iPhone Virtual Reality Training Model for Microsurgical Practice

BME 400
University of Wisconsin-Madison
Department of Biomedical Engineering
10/12/2022

Team Members:
Samuel Neuman - *Team Leader & Co-BWIG*
Alexander Vazquez - *Communicator & BPAG*
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 its 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, during the COVID-19 pandemics, there is a rising need for remote instruction. As smartphone cameras can provide high enough resolution and magnification on the surgery station, this project aims to develop a microsurgery training solution based on common smartphone models, while forming two visual channels to enable depth perception from the display. The team evaluates two proposed approaches to combine a hardware design and a software design from previous teams, including Single Frame and Combined Frame designs, and the Single Frame Design becomes the proposed final design for this team. Each of these designs utilizes multiple technologies to create the perception of depth of the object being viewed. The team ultimately considers a hybrid of the two designs to pursue that maximizes usability and minimizes the potential for constraint by hardware capabilities.

Table of Contents

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 future patients of the trainees. 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, more affordable while equivalent effective training method as substitutes to the microscopes. Cameras on smartphones can now provide sufficient magnification and high resolution for live streaming, and the clients see potential of utilizing the smartphone cameras to fulfill this commercial gap. However, as microsurgery requires binocular vision to enable depth perception of the microsurgeon, while most smartphone models can only activate one camera lens at a time, 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. And 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 surrounding environment when performing microsurgery training, and 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 magnficiations, 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 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/transfering data and the ability to interface with various types of software makes it 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 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].

2.4 Clients

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

2.5 Previous Work

2.5.1 Previous BME Team and Clients

This project first started as BME 400/402 capstone project in Fall 2020 under the advice of Dr. Willis Tompkins. Team members were Jason Wang, Xiaoxuan Ren, Martin Janiszewski, Jiong Chen, and Kartarina Milosavljevic.

In its second year, the project was run by a combination of BME 200/300 students in Fall 2021, including three junior students (Henry Plamondon, Nicholas Jacobson and Haochen Wang) and three sophomore students (Mitchell Benyukhis, Kenzie Germanson and Cameron Dimino). The project was under guidance of Prof. Walter Block. The juniors, joining Emma Kupitz and Frida Albiter Benitez, continued this project for their BME 301 course in Spring 2022 under guidance of Dr. Alireza Ashtiani and Dr. Aviad Hai.

Dr. Poore and Dr. Shaffrey were clients for both years, and Dr. Weifeng Zeng joined the client team during the second year.

2.5.2 Work From Previous Semester

Software Design

The first team delivered a software design via MATLAB. They used pre-recorded videos to test the functionality of the design. As shown in the Flowchart in Figure 2, the MATLAB script first loads the video from the address specified by the user. Then frames of the video are extracted. For each frame, the script duplicates it and performs parallel processing on the duplicates. Simultaneously, each copy of the frame is rotated 5 degrees away from the central axis, forming left and right views for later processing. The frames are then cropped and magnified to maintain its original width-height ratio, after which a red- and a cyan-colored anaglyph filter are added to left and right views respectively. The script then combines these filtered views, sending the final output for display. Microsurgery trainees will wear a pair of anaglyph 3D goggles (with red and blue lens for each eye) to watch the display. While each of the lenses can only transmit one of the views from the combined image, this design is expected to provide depth perception via the dual visual channels in display.

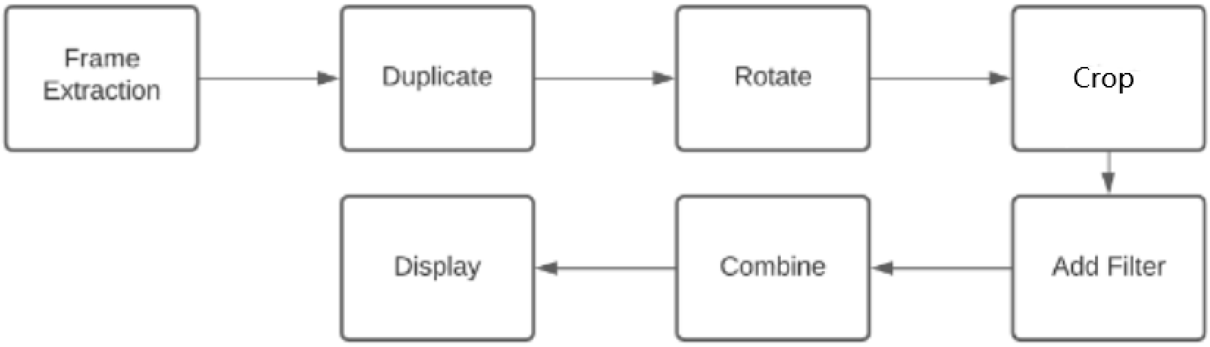


Figure 2: Flowchart for software design developed by the first BME design team. Image extracted from video capture will first get duplicated and rotated by 5 degrees. Then the duplicated images will be cropped to maintain the size ratio of the video source. Different filters (red or cyan) are added to each duplicated image, and the combination of filtered images is the output for display.

