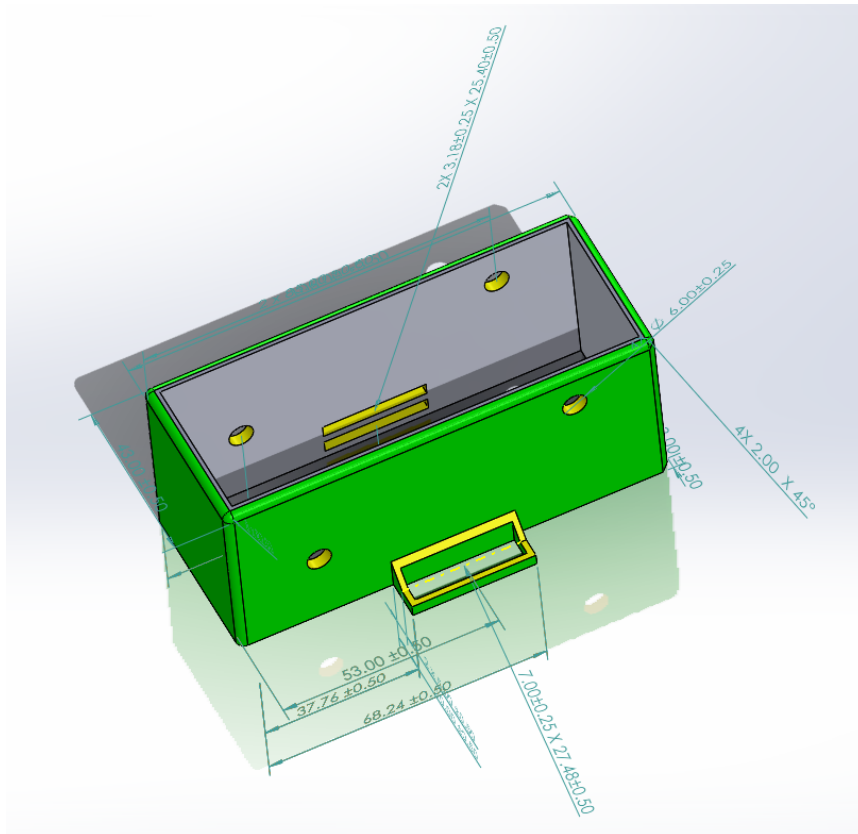


Figure 11. SOLIDWORKS formfactor model of the final design of the mirror array. All units are in mm.

6. Fabrication/Development Process

6.1 Overview of the Development Plan

Our planned first stage for this project involves image acquisition via smartphone camera with the mirror array attached. The video is transmitted to and processed by a laptop via Unity (Unity Software Inc.). Since Unity engine is specialized for VR experience rendering and game development, the team chooses it as the platform for video processing. At the end of this stage, video captured by the smartphone will be displayed on the laptop Unity project preview.

Currently we are transitioning to the second phase. By the end of this phase, we aim to finalize the Unity project with video acquisition and processing, and package the project into an APK file that can be installed on VR devices like Meta Quest 2 (Meta Platforms Inc.).

6.2 Video Acquisition

As is shown in the previous section, image acquisition involves the smartphone attachment and the smartphone itself. The user will first align the camera lens to the inner mirror (left non-reflected mirror in Figure 9), such that the mirror occupies the right half of the camera preview and forms the right visual channel. Meanwhile, the left half of the camera sensor directly captures an image of the same object, which forms the left visual channel.

Currently we are capturing real time video with a Motorola Edge 2021 smartphone (Motorola Mobility LLC) that runs on Android 12 (Google LLC). The team also uses an Android application DroidCamX Pro and its windows client, which uses wireless protocols to set up the smartphone camera as an external webcam to the laptop that runs on Windows 11 (Microsoft Corporation). DroidCamX Pro allows the laptop to use the smartphone as an external video capture device with up to 1080P resolution. The user can magnify or shrink the live video by pinching with fingers. It also features manual focusing with a finger tap. Additionally, video from smartphones can be streamed via USB wired connection or via WIFI as a wireless device. Since the smartphone is used as a wireless device in the local network and does not involve extra connection to routers and servers, we expect a significant reduction in delay for streaming video via DroidCamX, compared to the 0.21 second delay from streaming via Zoom last semester [10].

Using the default 3D core project template on Unity (version 2021.3.14f1), we keep the main scene with the default 3D template for directional light while adding a new camera object with OVR Camera Rig from the Oculus Integration Asset (version 46.0, developed by Reality Labs, formerly known as Oculus). Additionally, we add a Canvas layer with a CameraController object, a RawImage, and two button objects to the template.

CameraController object is a placeholder that holds a C# code “CameraScript,” which controls the behavior of buttons and displays the real time video to RawImage. For this project, we use C# language (version 9.0) that is used by Unity and code with VisualStudio (version 17.4.1).

When the program is running, the user can click the button “Start Camera,” CameraController turns on a camera hardware, changes the button text to “Stop Camera,” and renders the real time video as WebCamTexture onto the RawImage object. While the camera is

on, clicking the "Camera Switch" button searches for the next available video capture device and displays the real time video from it. At this time, clicking the "Stop Camera" will turn off the current video capture device, and the text on the button changes back to "Start Camera."