In addition to the MATLAB script for video processing, that team also developed a camera control app (CameraAccess) for the ios system in swift language. The application enables zooming and autofocus functions. Figure 3 shows a sample image processed by the MATLAB script that is captured with the CameraAccess app. The first team also reported that the MATLAB script was able to produce output video at 46.58 frames per second (fps) for a resolution of 320 x 240 pixels. Yet, when the resolution was raised to 1280 x 720 pixels, the processing speed dropped to 11 fps.

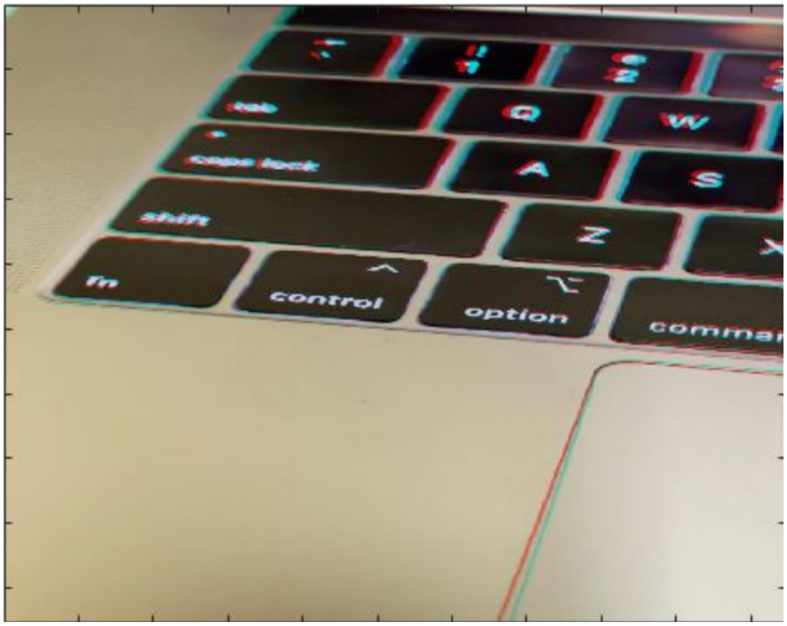


Figure 3: A sample image captured with CameraAccess app and processed with the software developed in MATLAB. Blurry image is a result of combined views with two different filters added. [Image adopted from final report of BME 400 team in Fall 2020].

Hardware Design

The second team appreciated the anaglyph filter design from the first team. However, high latency between an surgical operation and the display can arise from low processing speed and low fps in video output. Meanwhile, the team expressed concerns about the arbitrary setting on rotating angle to produce left and right views. Thus, aiming to reduce the runtime complexity of the software and to provide dynamic view separation, the team designed a portable smartphone attachment with two mirrors inside. As is shown in Figure 4, the image of the surgical station or tissue (represented by a red arrow pointing upward, which is placed on the right but not shown in the figure) first gets reflected onto the bottom half of the camera sensor by the outer mirror and the inner mirror. Meanwhile, the image can also directly project onto the top half of the sensor, effectively forming a second visual channel that is of a different viewing angle from the reflected view.

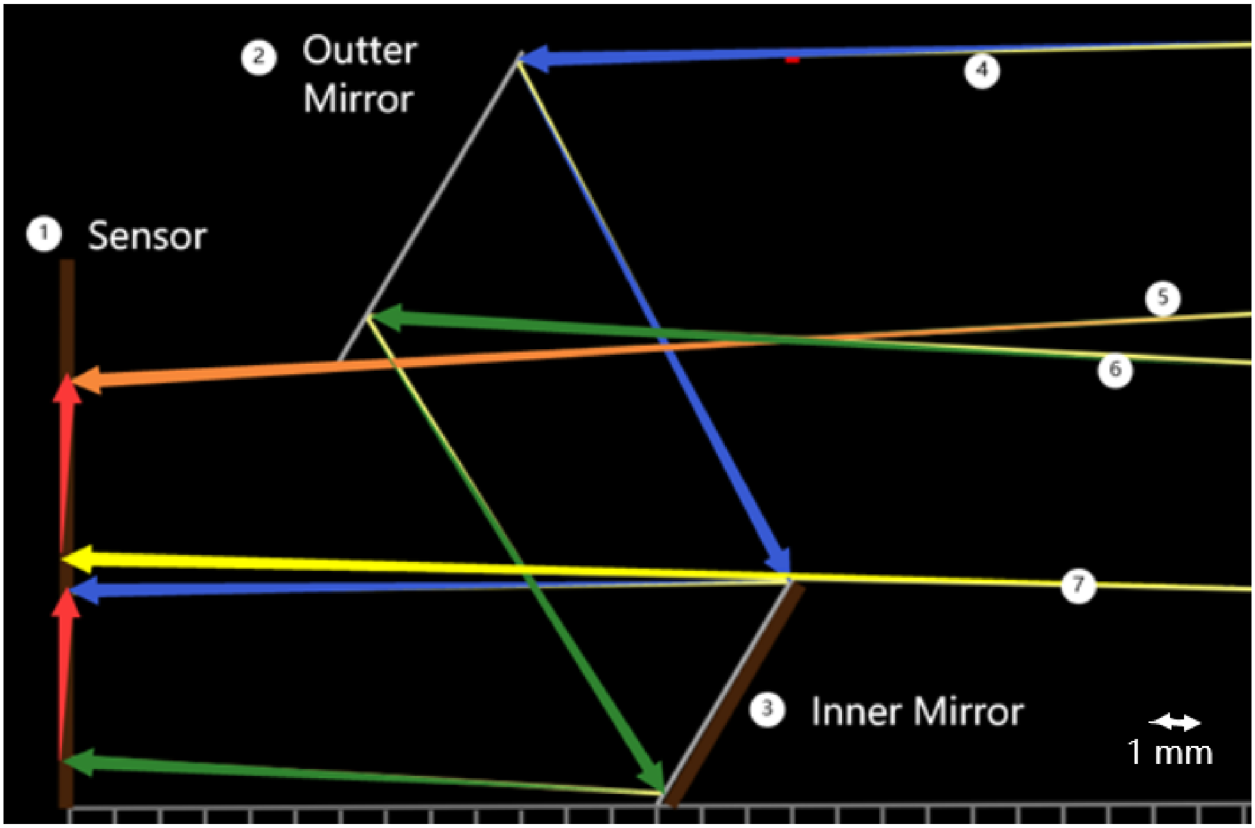


Figure 4: Ray tracing diagram for the hardware design. Light emitted by the object (a red arrow pointing upward, not shown in the figure) first gets reflected by the outer mirror, then by the inter mirror, onto the bottom half of the camera sensor. Meanwhile, another visual channel is formed by light directly emitted onto the top half of the camera sensor. Note that the area of the camera sensor is not fully utilized, and there is a vacancy between the bottom of the direct view (ray in yellow, 7) and the top of the reflected view (ray in blue, 4). Figure is drawn in scale. [Image created by Haochen, adopted from final report of BME 301 Spring 2022].

Also observe that in figure 4, though not represented with high accuracy, the surface area of the camera sensor is not fully utilized, and there is a gap between the visual channels. Both flaws were observed in the sample image provided by the team, as is shown in Figure 5. Due to misalignment of the mirror, there was misalignment and uneven distribution between the views. Therefore, though there was improvement in depth perception compared to smartphone cameras alone, it was still not significant enough to perform suturing practice with 7-0 sutures. Additionally, since the video was captured by one phone and streamed to another via zoom, there was noticeable display latency.