For demonstration purposes, we also add another overlay of VR frame, which helps determine the expected video output on the VR device for Stage 2. Figure 12 shows a screenshot of the user interface. Currently, the visual channels are not being separated, and when the canvas is brought to the eyes close enough on the VR device, it will still form a reasonable level of depth perception.

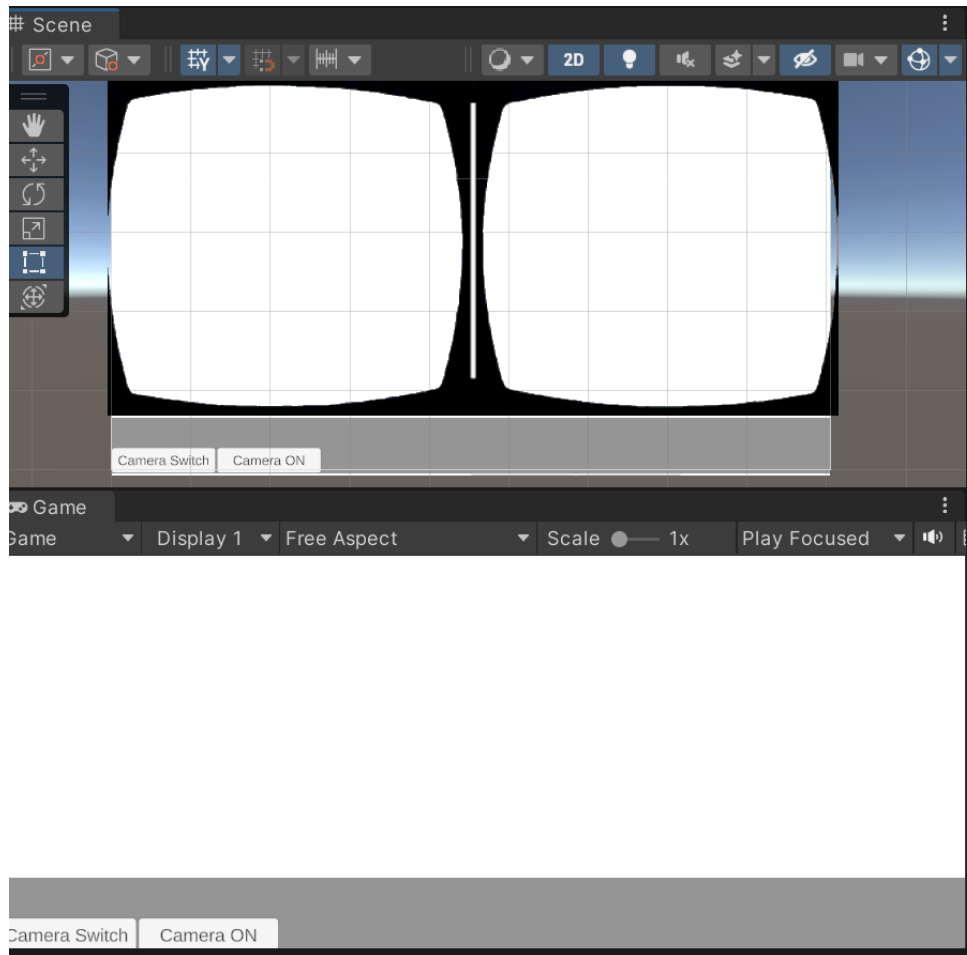


Figure 12. (Top) Expected display on a VR device. (Bottom) User interface for Unity project preview on PC. The buttons here are used to turn on/off and switch the video capture device connected to the PC. Both images obtained from Unity.

6.3 Video Processing

Because both visual channels are captured on the same display, we use the package OpenCV for Unity (version 2.4.9, developed by Enox Software). OpenCV (Open Source Computer Vision Library) is a software library is written in C/C++ language, which can be easily

incorporated into our C# codes [11]. It requires less memory allocation for the program, which is advantageous when running the program on RAM-limited devices like Meta Quest 2.

Using the setup of canvas from the Video Processing step as a template, a second RawImage is added to the canvas, which receives the right visual channel as a sub texture being cropped from the video source.

When the webcam is on, we first convert WebCamTexture to Mat (material) format for processing. The borderline between two visual channels is a vertical line at the middle of WebCamTexture, which requires calibration of the hardware to make the visual channels equally distributed on the screen. Meanwhile, since the right visual channel forms after reflection twice in the mirror array, there is distortion in the image. Instead of having linear transformation to adjust the width and height of the right visual channel, we will later adjust the angle and size of RawImage that host the right visual channel in Unity, thus reducing the amount of computing power that increases the delay in video streaming. For the left visual channel, it is rendered by the CameraControl script (see Appendix C), which also sends the right visual channel in Mat format to the CameraControlRight script (see Appendix D). CameraControlRight then displays the right visual channel on the second RawImage object.

Before modifying the Unity project, we tested the algorithm with python and casted real time stream directly from an external webcam (smartphone via USB connection to the PC). See Appendix E for python code used for testing.

6.4 Testing

Testing of the final design concept had two main goals. First, the team aimed to quantify the accuracy of the design in producing two images of the same object of interest from different perspectives. Second, of less importance, the team aimed to develop parameters that can be used to calibrate the device for translation to the VR environment.