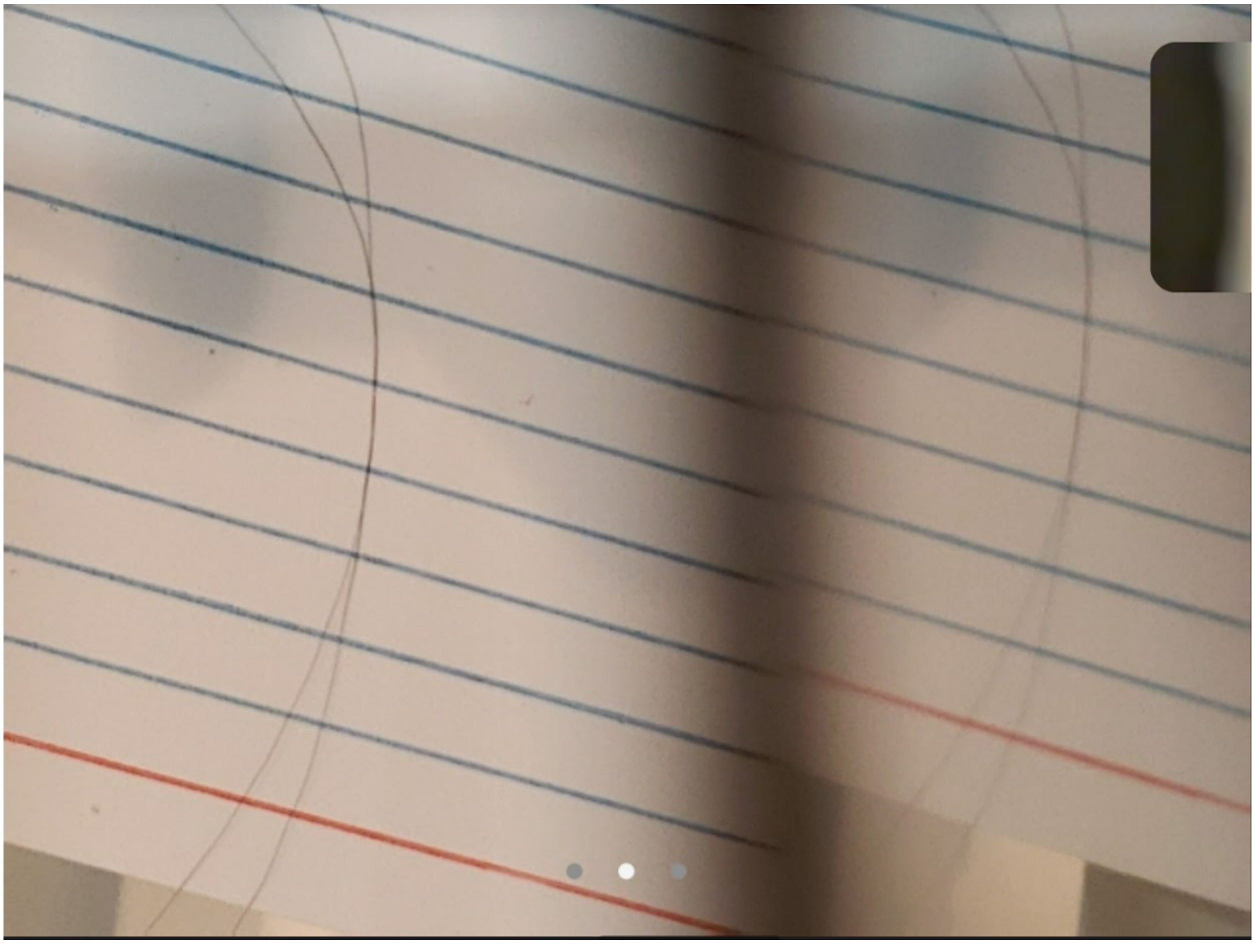


Figure 5: Screenshot from Zoom meeting during testing. The notecard with a piece of 7-0 suture was first captured with one camera, then the view was streamed to another smartphone inserted in Google Cardboard, a housing that allows combination of the views from the display. Observe that there is a gap, misalignment and even distribution between the left and the right view. [Screenshot obtained by Haochen from Spring 2022].

2.6 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.

2.7 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 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 6: The Olympus Orbeye 4k Video Microscope

The next alternative design that was found was the Microsoft Hololens 2. 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].