6.4.1 Testing Methodologies

The final design was evaluated based on two parameters: horizontal skew and vertical skew. Horizontal skew was defined as the discrepancy in horizontal (left to right relative to the user) distance within each individual hemisphere. Vertical skew was defined as the discrepancy in vertical (top to bottom) distance of the object within each individual hemisphere. To quantify horizontal and vertical skew, a 1cm x 1cm square object of interest was drawn on graph paper, and the mirror array was focused at a distance of 25cm above the object of interest such that the full square object was fully captured in both hemispheres. A line was drawn at the division between the left and right visual channels. Using ImageJ software, the distance between the inside edge of the squares in both hemispheres was measured from the dividing line of the two hemispheres. A second square was used as a control, and appropriate parameters were calculated ipsilateral to the non-reflected image channel. Additionally, the distance between the tops of the two squares in the two hemispheres was measured and compared to a control square drawn at the same vertical position. For both horizontal and vertical skew, the area of the squares was

normalized by drawing a square identical to the square in the non-reflected visual channel and placing the traced square to be centered on the square in the reflected visual channel. A two-tailed T test was used to compare both horizontal and vertical skew for both groups.

Additionally, multiple trials were run to quantify the real-time lag of the streamed video. This involved streaming a video, captured by the smartphone through the mirror array, of a stopwatch on a computer screen. A second camera recorded a video with both the streaming computer's video output of the recorded computer stopwatch as well as the computer stopwatch itself. Quantification of the time delay was purely proof-of-concept, and no statistics were run.

6.4.2 Results

Statistical significance was achieved in the degree of horizontal ($p < 0.001$) and vertical ($p < 0.001$) skew relative to an ipsilateral control object (Figure 13).

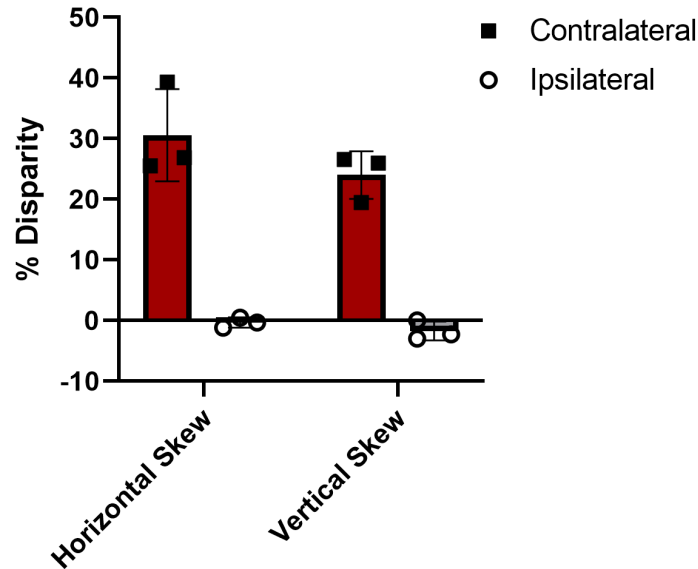


Figure 13. Quantification of the Horizontal and Vertical Skew between reflected and non-reflected hemispheres relative to the non-reflected hemisphere. Relevant controls include second squares drawn at the same horizontal or vertical position as the experimental square to control for vertical or horizontal skew, respectively, as a result of image compression or imperfections in the mirror. ($n=3,3$)

Testing on the algorithm with OpenCV and Python rendered a time delay of 0.23 +/- 0.01 seconds. This shows the relative time delay due to image processing. The delay of the video due to streaming was approximately 0.15 seconds -- this measurement is highly dependent on server capacity and network availability. The delay of the video equilibrated after about 10 seconds of use, and trials spanning >30 minutes showed consistent time delays with little variation. No lagging or frame-skipping was observed in any trials. An example cross-section of time delays over a 25 second stream is shown in Figure 14.

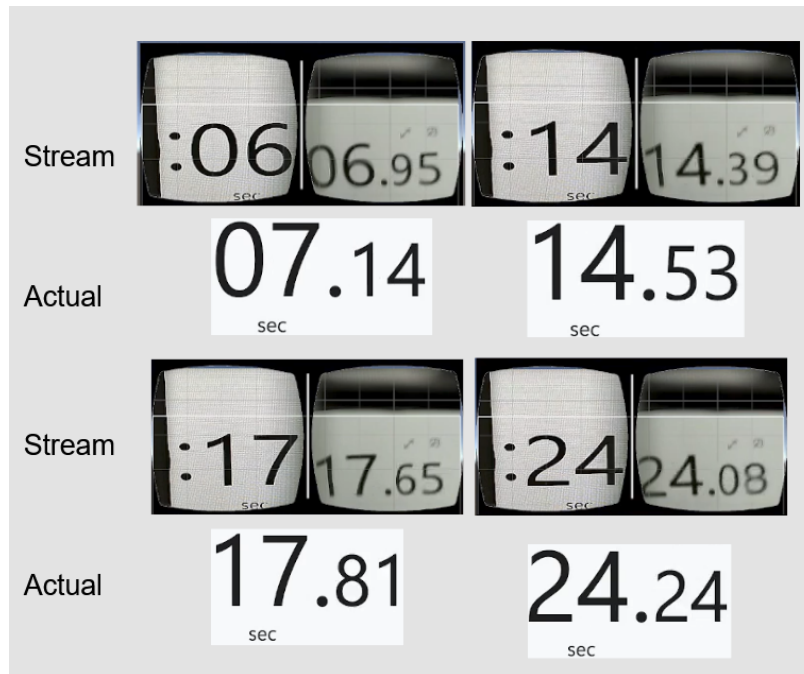


Figure 14. Time delay measurement for OpenCV on testing. The upper image shows separated left and right visual channels as they were displayed in Unity software, and the lower stopwatch shows the object being live-streamed.

6.4.3 Calibration Parameters

Calibration parameters were derived from experimentation to define the change in mirror angle to create images in which the object of interest is in roughly the same position in each visual channel with the same relative size (Figure 15). Notably, the calibration angle of the vertical direction cannot be easily changed due to the permanent fastening of the mirrors to the plastic mirror array, though the calibration angle of the horizontal direction can easily be adjusted in the current design concept.

Calibration Angles

$$\Delta\phi_x = F \times \arcsin(X_L - X_R)$$

$$\Delta\phi_y = F \times \arcsin(Y)$$

Normalization

$$\mathbf{A} = \begin{bmatrix} X_A & 0 \\ 0 & Y_A \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} X_B & 0 \\ 0 & Y_B \end{bmatrix}$$

$$\text{define } \mathbf{S} \in \mathbb{R} = \begin{bmatrix} \cos(\theta) & 0 \\ 0 & 1 \end{bmatrix}$$

$$\det(\mathbf{A}) = \det(\mathbf{B} \times \mathbf{S}) = \text{Area} = xy\cos(\theta)$$

Figure 15. Change in calibration angle needed to focus the mirrors to create centered objects of interest in each visual channel and matrix transformation to normalize the area of the object of interest in both hemispheres.

6.4.4 Preliminary work on VR Application

Informal experimentation was done to push live streamed videos captured by a smartphone to a Meta Quest 2 VR headset. Difficulties were observed in the ping of the smartphone, and further work needs to be done to allow the VR headset to recognize the compatible smartphone (Figure 16).

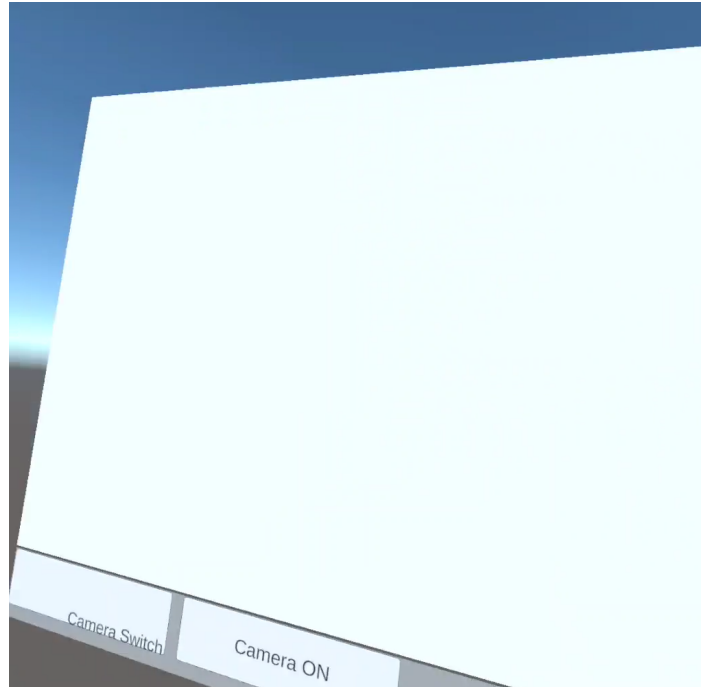


Figure 16. Program user interface running on Meta Quest 2. The image shown is projected into the virtual reality environment that the user sees while using the headset.

7. Discussion

The ongoing purpose of this project is to develop a device and corresponding software that can harness the camera recording and processing capabilities of a cell phone to replicate the function of a surgical microscope. Previous semesters have seen the development of a prototype of the mirror apparatus that can be used to refract the focal rays of the camera onto the object of interest. The primary goal of our team this semester was to generate working code that can perform the necessary manipulations to the images captured through this mirror apparatus such that a three-dimensional image can be perceived by the user. This involves some combination of a linear transformation to flatten the perceived inclined image planes, cropping of the two images acquired from the mirror apparatus, and stitching of the central regions of the images to create a coherent picture. Our proposed design presents a single image representing contributions from both hemispheres at an optimal lag time of less than 0.25s. While the current iteration of this design was shown to have issues with alignment, we have developed parameters to calibrate future versions of the design concept and account for innate inaccuracies in both the code and manufacturing of the mirror array. Current limitations include the processing power of the device used for image processing; it is crucial that steps are taken to minimize the run time of the code and simplify all computations. Governors may be needed to allow for the skipping of frames and image compression on intermittent occasions to avoid long-term frame lag in settings of limited server availability.

7.1 Future Work

Continuing on our previous work in developing these preliminary designs, we will begin by generating working code that can complete each of the three steps theorized in the second phase of image processing of our final preliminary design. This will include block coding and testing using trial by error. Once a final code capable of completing each step of our design is completed, testing will begin using some variation on examinations of speed that compare the processed images to unassisted hand-eye coordination. Tests involving suturing would provide the greatest external validity for our purposes and would serve as the best evidence towards the usefulness and practical application of this design concept in the classroom settings.