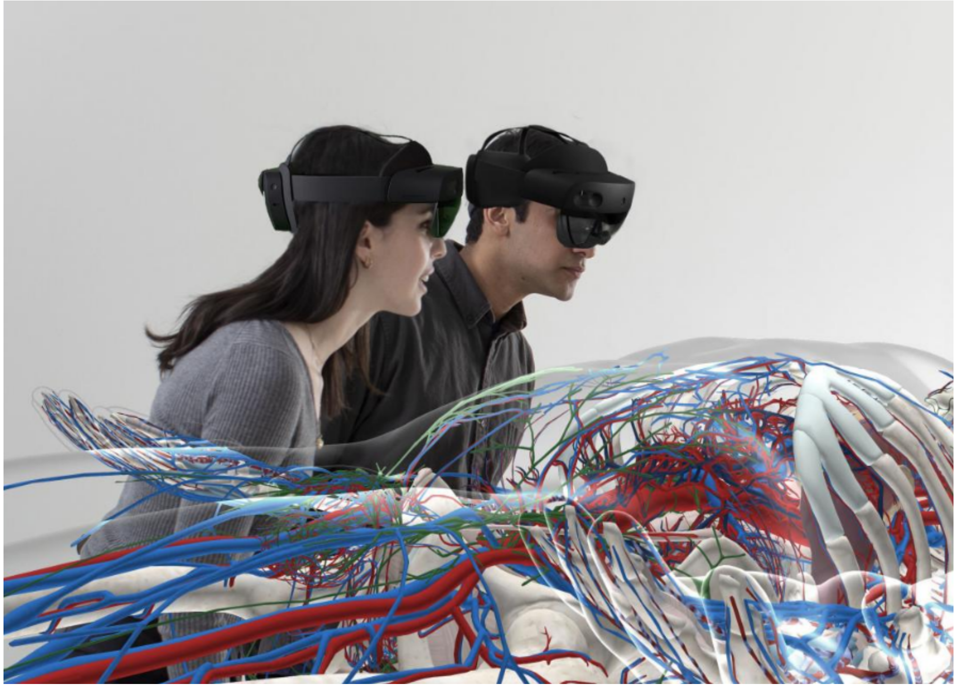


Figure 7: The Microsoft Hololens 2

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 8.

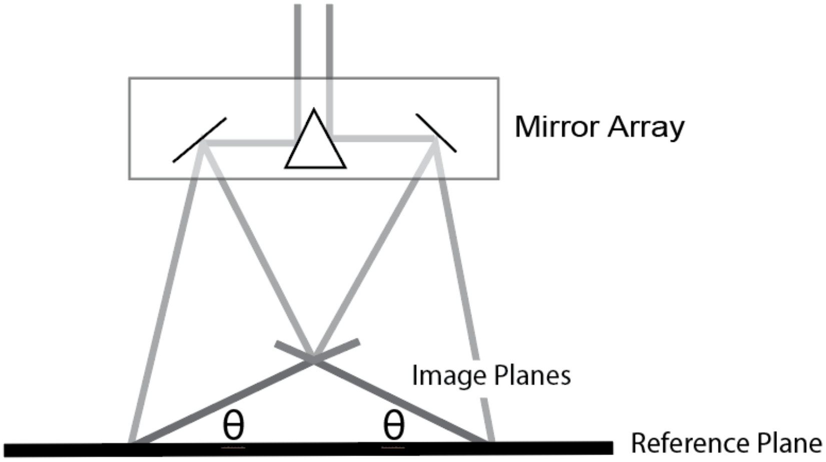


Figure 8. The geometry of the inclined image planes created by the mirror array. The offset of the mirrors in the mirror array will be set such that the camera focuses only at a fixed distance from the mirror array. With this known distance, the angle θ can be derived as the angle between the two image 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.

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

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 9. The full application of this code took roughly 30 seconds to process the two 1080p images, suggesting possible bottlenecks in future working codes.

Figure 9. 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 10.

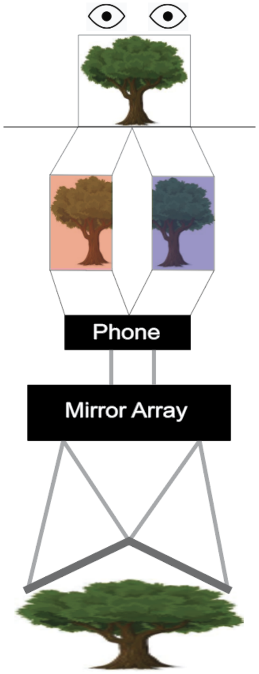


Figure 10. The schematic of the single frame view. First, the camera acquires the images of the two inclined planes of the object. The phone first applies the linear transformation to flatten the image and crops out the overlapping regions of the two halves of the image. Finally, the phone applies an 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 11.