To achieve these goals, code must be developed that allows the VR device to access the webcam to perform real time processing of the live streamed video by the VR headset. This will require the use of the OpenCV package code to finalize the steps of visual channel separation. Preliminary work has shown the currently used package to be computationally expensive, and the updated package is more consolidated and better suited for our applications. Exploration is also planned for the coming semester on mass-appeal applications of the design concept, such as the development of a professional smartphone application.

Conclusions

Micro surgical techniques are an important part of modern medicine, especially as more research is conducted on the smaller compartments of organisms. Providing an affordable and accessible method of training surgeons in this practice is necessary to ensure availability of these procedures around the world. The high cost of professional equipment highlights the importance of a design that can offer an effective alternative. The main challenge of this is the production of a prototype that can give the user a comparable sensation of depth perception at a low latency. A combination of techniques used by previous groups along with a heavier software end seems to offer the best possibility of producing this.

Our group will work on manipulating the mirror distance and angle with respect to the phone camera to produce a high resolution image. The mirror mount will be modified to give the user a sense of stereoscopic vision similar to natural human perception. Lastly, we will develop a software method of processing the image to create a seamless combined view.

References

- [1] A. Beris, I. Kostas-Agnantis, I. Gkiatas, D. Gatsios, D. Fotiadis, and A. Korompilias, “Microsurgery training: A combined educational program,” vol. 51, pp. S131–S134, 2020, doi: <https://doi.org/10.1016/j.injury.2020.03.016>.
- [2] K. Uluç, G. C. Kujoth, and M. K. Başkaya, “Operating microscopes: past, present, and future,” vol. 27, no. 3, p. E4, 2009, doi: 10.3171/2009.6.FOCUS09120
- [3] Olympus, “Revolutionize your surgical microscopy - Olympus,” Olympus - Orbeye. [Online]. Available: https://www.olympus-europa.com/medical/rmt/media/en/Content/Content-MSD/Images/SRP-Pages/SRP-ORBEYE/ORBEYE_concept_brochure_53297.pdf.
- [4] D. Grady, “Orbeye 3-D system turns surgery into dizzying, immersive experience,” *The Columbus Dispatch*, 12-Jan-2018. [Online]. Available: <https://www.dispatch.com/news/20180114/orbeye-3-d-system-turns-surgery-into-dizzying-immersive-experience#:~:text=Mark%20Miller%2C%20a%20spokesman%20for,cost%20about%20%24400%2C000%2C%20Langer%20said>.
- [5] “3D Types.” http://www.etc.cmu.edu/projects/lm-3d/3D_Types.html (accessed Oct. 11, 2022).
- [6] B. M. A. A; “The surgical suture,” *Aesthetic surgery journal*, Apr-2019. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/30869751/>. [Accessed: 24-Sep-2022].
- [7] R. Rotermund, J. Regelsberger, K. Osterhage, J. Aberle, and J. Flitsch, “4K 3-dimensional video microscope system (orbeye) for transsphenoidal pituitary surgery,” *Acta Neurochir*, vol. 163, no. 8, pp. 2097–2106, Aug. 2021, doi: 10.1007/s00701-021-04762-x.
- [8] “ORBEYE,” Olympus America | Medical. <https://medical.olympusamerica.com/products/orbeye> (accessed Oct. 11, 2022).
- [9] “HoloLens 2—Overview, Features, and Specs | Microsoft HoloLens.” <https://www.microsoft.com/en-us/hololens/hardware> (accessed Oct. 11, 2022).
- [10] Zoom, “Overview core concepts and components - zoom,” *Zoom Connection Process Whitepaper*. [Online]. Available:

https://explore.zoom.us/docs/doc/Zoom_Client_Connection%20Process_Whitepaper.pdf.
[Accessed: 13-Dec-2022].

[11] “About,” *OpenCV*, 04-Nov-2020. [Online]. Available: <https://opencv.org/about/>. [Accessed: 12-Dec-2022].

Appendix

A. Product Design Specification

Product Design Specification (PDS)

Title: iPhone Virtual Reality Training Model for Microsurgical Practice.

Sam Neuman, Alex Vazquez, Haochen Wang

September 23th, 2022

Function:

This training model will make microsurgical training less expensive and more accessible to a wide range of users. It eliminates the need for an expensive surgical microscope by replacing it with a smartphone. The prototype will utilize the zoom functionality of the smartphone for the surgeon to clearly see sutures and tissues up close. By using a smartphone, it is also possible to stream the training to Zoom or a similar platform so training can occur virtually. The design will minimize lag time between the recording phone and projecting device for simultaneous view of both the trainee and observers, while increasing spatial awareness and depth perception via binocular live video.

Client requirements:

- Must allow for depth perception with regard to where the trainee's hands are in relation to the work site.
- Must create an image with high enough zoom and resolution to see sutures (0.070 mm in diameter) clearly [1]
- Must remain inexpensive so it is widely accessible to training surgeons
- Must produce a streaming resolution of at least 10.2 megapixels
- Must have a frame rate of at least 24 frames per second
- Must have a stream delay of no more than 0.5 seconds
- Should utilize full magnification power of the smartphone

Design requirements:

1. Physical and Operational Characteristics

a. Performance requirements:

- i. The device must be able to provide a clear image of the subject in a clinical environment. The device must be able to handle daily use and must be able to handle a load of at least 400g, the weight of the heaviest available smartphones.

b. Safety:

- i. The device should be out of the way of the surgeon to prevent interference during practice and contact with hazardous material encountered during the simulated surgery. The device also needs to be able to be sterilized in an efficient manner before and after each use.

c. Accuracy and Reliability:

- i. The device should be able to consistently maintain a magnification of 2x and the displayed magnification should be accurate with repeated trials. The device should display an accurate and clear image of the surgery area with minimal latency.

d. Life in Service:

- i. The device should withstand continued use over the duration of the training process, the longest of which can last up to 12 hours. The device should be able to withstand this use everyday over its lifespan, as many different trainees may use the device.

e. Shelf Life:

- i. The device should be stored in normal interior conditions for an indefinite amount of time. This means that the device should not be made out of materials that degrade over a short period of time (6 months, the duration of one semester) in absence of normal use. With continued use, the team would expect the smartphone being the limiting factor for the whole design. Thus, the final deliverable should have at least one year of lifespan, which matches the lithium battery warranty provided by Apple. [2]

f. Operating Environment:

- i. The product will most likely be used in a domestic or indoor environment, so the device will not be exposed to extreme conditions.
- ii. 0-35 ° C operating temperature, - 20-45 ° C nonoperating temperature, 5-95% non-condensing, relative humidity (the specification of iPhone 8, and more restriction may be applied as other hardware is introduced to the final deliverable) [3]
- iii. The person who will use this will be the trainee, which is the person who is practicing surgery using the iphone, and the trainer(s) who is/are watching the trainee on the headset.
- iv. Potential splash of food dye, blood, in vitro tissues, etc. [4]
- v. Components that are exposed to the operation station shall not be malfunctioned upon such splash
- vi. Potential scratches from the surgical equipment, such as tweezers or needles.
- vii. The final deliverable should at least endure accidental damage from the aforementioned scenarios, while maintaining the resolution to recognize the suture

g. Ergonomics:

- i. The product can involve delicate technology, such as smart phones and laptops, so the same restrictions of force that cause those devices not to be damaged or break apply here.
- ii. For the iPhone 8, do not submerge in water greater than 1 meter and for longer than 30 minutes. [3]

h. Size:

- i. Should be able to be set up in an indoor living space (i.e. 10 x 10 sqft, approximately 3 x 3 meters)

i. Weight:

- i. Optimum weight: < 10lbs (approximately 4.5 kg). Must be easily transportable

j. Materials:

- i. No restrictions on material mechanics
- ii. Cannot be toxic upon skin contact or inhalation
- iii. Shall have minimal degradation resistance, such as from sunlight

k. Aesthetics, Appearance, and Finish:

- i. The color of the product should be dull so that it doesn't distract from the microsurgical practice it is intended for. The shape and form should be adjustable so that each user/consumer can place it into alternate positions to get a better and more comfortable practice for themselves. The texture of the finish should be flat and soft in order for it to be comfortable for the user and in order for it to not be a distraction.
- ii. Should simulate the working condition of an operation room with microscopes
- iii. Must not interfere with the operation and training performance of the user

2. Production Characteristics

a. Quantity:

- i. Tens of Thousands of units will be needed so that this can replace all current expensive training mechanisms for microsurgical practice for medical residents.

b. Target Product Cost:

- i. The target cost of the product is undetermined thus far until clients discuss but it will need to allow for an iPhone, a stand, and any attachment that is necessary to put over the camera to replicate microsurgery practice as best as possible. There are existing products whose costs are at least \$100,000 [5] which is drastically greater than the target cost. The prototype is a cheap alternative for medical students to use for remote training, using materials that are commonly owned.

3. Miscellaneous

c. Standards and Specifications:

- i. ISO 10936-1:2017
 - Specifies the requirements for microscopes used during surgical procedures, so the team must adhere to these specifications when creating a design. However, since this prototype will be used for practice purposes, the requirements may not all apply [6]
- ii. Code of Federal Regulations Title 21, Volume 8, Sec. 882.4525 Microsurgical instrument [7]
 - The final deliverable will fall into the Class I medical device category, which is exempt from the premarket notification procedures 510(k)
- iii. Code of Federal Regulations Title 21, Volume 8, Sec. 878.4700 Surgical microscope and accessories [8]
 - The final deliverable, under definition of this section, will be a Class I device. However, since the recording device in this design will be a DC powered smartphone, no more actions shall be made upon this regulation

d. Customer:

- i. The customer would prefer the delay of relaying the image to the headset to be minimized for enhance practicing technique (less than 0.5 s)
- ii. The quality of the camera while zooming should be clear enough to clearly see the material being worked upon. 2x zoom using an iPhone 11 Pro was tested to be the most practical. The requirement is that the trainee is able to see the suture, which is 0.070 mm [1]
- iii. The camera should be able to show the depth of the workspace in order to help determine the distance between the instruments being utilized and the suture on the workbench. This may require the use of two lenses to allow for a binocular view
- iv. The device should be comfortable to wear for extended periods of time

e. User-related concerns:

- i. As this is a device used for practice, there will be no requirements for patient confidentiality.
- ii. Sterilization should not be an issue with regard to the camera setup. However, it may be practical to clean the headset with a wipe between uses.
- iii. The design should be able to receive accommodations for users with visual impairments.

f. Competition:

- i. Augmented Reality (Mixed Reality)
The Microsoft HoloLens is a very complex device which allows for similar types of practice.