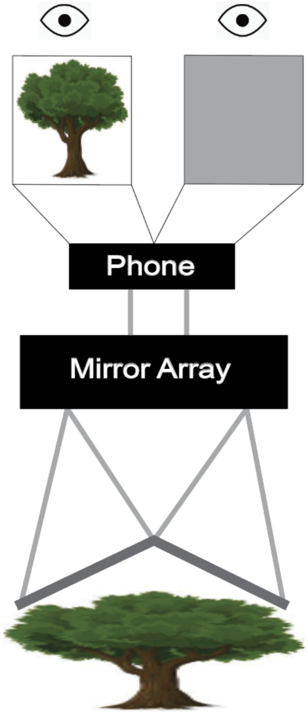


Figure 11. The combined view preliminary design. Both hemispheres of the images will be acquired simultaneously and cropped by the software to contain the same landmarks and thus be the same shape and encompass the same features. The phone screen will display only one of the left or right images to both eyes at one time in rapid succession, creating the illusion of each eye contributing its perspective to resolve a three dimensional image.

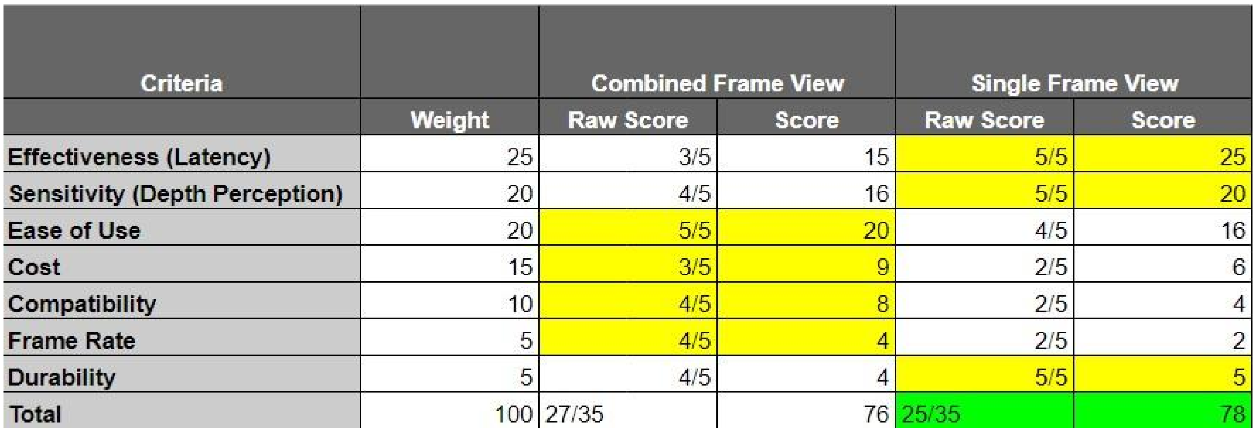
Even more than the single frame view design, hardware capabilities will be a significant constraint on this design. Operating at a frame rate of 120 fps is optimistic, albeit in excess of the specifications of most computer monitors and phone screens. Otherwise, no accelerations of the computations will 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		Combined Frame View		Single Frame View	
	Weight	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 prognosticate 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.

4.3 Proposed Final Design

The proposed final design will be a hybrid of the two preliminary designs. We will incorporate the linear transformation and stitching of the single frame view design to create a single frame encompassing the context of the imaging field. No anaglyph filter will be applied, as the loss of color resolution is a limiting factor to translation of this technology to its applications. The single frame view also minimizes concerns of the framerate capabilities of the screen being used to image, though it is likely that the computing components of the device will not be able to handle such computations. Thus, the final design will likely incorporate code to skip frames when needed to prevent lag accumulation.

5. Fabrication/Development Process

5.1 Methods

All coding will be performed in Python using the Jupiter IDE for block running. Python is a ubiquitous programming language that allows for easy conversion to C++ and Java for iphone/android application development. Proof of concept code may be written in MATLAB for ease of access.

5.2 Testing

Testing will be performed on a qualitative basis until a final design is achieved. This will involve multiple interactions of coding a piece of the project, testing it with either single frames or short videos, and making minor adjustments to the code. Once a final draft of the code is written, tactile tests will be performed by either ourselves, a lay third party, or persons with some degree of medical training. These tests will involve some variation on a speed test for a basic suture both with our image processing software and without our image processing software. A third trial utilizing the clients' microscope would be a suitable positive control if it is possible this semester.

Discussion

The purpose of this project is to develop a device and corresponding software that can harness the camera and processing capabilities of a cell phone to replicate the function of a surgical microscope. Previous semesters have developed 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 is 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 will present a single image representing contributions from both hemispheres at an optimal frame rate of 60 fps. 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 will be needed to allow for the skipping of frames and image compression on intermittent occasions to avoid long-term frame lag.

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 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.

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 producing 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).

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 many 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].