However, the Hololens is much less accessible and much more expensive. This will be an alternative that is possible to use from many different remote locations. Meanwhile, mixed reality provided by Hololens is rather redundant for the purpose of the clients. [9]

ii. Exoscopic Platforms

Zeiss, Olympus and Mitaka are well known medical device providers for exoscopes, featuring high definition images of the field with 8x to 30x magnifying capability. However, the price varies from 0.2 to 1.5 million dollars, resulting in limited access for trainees from less developed regions [5].

Reference

- [1] B. M. A. A; “The surgical suture,” *Aesthetic surgery journal*, Apr-2019. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/30869751/>. [Accessed: 18-Sep-2022].
- [2] Apple, “iPhone battery and performance,” *Apple Support*, 24-Sep-2021. [Online]. Available: <https://support.apple.com/en-us/HT208387>. [Accessed: 19-Sep-2022].
- [3] Apple, “iPhone 8 - Technical Specifications,” Official Apple Support, 25-Mar-2021. [Online]. Available: https://support.apple.com/kb/sp767?locale=en_US. [Accessed: 22-Sep-2022].
- [4] “Dr. Sam POORE and team featured ON BTN LiveBIG Wisconsin,” Department of Surgery, 13-Jul-2020. [Online]. Available: <https://www.surgery.wisc.edu/2020/07/13/dr-sam-poore-and-team-featured-on-btn-livebig-wisconsin/>. [Accessed: 20-Sep-2022].
- [5] D. J. Langer, T. G. White, M. Schulder, J. A. Boockvar, M. Labib, and M. T. Lawton, “Advances in INTRAOPERATIVE Optics: A brief review of Current Exoscope platforms,” *Operative Neurosurgery*, vol. 19, no. 1, pp. 84–93, 2019.
- [6] “ISO 10936-1:2017,” ISO, 26-Jul-2017. [Online]. Available: <https://www.iso.org/standard/69819.html>. [Accessed: 18-Sep-2022].
- [7] CFR - Code of Federal Regulations Title 21 Sec. 882.4525, vol. 8. 2020.
- [8] CFR - Code of Federal Regulations Title 21 Sec. 878.4700, vol. 8. 2020.
- [9] “Microsoft mixed reality – Healthcare,” Microsoft Mixed Reality – Healthcare. [Online]. Available: <https://www.microsoft.com/en-us/hololens/industry-healthcare>. [Accessed: 17-Sep-2022].

B. CameraScript.cs used in Image Acquisition Phase

using System.Collections;

using System.Collections.Generic;

using TMPro;

using UnityEngine;

using UnityEngine.UI;

public class CameraScript : MonoBehaviour

{

int currentCamIndex = 1;

//int frameWidth;

private static WebCamTexture tex;

```

public RawImage display;

public TextMeshProUGUI startStopText;

public void SwapCam_Clicked()
{
    if (WebCamTexture.devices.Length > 0)
    {
        currentCamIndex += 1;
        currentCamIndex %= WebCamTexture.devices.Length;

        // If tex is not null
        // Stop the webcam and start new cam
        if (tex != null)
        {
            StopWebcam();
            StartStopCam_Clicked();
        }
    }
}

public void StartStopCam_Clicked()
{
    if (tex != null) // Stop the camera
    {
        StopWebcam();
        startStopText.text = "Start Camera";
    }
    else // Start the camera
    {
        WebCamDevice device = WebCamTexture.devices[currentCamIndex];
        tex = new WebCamTexture(device.name);
        // frameWidth = tex.requestedWidth / 2;
        display.texture = tex;

        tex.Play();
        startStopText.text = "Stop Camera";
    }
}

public void StopWebcam()
{

```

```

    display.texture = null;
    tex.Stop();
    tex = null;
}

public static WebCamTexture Tex
{
    get { return tex; }
    set { tex = value; }
}
}

```

C. CameraControl.cs to be used in Video Processing

```

using System.Collections;
using System.Collections.Generic;
using TMLPro;
using UnityEngine;
using UnityEngine.UI;
using OpenCvSharp;
using System;

public class CameraControl : MonoBehaviour
{
    public static int frameHeight;
    public static int frameWidth;
    int currentCamIndex = 1;
    public static WebCamTexture tex;
    public RawImage displayL;
    //public RawImage displayR;
    public TextMeshProUGUI startStopText;
    public Mat imageL;
    public Mat imageR;
    public Mat image;
    // private CameraControlRight controlRight = new CameraControlRight();

    public void SwapCam_Clicked()
    {
        if (WebCamTexture.devices.Length > 0)
        {
            currentCamIndex += 1;
            currentCamIndex %= WebCamTexture.devices.Length;

            // If tex is not null

```

```

// Stop the webcam and start new cam
if (tex != null)
{
    StopWebcam();
    StartStopCam_Clicked();
}
}
}

public void StartStopCam_Clicked()
{
    if (tex != null) // Stop the camera
    {
        StopWebcam();
        startStopText.text = "Start Camera";
    }
    else // Start the camera
    {
        WebCamDevice device = WebCamTexture.devices[currentCamIndex];
        tex = new WebCamTexture(device.name);
        image = OpenCvSharp.Unity.TextureToMat(tex);
        if (frameWidth == 0 || frameHeight == 0)
        {
            frameWidth = tex.requestedWidth;
            frameHeight = tex.requestedHeight;
        }
        imageL = image.SubMat(0, frameHeight, 0, frameWidth / 2);
        imageR = image.SubMat(0, frameHeight, frameWidth / 2, frameWidth);

        displayL.texture = OpenCvSharp.Unity.MatToTexture(imageL);
        // displayR.texture = OpenCvSharp.Unity.MatToTexture(imageR);

        //CameraControlRight.CameraShowRight(this.imageR);
        startStopText.text = "Stop Camera";
    }
}

public void StopWebcam()
{
    displayL.texture = null;
    //CameraControlRight.displayR.texture = null;
    // displayR.texture = null;
    tex.Stop();
}

```

```

    tex = null;
}
}

```

D. CameraControlRight.cs to be used in Video Processing

```

using System.Collections;
using System.Collections.Generic;
using TMPro;
using UnityEngine;
using UnityEngine.UI;
using OpenCvSharp;

public class CameraControlRight : CameraControl
{
    public RawImage displayR;

    public void CameraShowRight()
    {
        displayR.texture = OpenCvSharp.Unity.MatToTexture(imageR);
    }
}

```

E. Python code with OpenCV

```

import cv2
import sys
import numpy as np
import time

# implement time tracking function
_start_time = time.time()

def tic():
    global _start_time
    _start_time = time.time()

def tac():
    t_sec = round(time.time() - _start_time)
    (t_min, t_sec) = divmod(t_sec,60)
    (t_hour,t_min) = divmod(t_min,60)
    print("Time passed: {}hour: {}min: {}sec'.format(t_hour,t_min,t_sec))

# read the video from specified path

```

```

cap = cv2.VideoCapture(1)

while(cap.isOpened()):
    # tic();
    # extract the frames
    ret, frame = cap.read()

    if frame is not None:
        # read the size of the image, dimensions[2] is channel
        dimensions = frame.shape
        height = dimensions[0]
        width = dimensions[1]
        cut = int(width/2)

        # crop the frame
        leftView = frame[0:height, 0:cut]
        rightView = frame[0:height, cut:width]

        cv2.imshow('leftView',leftView)
        cv2.imshow('rightView',rightView)

        # tac();

    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

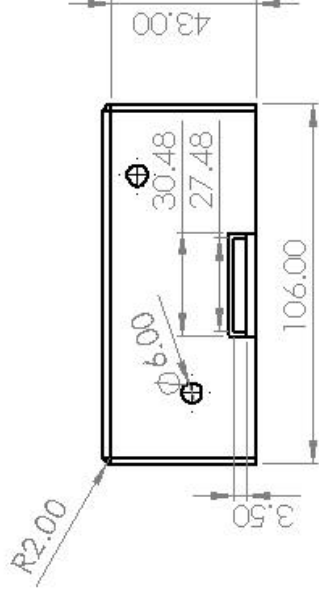
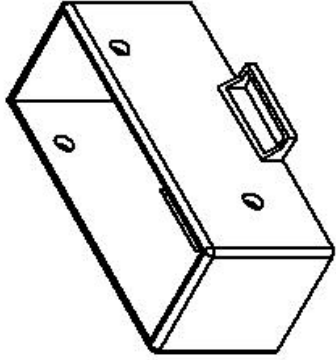
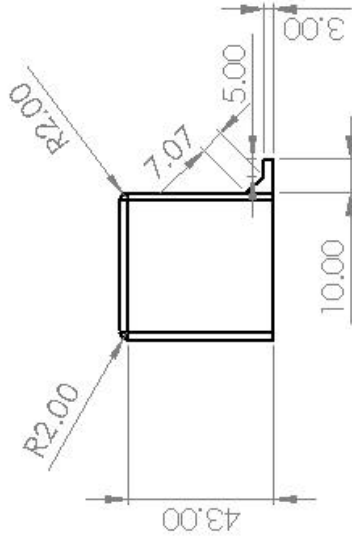
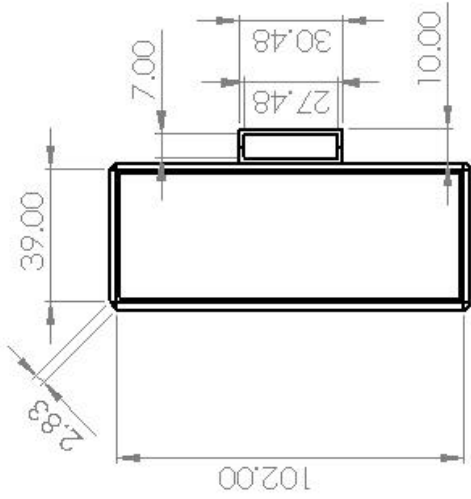
cap.release()
cv2.destroyAllWindows()

```

F. CAD Drawing for the Smartphone Attachment

Figures below show the CAD drawing for the mirror housing (top) and for two pins (bottom)

6 5 4 3 2 1



D C B A

UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN MILLIMETERS		FINISH:		DEBUR AND BREAK SHARP EDGES		DO NOT SCALE DRAWING		REVISION	
SURFACE FINISH:		TOLERANCES:		NAME		SIGNATURE		DATE	
LINEAR:		ANGULAR:		DRAWN		CHKD		APPRD	
				MFG		QA		MATERIAL:	
								WEIGHT:	
								TITLE:	
								DWG NO.	
								SHEET 1 OF 1	
								SCALE: 1:1	
								Mirror Housing V2 ^{A4}	

